Introduction

Osteoporosis is a skeletal disorder of low bone mass and microarchitectural deterioration of bone, leading to bone fragility and increased risk of fracture.\(^1\) Osteoporosis affects more than 10 million U.S. adults aged 50 years or older.\(^2\) About two million U.S. adults experience an osteoporotic or other low-or no-trauma fracture each year.\(^3\) These fractures frequently cause pain, disability, and impaired quality of life;\(^4,5\) and hip and clinical vertebral fractures, specifically, are associated with increased mortality.\(^5,6\) Because risk of most fractures rises steeply with age, and the population is aging, fracture burden is projected to increase in coming decades.

In short-term (18 to 36 months) randomized controlled trials (RCTs) for osteoporosis treatment, bisphosphonates (alendronate, zoledronate, risedronate, ibandronate), denosumab, teriparatide, and abaloparatide lower risk of nonvertebral fractures, clinical vertebral fractures (usually diagnosed in the community because of back pain, with study comparison of study and community radiographs), and radiographic vertebral fractures (identified in studies by comparing vertebral heights on scheduled serial vertebral radiographs; mostly unrecognized in the community).\(^7,8\) Several bisphosphonates (alendronate, zoledronate, risedronate) and denosumab also lower risk of hip fractures.\(^7\)

Purpose of Review

To summarize the effects of long-term osteoporosis drug treatment and of osteoporosis drug treatment discontinuation and holidays.

Key Messages

- Evidence on the effects of long-term osteoporosis drug treatment and drug continuation versus discontinuation is mostly limited to white, healthy, postmenopausal women.
- Long-term alendronate reduces radiographic vertebral and nonvertebral fractures in women with osteoporosis; long-term zoledronate reduces vertebral and nonvertebral fractures in women with osteopenia or osteoporosis.
- Long-term bisphosphonates may increase atypical femoral fractures and osteonecrosis of the jaw, although both are rare.
- In women with osteoporosis, long-term raloxifene reduces vertebral fractures, but not hip or nonvertebral fractures, and increases venous thromboembolism.
- Long-term oral hormone therapies reduce hip and clinical fractures but increase multiple serious harms.
- Evidence is insufficient about the effects of long-term denosumab, risedronate, ibandronate, teriparatide, and abaloparatide on fractures and harms.
- Continuing bisphosphonates after 3–5 years versus discontinuation reduces some measures of vertebral fractures, but not nonvertebral fractures.
Despite the evidence on the efficacy of short-term osteoporosis drug treatment for reducing fracture risk in appropriate patients, there is uncertainty about the balance of benefits and harms of long-term or continued treatment. A recent American College of Physicians clinical practice guideline recommended treatment of osteoporotic women with a bisphosphonate or denosumab for 5 years to reduce risk of hip and vertebral fractures, and suggested that high-risk patients may benefit from more than 5 years of treatment.\(^8\)

However, concerns that long-term bisphosphonate persistence in bone might increase fracture risk by inhibiting normal repair of bone microdamage\(^9,10\) have led to the suggestion that bisphosphonate treatment be discontinued periodically.\(^7\) Several groups advocate bisphosphonate "drug holidays" to minimize harms while preserving as much fracture benefit as possible. But, there is no consensus about who should get them, when they should start, how long they should last, and the criteria for restarting treatment.\(^9,11,12\) By contrast, drug holidays are not recommended after denosumab, because bone loss increases rapidly after discontinuation, possibly increasing risk of radiographic vertebral fractures.\(^13\)

Uncertainties about the most appropriate use of long-term osteoporosis drug treatment and of osteoporosis drug holidays led to scheduling a National Institutes of Health (NIH) Office of Disease Prevention (ODP) Pathways to Prevention (P2P) workshop. The goals of the workshop were to present an evidence-based synthesis of the pertinent research base, and to suggest future research to assist patients, clinicians, and other healthcare decision makers.

To further these aims, we conducted this systematic review to address the following questions: (1) What are the effects of long-term (>3 years) osteoporosis drug treatment versus control on risks of incident fractures and harms; (2) Do effects of long-term osteoporosis drug treatment vary as a function of patient, bone, or osteoporosis drug characteristics; (3) Among individuals receiving osteoporosis drug treatment to prevent fracture, what are the effects of continuing versus discontinuing treatment (i.e., osteoporosis drug holiday) on risks of incident fractures and harms; and (4) Do these outcomes of drug holidays vary as a function of patient, bone, or osteoporosis drug characteristics?

**Methods**

The review was conducted following the Agency for Healthcare Research and Quality (AHRQ) methods guidance. The protocol is available at https://effectivehealthcare.ahrq.gov/topics/osteoporosis-fracture-prevention/research-protocol and is registered in PROSPERO. We detail our literature search strategy, study selection criteria, and data extraction and synthesis methods in the full report.

**Results**

We identified 8,356 unique publications through October 2018, of which 61 met eligibility criteria and were included in the review. Of 48 publications with low or medium risk of bias (ROB), there were 35 randomized or controlled clinical trials (9 unique studies) and 13 controlled observational studies (11 unique studies) (Appendix C of the full report). Most publications were based on three RCTs of alendronate, zoledronic acid, and raloxifene, respectively, and their extension studies.

All trials enrolled only postmenopausal women, with most limited to women with osteoporosis as defined by bone mineral density (BMD) and vertebral fracture history, and some including women with osteopenia. Observational studies included 84 to 100 percent women. Mean participant age was 72 years, with all but two studies reporting mean age <80 years. Most observational studies presumed participants had osteoporosis because of past fracture or use of osteoporosis drugs, but none reported BMD status.
**Long-Term Osteoporosis Drug Treatment**

**Efficacy**

Seven eligible placebo-controlled RCTs with low or medium ROB examined the effect of long-term treatment, one each for alendronate, zoledronate, raloxifene, denosumab, and estrogen, and two for estrogen/progestin.

In women with osteopenia or osteoporosis by BMD, but with no past vertebral fracture, 4 years of alendronate versus placebo reduced incident radiographic vertebral fractures (hazard ratio [HR] 0.56 [95% confidence interval (CI), 0.39, 0.80]) (high strength of evidence [SOE]) (Table A), while absolute risk reductions for incident hip and nonvertebral fracture were small and not statistically significant (low SOE). In women with osteoporosis by BMD or past fracture, 4 years of raloxifene versus placebo reduced incident radiographic (relative risk [RR] 0.64 [0.53, 0.76]) and clinical vertebral fractures (RR 0.58 [95% CI 0.43, 0.79]) (both high SOE). However, raloxifene did not reduce incident hip or nonvertebral fracture (moderate and high SOE, respectively). In older women with osteopenia or osteoporosis, 6 years of zoledronate versus placebo reduced incident clinical fractures (HR 0.73 [95% CI 0.60, 0.90]) (low SOE), incident nonvertebral fractures (HR 0.66 [95% CI 0.51, 0.85]) (moderate SOE) and incident clinical vertebral fractures (HR 0.41 [95% CI 0.22, 0.75]) (moderate SOE). Nonvertebral fractures appeared similarly reduced in the subset of women with osteopenia. In women with unknown osteoporosis or osteopenia status, incident clinical fractures (high SOE) and incident hip fractures (moderate SOE) both were reduced with hormone therapy compared to placebo, with 5.6 years of oral estrogen/progestin in women with an intact uterus, and with 7 years of unopposed oral estrogen in women with a hysterectomy.

Evidence was insufficient to compare fracture risk between women on long-term denosumab versus placebo, and there were no data from eligible trials about the long-term fracture efficacy of sequential osteoporosis drug therapy (e.g., anabolic followed by anti-resorptive, or denosumab followed by bisphosphonate).

Alendronate, zoledronate, denosumab, and raloxifene for long-term treatment each increased hip and lumbar spine BMD compared to placebo.

**Variation in Efficacy as a Function of Patient, Bone, or Osteoporosis Drug Characteristics**

Efficacy of long-term alendronate appears to vary as a function of baseline BMD. Relative risk of incident clinical fractures was significantly reduced in women with osteoporotic BMD (femoral neck BMD <-2.5 (moderate SOE), but not in women with osteopenic BMD (femoral neck BMD -1.6 to >-2.5) (low SOE) (Table A; Appendix Table D8). In women with osteoporosis, relative risk of incident radiographic vertebral fracture with long-term alendronate was halved versus placebo (HR 0.50 [95% CI 0.31, 0.82]) (moderate SOE). Although women with femoral neck BMD -2.5 to -2 had a similar relative reduction in these fractures versus placebo (HR 0.54 [95% CI 0.28, 1.04]) (low SOE), a lower proportion of women with osteopenia had incident radiographic vertebral fractures, and results were not statistically significant. No tests of interaction were reported for these BMD stratified results.

In a post hoc analysis, women with osteoporosis had a reduced risk of incident hip fracture with long-term alendronate versus placebo, but women with osteopenia had no reduced risk (p-value for interaction not reported). In additional post hoc analyses, some conducted in women with osteopenia, neither past nonvertebral fracture, 10-year major osteoporotic fracture probability calculated with femoral neck BMD, nor pretreatment levels of bone turnover markers significantly modified the effect of long-term alendronate versus placebo on risk of any incident fracture outcome.

Neither age, baseline BMD, nor baseline radiographic vertebral fracture modified
the effect of long-term raloxifene versus placebo on risk of incident fractures. Two large trials of oral hormone therapy versus placebo in women with unknown osteoporosis or osteopenia status reported inconsistent findings about whether treatment effect on risk of incident hip and clinical fractures differed as a function of age or time since menopause.\textsuperscript{19, 28} However, authors minimized their one significant interaction for age because of the many interactions examined.\textsuperscript{19}

### Table A. Evidence on efficacy of long-term (\textgreater{}3 years) osteoporosis drug treatment

<table>
<thead>
<tr>
<th>Comparison</th>
<th># Studies by Design</th>
<th>Treatment Duration</th>
<th>Participant Characteristics</th>
<th>Incident Fracture Outcome</th>
<th>Relative and Absolute Risk Differences (95% CI)</th>
<th>Strength of Evidence* (Justification)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alendronate vs. placebo</td>
<td>1 RCT\textsuperscript{14}</td>
<td>4 yr</td>
<td>4,432 PM women with osteopenia or osteoporosis (T-score &lt;-1.6) and no RVF</td>
<td>CF</td>
<td>No difference: HR=0.86 [0.73, 1.01]; ARR=-2 [-4, 0]</td>
<td>Low (IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NVF</td>
<td>No difference: HR=0.88 [0.74, 1.04]; ARR=-1 [-3, 0]</td>
<td>Low (IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hip</td>
<td>No difference: HR=0.79 [0.43, 1.44]; ARR=-0.2 [-0.8, 0.4]</td>
<td>Low (h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RVF</td>
<td>Lower risk: HR=0.56 [0.39, 0.80]; ARR=-2 [-3, -1]</td>
<td>High</td>
</tr>
<tr>
<td>Zoledronate vs. placebo</td>
<td>1 RCT\textsuperscript{15}</td>
<td>6 yr</td>
<td>1,631 PM women with osteoporosis by BMD (T-score &lt;-2.5) and no RVF</td>
<td>CF</td>
<td>Lower risk: HR=0.64 [0.50, 0.82]; ARR=-7 [-10, -3]</td>
<td>Moderate (RB)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RVF</td>
<td>Lower risk: HR=0.50 [0.31, 0.82]; ARR=-3 [-5, -1]</td>
<td>Moderate (RB)</td>
</tr>
<tr>
<td>Denosumab\textsuperscript{**} vs. placebo</td>
<td>1 RCT\textsuperscript{18}</td>
<td>4 yr</td>
<td>2,000 PM women \geq 65 with osteoporosis or osteopenia</td>
<td>CF</td>
<td>Lower risk: HR=0.73 [0.60, 0.90]; ARR=-5 [-9, -2]</td>
<td>Moderate (IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NVF</td>
<td>Lower risk: HR=0.66 [0.51, 0.85]; ARR=-5 [-8, -2]</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hip</td>
<td>No difference: HR=0.66 [0.27, 1.16]; ARR=-0.4 [-1, 0.5]</td>
<td>Low (h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CVF</td>
<td>Lower risk: HR=0.41 [0.22, 0.75]; ARR=-2 [-3, -1]</td>
<td>Moderate (IM)</td>
</tr>
<tr>
<td>Raloxifene vs. placebo</td>
<td>1 RCT with 1 CCT extension\textsuperscript{16, 17, 26-37}</td>
<td>4 to 8 yr</td>
<td>365 PM women with osteopenia or osteoporosis by BMD</td>
<td>CF</td>
<td>RR=0.97 [0.40, 2.35]; ARR=-0.4 [-10, 9]</td>
<td>Insufficient (RB, IN, h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NVF</td>
<td>No difference: 4 yr: RR=0.93 [0.81, 1.06]; ARR NA 8 yr: RR=1.00 [0.82, 1.21]; ARR NA</td>
<td>4 yr: High 8 yr: Moderate (RB)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hip</td>
<td>No difference: 4 yr: RR=0.97 [0.62, 1.52]; ARR=0 [-0.6, 0.5]</td>
<td>Moderate (IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CVF</td>
<td>Lower risk: 4 yr: RR=0.58 [0.43, 0.79]; ARR=-2 [-3, -1]</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RVF</td>
<td>Lower risk: 4 yr: RR=0.64 [0.53, 0.76]; ARR=-5 [-6, -3]</td>
<td>High</td>
</tr>
</tbody>
</table>
Table A. Evidence on efficacy of long-term (>3 years) osteoporosis drug treatment (continued)

<table>
<thead>
<tr>
<th>Comparison</th>
<th># Studies by Design</th>
<th>Participant Characteristics</th>
<th>Incident Fracture Outcome</th>
<th>Relative and Absolute Risk Differences (95% CI)</th>
<th>Strength of Evidence* (Justification)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estrogen vs. placebo</strong></td>
<td>1 RCT</td>
<td>10,739 PM women with past hysterectomy</td>
<td>CF</td>
<td>Lower risk: HR=0.71 [0.64, 0.80]; ARR= -4 [-5, -3]</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hip</td>
<td>Lower risk: HR=0.65 [0.45, 0.94]; ARR= -0.5 [-0.9, -0.08]</td>
<td>Moderate (IM)</td>
</tr>
<tr>
<td><strong>Estrogen/progestin vs. placebo</strong></td>
<td>1 RCT</td>
<td>3,816 PM women with past hysterectomy and past clinical fracture</td>
<td>CF</td>
<td>Lower risk: HR=0.73 [0.62, 0.86]; ARR= -5 [-7, -2]</td>
<td>Low (RB, IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hip</td>
<td>Lower risk: HR=0.55 [0.32, 0.94]; ARR= -1 [-2, 0]</td>
<td>Low (RB, IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>53 PM women with past hysterectomy and osteoporosis by BMD</td>
<td>CF</td>
<td>HR=0.83 [0.17, 3.91]; ARR NA</td>
<td>Insufficient (RB, h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CVF</td>
<td>Lower risk: HR=0.65 [0.46, 0.92]; ARR= -0.3 [-0.5, -0.02]</td>
<td>Moderate (IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>363 PM women with past hysterectomy and osteopenia by BMD</td>
<td>CF</td>
<td>HR=0.83 [0.49, 1.40]; ARR NA</td>
<td>Insufficient (RB, h-IM)</td>
</tr>
<tr>
<td><strong>Estrogen/progestin vs. placebo</strong></td>
<td></td>
<td>16,608 PM women with intact uterus</td>
<td>CF</td>
<td>Lower risk: HR=0.76 [0.69, 0.83]; ARR= -2.4 [-3.3, -1.5]</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hip</td>
<td>Lower risk: HR=0.67 [0.47, 0.96]; ARR= -0.3 [-0.6, -0.03]</td>
<td>Moderate (IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CVF</td>
<td>Lower risk: HR=0.65 [0.46, 0.92]; ARR= -0.3 [-0.5, -0.02]</td>
<td>Moderate (IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5,897 PM women with intact uterus and past clinical fracture</td>
<td>CF</td>
<td>Lower risk: HR=0.78 [0.68, 0.91]; ARR= -3 [-5, -1]</td>
<td>Low (RB, IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hip</td>
<td>No difference: HR=0.77 [0.48, 1.22]; ARR = -0.3 [-0.9, 0.2]</td>
<td>Low (RB, IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PM women with intact uterus and osteoporosis by BMD n not reported</td>
<td>CF</td>
<td>HR=0.53 [0.25, 1.10]; ARR NA</td>
<td>Insufficient (RB, h-IM)</td>
</tr>
</tbody>
</table>

(RB, h-IM) Insufficient
Table A. Evidence on efficacy of long-term (>3 years) osteoporosis drug treatment (continued)

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Participant Characteristics</th>
<th>Incident Fracture Outcome</th>
<th>Relative and Absolute Risk Differences (95% CI)</th>
<th>Strength of Evidence* (Justification)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estrogen/progestin vs. nonplacebo control</td>
<td>36 PM women with osteoporosis by BMD (T-score &lt;-2) and RVF</td>
<td>NVF</td>
<td>RR=0.93 [0.06, 13.5]; ARR= -0.5 [-19, 18]</td>
<td>Insufficient (RB, h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RVF</td>
<td>RR=0.37 [0.09, 1.62]; ARR= -22 [-53, 8]</td>
<td>Insufficient (RB, h-IM)</td>
</tr>
</tbody>
</table>

Abbreviations: ARR=absolute risk reduction; BMD=bone mineral density; CCT=controlled clinical trial; CF=clinical fracture; CI=confidence intervals; CVF=clinical vertebral fracture; h-IM=highly imprecise; HR=hazard ratio; IM=imprecise; IN=indirect; NA=not available; NVF=nonvertebral fracture; PM=postmenopausal; RB=medium risk of bias; RCT=randomized controlled trial; RR=risk ratio; RVF=radiographic vertebral fracture.

*Definitions of terms for strength of evidence grades and domains ratings are detailed in the section of the main report titled, 'Strength of Evidence for Major Comparisons and Outcomes.'
†Results reported for raloxifene 60 mg/d and 120 mg/day groups pooled together.
‡Results reported for raloxifene 60 mg/d dose group.
**Analyses pooled all participants initially assigned to denosumab, which included both those who received long-term and short-term denosumab.

Harms

Due to few events, RCT data provided insufficient evidence about whether long-term alendronate or zoledronate increase risk of radiologically confirmed atypical femoral fracture (AFF), subtrochanteric or femoral shaft fractures without radiologically confirmed AFF features, or osteonecrosis of the jaw (ONJ). Data from controlled, long-term observational studies suggest that alendronate and bisphosphonates as a class increase both radiologically confirmed AFF (low SOE), and subtrochanteric and femoral shaft fractures without radiologically confirmed with AFF features (low SOE) (Table B). Relative risks for these outcomes varied from 1 to >100 across studies, likely related to heterogeneity in designs. Relative risks appeared higher for radiologically confirmed AFF than for cases defined only by diagnostic codes. However, some AFF risk estimates were calculated using controls with subtrochanteric or femoral shaft fractures without AFF features. In those cases, risk estimates reflect the probability that a subtrochanteric or femoral shaft fracture will have AFF features, and not the relative risk of sustaining an AFF. Studies also differed in whether fractures cases were excluded for cancer and excess trauma; in whether current bisphosphonate use was compared to no use, limited past use, or nonbisphosphonate osteoporosis drug use; and in how they addressed potential confounding.

Few observational studies provided data about risk of ONJ. Different studies provided low and insufficient strength evidence, respectively, about whether long-term alendronate increases risk of ONJ. Relative risk estimates varied widely between studies, likely due to heterogeneity in case definitions, treatment control groups, and covariate modeling.
We found insufficient evidence about whether long-term zoledronate increases risk of AFF or ONJ. The single trial of long-term zoledronate versus placebo reported no cases of AFF or ONJ in either treatment group, and we identified no eligible observational studies that evaluated risk of these harms with long-term zoledronate.

Due to its pooling of results for both short- and long-term denosumab treatment, it was not possible to conclude anything about the risk of harms of long-term denosumab compared with placebo from the one study that met eligibility for this review. In long-term trials of oral hormone therapy, specifically estrogen/progestin and estrogen versus placebo in postmenopausal women with unknown osteoporosis or osteopenia status, risk was significantly increased for cardiovascular disease, and mild cognitive impairment or dementia. Risk for a composite outcome measure defined to weigh risk of incidence of any of several serious harms (coronary heart disease, invasive breast cancer, stroke, pulmonary embolism, endometrial cancer, colorectal cancer, or death due to other causes) versus hip fracture did not differ between estrogen/progestin and placebo or between estrogen and placebo. Results were similar regardless of participants' baseline fracture risk, indicating that risk of one or more of these harms offset the reduction in hip fractures even in participants at highest baseline fracture risk. In addition, estrogen/progestin was associated with an increased risk of invasive breast cancer. Long-term raloxifene versus placebo in treatment for osteoporosis significantly increased the risk of deep vein thrombosis and pulmonary embolism by about 3 to 4-fold.

Variation in Harms as a Function of Patient, Bone, or Osteoporosis Drug Characteristics

We found little evidence about factors that modify risk of harms with long-term osteoporosis drug treatment. One study was inconclusive about whether relative risk for AFF associated with bisphosphonate use increased with age. Three controlled observational studies reported that >5 years of bisphosphonate use increased risk of subtrochanteric or femoral shaft fractures (ST/FSF) or radiologically confirmed AFF more than did 3-5 years of use. However, none of these studies reported tests for interaction by treatment duration. For long-term raloxifene versus placebo, one study reported that risk of deep venous thrombosis and pulmonary embolism did not vary as a function of baseline cardiovascular risk, and another that risk of incident stroke was lower with raloxifene versus placebo in women with increased cardiovascular risk. Trials of long-term oral hormone therapy evaluated whether risk of harms varied by a long list of patient characteristics. Though results suggested that risk of breast cancer with estrogen/progestin compared to placebo may be greater with increased duration of prior postmenopausal hormone use, this was the only significant result out of many examined, and may have been due to chance. Strength of evidence was not assessed for effect modifiers.
<table>
<thead>
<tr>
<th>Comparison</th>
<th># Studies by Design</th>
<th>Participant Characteristics</th>
<th>Harms</th>
<th>Relative and Absolute Risk Differences (95% CI)</th>
<th>Strength of Evidence* (Justification)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alendronate vs. placebo</td>
<td>1 RCT&lt;sup&gt;90&lt;/sup&gt; 3 to 4.5 yr</td>
<td>6,459 PM women with osteopenia or osteoporosis (T-score &lt;-1.6) with or without RVF</td>
<td>ST or FS fracture DX with rare x-ray review for confirmation of AFF features (n=2 cases)</td>
<td>HR=1.03 [0.06, 16.46]; ARR=0 [-0.09, 0.09]</td>
<td>Insufficient (h-IM)</td>
</tr>
<tr>
<td>Alendronate vs. no osteoporosis drug treatment</td>
<td>2 retrospective cohort observational studies&lt;sup&gt;41, 51, 52&lt;/sup&gt; 3.8 yr (mean) and ≥6 yr</td>
<td>534 adults &gt;60 yr with nonhip fracture (90% women)</td>
<td>ST or FS fracture DX codes without x-ray confirmation of AFF features (n=5 cases)</td>
<td>≥6 yr; HR=1.37 [0.22, 8.62]; ARR NA</td>
<td>Insufficient (RB, h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>220,360 adults (85% women) exposed to alendronate or no osteoporosis drug general population controls from national database</td>
<td>ST or FS fracture DX codes without x-ray confirmation of AFF features (n=309 cases)</td>
<td>Higher risk: 3.8 yr; ST: 0.17% vs. 0.06%; HR=2.41 [1.78, 3.27]; ARR=0.11 [0.08, 0.15] 3.8 yr; FS: 0.12% vs. 0.03%; HR=2.90 [1.97, 4.26]; ARR 0.09 (0.06, 0.12) 3.8 yr: ST/FS: 0.29% vs. 0.09%; ARR 0.20 (0.15, 0.25)</td>
<td>Low (RB, IM, LE)</td>
</tr>
<tr>
<td>Alendronate vs. raloxifene</td>
<td>1 retrospective cohort observational study&lt;sup&gt;43&lt;/sup&gt; ~4 yr (mean)</td>
<td>8,354 women aged ≥50 yr from database of 1 hospital</td>
<td>ONJ DX codes with x-ray and pathology features (n=40 cases)</td>
<td>Higher risk with alendronate: HR=7.42 [1.02, 54.09]; ARR NA</td>
<td>Low (RB, IM, LE)</td>
</tr>
<tr>
<td>Alendronate vs. raloxifene or calcitonin</td>
<td>1 retrospective cohort observational study&lt;sup&gt;42&lt;/sup&gt; Up to 6 yr</td>
<td>43,645 adults aged &gt;50 yr (84% women) with recent hip or vertebral fracture now on osteoporosis drug treatment from national database</td>
<td>ONJ DX codes without x-ray or pathology review (n=46 cases)</td>
<td>HR=0.86 [0.44, 1.69]&lt;sup&gt;†&lt;/sup&gt;; ARR NA</td>
<td>Insufficient (RB, h-IM)</td>
</tr>
<tr>
<td>Zolendronate vs. placebo</td>
<td>1 RCT&lt;sup&gt;95&lt;/sup&gt; 6 yr</td>
<td>2,000 PM women ≥65 with osteoporosis or osteopenia</td>
<td>SAE</td>
<td>No difference: OR=0.84 [0.70, 1.00]; ARR= -4 [-9, 0]</td>
<td>Low (IM)</td>
</tr>
</tbody>
</table>
### Table B. Evidence on harms of long-term (>3 years) osteoporosis drug treatment (continued)

<table>
<thead>
<tr>
<th>Comparison # Studies by Design</th>
<th>Participant Characteristics</th>
<th>Harms</th>
<th>Relative and Absolute Risk Differences (95% CI)</th>
<th>Strength of Evidence* (Justification)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bisphosphonate‡ vs. no bisphosphonate 3 observational studies‡, 38, 39, 53 &gt;3 yr</td>
<td>~2.8 million (retrospective cohort) and 1.124 (case-control) adults aged &gt;55 yr from national database (87% women cases and 52% women controls in cohort analysis; 86% women in case-control analysis)</td>
<td>AFF with radiologic features (n=172 cases)</td>
<td>Higher risk: Cohort &gt;4 yr: RR=126 [55, 288]; ARR NA Case-control 3-4 yr: OR=40 [17, 91]; ARR=NA 4-5 yr: OR=116 [58, 234]; ARR=NA &gt;5 yr: OR=93 [66, 132]; ARR=NA</td>
<td>Low (RB, CO, LE)</td>
</tr>
<tr>
<td></td>
<td>264 women aged &gt;65 yr from national primary practice database (case-control)</td>
<td>ST or FS fracture DX codes without x-ray confirmation of AFF features (n=44 cases)</td>
<td>Higher risk: &gt;3 yr: OR=9.46 [2.17, 41.3]; ARR=NA</td>
<td>Low (RB, LE)</td>
</tr>
<tr>
<td></td>
<td>6,644 women aged &gt;50 yr with hip or femoral fracture from 8 hospital medical records databases (nested case-control)</td>
<td>AFF with radiologic features (n=196 cases)</td>
<td>Higher risk: Mean use 5.2 yr: OR=25.65 [10.74, 61.28]; ARR=NA</td>
<td>Low (RB, LE)</td>
</tr>
<tr>
<td>Current vs. past bisphosphonates‡ 2 case-control observational studies‡, 48, 54 &gt;3 yr</td>
<td>172 PM women with &gt;1 yr bisphosphonate use from 1 hospital database</td>
<td>AFF with radiologic features (n=43 cases)</td>
<td>Higher risk with current bisphosphonate: HR=3.36 [1.77, 11.91] to 5.17 [2.0, 13.36]; ARR NA</td>
<td>Low (RB, LE)</td>
</tr>
<tr>
<td></td>
<td>1,855 women aged &gt;68 yr from a provincial database</td>
<td>ST or FS fracture DX codes without x-ray review (n=325 cases)</td>
<td>Higher risk with current bisphosphonate: 3-5 yr: OR=1.59 [0.80, 3.15]; ARR=NA &gt;5 yr: OR=2.74 [1.25, 6.02]; ARR=NA</td>
<td>Low (RB, IM)</td>
</tr>
</tbody>
</table>
### Table B. Evidence on harms of long-term (>3 years) osteoporosis drug treatment (continued)

<table>
<thead>
<tr>
<th>Comparison # Studies by Design Treatment Duration</th>
<th>Participant Characteristics</th>
<th>Harms</th>
<th>Relative and Absolute Risk Differences (95% CI)</th>
<th>Strength of Evidence* (Justification)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bisphosphonates† vs. pooled raloxifene or calcitonin 1 retrospective cohort observational study49 &gt;3 yr</td>
<td>4,097 Medicare beneficiaries (97% women)</td>
<td>ST or FS fracture DX codes without x-ray confirmation of AFF features (n=34 cases)</td>
<td>3-5 yr: HR=1.20 [0.55, 2.61]; ARR=0.1 [-0.3, 0.5] &gt;5 yr: HR=2.02 [0.41, 10.0]; ARR=0.1 [-0.1, 0.4]</td>
<td>Insufficient (RB, h-IM)</td>
</tr>
<tr>
<td>Denosumab†† vs. placebo 1 RCT18 4 yr</td>
<td>365 PM women with osteopenia or osteoporosis by BMD</td>
<td>SAE</td>
<td>RR=1.64 [0.69, 3.88]; ARR=7 [-3, 17]</td>
<td>Insufficient (RB, IN, h-IM)</td>
</tr>
<tr>
<td>Raloxifene vs. placebo 1 RCT with 1 CCT extension16, 17, 26-37 4 to 8 yr</td>
<td>6,828 PM women with osteoporosis by BMD or RVF</td>
<td>SAE</td>
<td>No difference: 8 yr: RR=0.93 [0.86, 1.00]**; ARR=-3 [-6, 0]</td>
<td>Low (RB, IM)</td>
</tr>
<tr>
<td>Raloxifene vs. no treatment 1 retrospective cohort observational study41, 52 3.8 yr (mean)</td>
<td>19,324 adults (85% women) exposed to raloxifene or no osteoporosis drug general population controls from national database</td>
<td>ST or FS fracture DX codes without x-ray confirmation of AFF features (n=25 cases)</td>
<td>ST: HR=1.06 [0.34, 3.32]; ARR=0.04 [-0.06, 0.14] FS: HR=0.82 [0.21, 3.20]; ARR 0.01 [-0.07, 0.09]</td>
<td>Insufficient (RB, h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ONJ DX codes without x-ray or pathology review (n=2 cases)</td>
<td>2 cases, only in control group</td>
<td>Insufficient (RB, h-IM)</td>
</tr>
</tbody>
</table>

Abbreviations: AFF=atypical femoral fracture; ARR=absolute risk reduction; BMD=bone mineral density; CCT=controlled clinical trial; CI=confidence intervals; CO=consistent; DX=diagnosis; FS=femoral shaft; h-IM=highly imprecise; HR=hazard ratio; IM=imprecise; IN=indirect; LE=large effect; NA=not available (data not reported); ONJ=osteonecrosis of the jaw; OR=odds ratio; PM=postmenopausal; RB=medium risk of bias; RCT=randomized controlled trial; RR=risk ratio; RVF=radiographic vertebral fracture; SAE=serious adverse event; ST=subtrochanteric

*Definitions of terms for strength of evidence grades and domains ratings are detailed in the section of the main report titled, 'Strength of Evidence for Major Comparisons and Outcomes.'

†Because the higher adjusted incidence rates in the alendronate group (0.15%) compared with the raloxifene-calcitonin group (0.08%) suggested a possibly increased risk, we manually recalculated the estimate of effect and found RR 1.20 (95% CI 0.59, 2.56). Authors were contacted for clarification, but did not reply.

‡Included bisphosphonates varied by study. All studies included alendronate, risedronate, and one or more of the following: ibandronate, etidronate, and zoledronate.

**Results reported for raloxifene 60 mg/d dose group.

††Analyses pooled all participants initially assigned to denosumab, which included both those who received long-term and short-term denosumab.
Osteoporosis Drug Holidays

Effect

In postmenopausal women who previously received 3-5 years of bisphosphonate, two trials compared continued versus discontinued alendronate for 5 more years, and one compared continued versus discontinued zoledronate for 3 more years.\(^55-57\) None found a reduction in incident nonvertebral fractures (Table C).

However, these trials collectively suggested a reduction in incident vertebral fractures. One enrolled women who previously received 5 years of alendronate in the active treatment arm of a trial for osteopenia or osteoporosis and a subsequent extension, and reported that alendronate continuation for 5 years versus placebo (drug holiday) reduced incident clinical vertebral fractures (HR 0.45 [95% CI 0.24, 0.85]) (moderate SOE), but not incident radiographic vertebral fractures (HR 0.86 [95% CI 0.60, 1.22]) (moderate SOE).\(^55\) A second trial enrolled women who previously received 3 years of zoledronate in the active treatment arm of a trial for osteoporosis, and reported that zoledronate continuation for 3 years versus placebo (drug holiday) reduced incident radiographic vertebral fractures (HR 0.51 [95% CI 0.26, 0.95]) (low SOE), but that evidence was insufficient about incident clinical vertebral fracture.\(^54\) In a third trial that enrolled women who previously received 5 years of alendronate in the active treatment arm of a trial for osteoporosis and a subsequent extension, and then nonrandomly assigned them to alendronate continuation for 2 years and 5 years versus placebo (drug holiday), evidence was insufficient to draw conclusions about differences in risk of incident clinical vertebral fractures or incident radiographic vertebral fractures.\(^57-59\) Similarly, we could not draw conclusions from a small, 4-year denosumab dose-finding trial, because fracture results were pooled between the denosumab continuation and discontinuation treatment arms.\(^18\)

In women who previously received 3-5 years of bisphosphonate treatment, continued bisphosphonate treatment for an additional 3-5 years was associated with stable or slightly decreased hip BMD, whereas women assigned to discontinue treatment (drug holiday) had significantly larger declines in hip BMD. A 4-year denosumab trial reported that compared to baseline, hip and spine BMD were most increased in women assigned denosumab for 4 years, back to pretreatment baseline in women assigned denosumab for 2 years followed by discontinuation for 2 years, and immediately increased in women assigned denosumab for 2 years, placebo for 1 year, then denosumab for 1 year.

Variation in Effect as a Function of Patient, Bone, or Osteoporosis Drug Characteristics

In post hoc analyses, the effect of alendronate continuation versus discontinuation (drug holiday) on risk of incident clinical fractures, which did not differ overall, did not vary as a function of baseline BMD or radiographic vertebral fracture status.\(^55, 60\) Further post hoc subgroup analyses suggested that in women without a prevalent radiographic vertebral fracture, continued alendronate versus discontinuation reduced risk of incident nonvertebral fractures in women with osteoporotic BMD but not in those with osteopenia. However, risk for incident nonvertebral fractures between alendronate continuation and discontinuation appeared not to differ in women with prevalent radiographic vertebral fractures. Further, risk of incident vertebral fractures appeared no different between alendronate continuation and discontinuation groups, regardless of baseline BMD or radiographic vertebral fracture status. The single positive outcome may have been due to chance. We found no evidence about possible modifiers of the effect of continuing any other osteoporosis drug treatment versus discontinuation on risk of incident fracture.
**Harms**

Trials of alendronate and zoledronate continuation versus discontinuation reported no difference between treatment groups in risk of serious adverse events (Table D).\textsuperscript{55, 56, 58-62} Too few cases of AFF with confirmed radiologic features, subtrochanteric or femoral shaft fractures without confirmed AFF features, or ONJ, occurred in these trials to draw conclusions about differences in their risk between treatment continuation and discontinuation groups.\textsuperscript{50, 58} One retrospective cohort study reported that incidence of AFF was significantly higher in bisphosphonate users (99% alendronate) who continued versus discontinued use (0.15% vs. 0.03%; estimated OR 6.03 [95% CI 1.87, 19.42]).\textsuperscript{63} However, this analysis did not radiologically confirm AFF diagnoses and did not describe accounting for potentially confounding variables. Though atrial fibrillation appeared more frequently with zoledronate continuation versus discontinuation, the absolute number of events was low and possible differences between treatment groups were not statistically significant.\textsuperscript{58, 59} It was not possible to draw conclusions about differences in harms between the denosumab continuation and discontinuation arms in one trial reporting because harms results for these two groups were pooled.\textsuperscript{18}

**Variation in Harms as a Function of Patient, Bone, or Osteoporosis Drug Characteristics**

We found no evidence about whether the risk of harms between continuation of any osteoporosis drug treatment and discontinuation varied as a function of patient, bone or drug characteristics.

| Table C. Evidence on effects of osteoporosis drug continuation versus discontinuation* on incident fractures |
|---|---|---|---|
| **Comparison** | **Participant Characteristics** | **Incident Fracture Outcome** | **Relative and Absolute Risk Differences (95% CI)** | **Strength of Evidence\textsuperscript{a}** |
| Alendronate continuation vs. discontinuation (AL x 10 yr vs. AL x 5 yr followed by PBO x 5 yr) | 1,099 PM women previously received alendronate 5 yr for osteopenia or osteoporosis (T-score <-1.6) | CF | No difference: RR=0.93 [0.71, 1.21]; ARR= -1 [-6, 4] | Moderate (IM) |
| | | NVF | No difference: RR=1.00 [0.76, 1.32]; ARR= -0.1 [-5, 5] | Moderate (IM) |
| | | Hip | RR=1.02 [0.51, 2.10]; ARR=0 [-2, 2] | Insufficient (h-IM) |
| | | CVF | Lower risk with continuation: RR=0.45 [0.24, 0.85]; ARR= -3 [-5, -0.5] | Moderate (IM) |
| | | RVF | No difference: RR=0.86 [0.60, 1.22]; ARR= -1 [-5, 2] | Moderate (IM) |
### Table C. Evidence on effects of osteoporosis drug continuation versus discontinuation* on incident fractures (continued)

<table>
<thead>
<tr>
<th>Comparison Study</th>
<th>Design by Treatment Duration</th>
<th>Participant Characteristics</th>
<th>Incident Fracture Outcome</th>
<th>Relative and Absolute Risk Differences (95% CI)</th>
<th>Strength of Evidence†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alendronate</td>
<td>continuation vs. discontinuation (AL x 7 yr [A7] vs. AL x 5 yr followed by PBO x 2 yr [A5/P2]; AL x 10 yr [A10] vs. AL x 7 yr + PBO x 3 yr [A7/P3])</td>
<td>350 PM women previously received alendronate 5 yr for osteoporosis (T-score ≤-2.5) (n=350 for A7 vs. A5/P2; n=247 for A10 vs. A7/P3)</td>
<td>NVF</td>
<td>A7 vs. A5/P2: RR=0.87 [0.40, 1.91]; ARR= -1 [-7, 5] A10 vs. A7/P3: RR= 0.81 [0.38, 1.71]; ARR= -2 [-11, 6]</td>
<td>A7 vs. A5/P2: Insufficient (h-IM) A10 vs. A7/P3: Insufficient (RB, h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CVF</td>
<td>A7 vs. A5/P2: RR= 0.92 [0.40, 2.10]; ARR= -1 [-6, 5]</td>
<td>Insufficient (h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RVF</td>
<td>AL10 vs. AL5/P5: RR=1.40 [0.52, 3.74]; ARR=2.6 [-4.6, 9.9]</td>
<td>Insufficient (RB, h-IM)</td>
</tr>
<tr>
<td>Zoledronate</td>
<td>continuation vs. discontinuation (Z x 2 yr vs. Z x 1 yr followed by PBO x 1 yr)</td>
<td>379 PM women with osteopenia</td>
<td>CF</td>
<td>RR=1.37 [0.39, 4.78]; ARR=1 [-2, 4]</td>
<td>Insufficient (h-IM)</td>
</tr>
<tr>
<td>Zoledronate</td>
<td>continuation vs. discontinuation (Z x 6 yr vs. Z x 3 yr followed by PBO x 3 yr)</td>
<td>1,233 PM women previously received zoledronic acid 3 yr for osteoporosis by BMD or RVF</td>
<td>CF</td>
<td>No difference: HR=1.04 [0.71, 1.54]; ARR NA</td>
<td>Moderate (IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NVF</td>
<td>No difference: HR= 0.99 [0.7, 1.5]; ARR= -0.3 [-3, 3]</td>
<td>Moderate (IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hip</td>
<td>HR= 0.90 [0.33, 2.49]; ARR= -0.2 [-1, 1]</td>
<td>Insufficient (h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CVF</td>
<td>HR=1.81 [0.53, 6.2]; ARR NA</td>
<td>Insufficient (h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RVF</td>
<td>Lower risk with continuation: OR=0.51 [0.26, 0.95]; ARR= -3 [-6, -1]</td>
<td>Low (h-IM)</td>
</tr>
<tr>
<td>Zoledronate</td>
<td>continuation vs. discontinuation (Z x 9 yr vs. Z x 6 yr followed by PBO x 3 yr)</td>
<td>190 PM women previously received zoledronic acid 6 yr for osteoporosis by BMD or RVF</td>
<td>CF</td>
<td>HR=1.11 [0.45, 2.73]; ARR=1 [-7, 10]</td>
<td>Insufficient (h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RVF</td>
<td>OR=0.58 [0.13, 2.55]; ARR= -2 [-8, 4]</td>
<td>Insufficient (h-IM)</td>
</tr>
</tbody>
</table>
Table C. Evidence on effects of osteoporosis drug continuation versus discontinuation* on incident fractures (continued)

<table>
<thead>
<tr>
<th>Comparison # Studies by Design Treatment Duration</th>
<th>Participant Characteristics</th>
<th>Incident Fracture Outcome</th>
<th>Relative and Absolute Risk Differences (95% CI)</th>
<th>Strength of Evidence†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denosumab continuation vs. discontinuation (D x 4 yr vs. D x 2 yr followed by PBO x 2 yr) 1 RCT</td>
<td>314 PM women with osteopenia or osteoporosis by BMD</td>
<td>CF</td>
<td>No numerical data</td>
<td>Insufficient (no data)</td>
</tr>
</tbody>
</table>

Abbreviations: AL=alendronate; ARR=absolute risk reduction; BMD=bone mineral density; CF=clinical fracture; CI=confidence intervals; CVF=clinical vertebral fracture; D=denosumab; h-IM=highly imprecise; HR=hazard ratio; IM=imprecise; NA=not available (no data reported); NVF=nonvertebral fracture; OR=odds ratio; PBO=placebo; PM=postmenopausal; RCT=randomized controlled trial; RB=medium risk of bias; RR=risk ratio; RVF=radiographic vertebral fracture; Z=zoledronate

*Discontinuation ≥1 year after prior treatment ≥1 year.
†Definitions of terms for strength of evidence grades and domains ratings are detailed in the section of the main report titled, 'Strength of Evidence for Major Comparisons and Outcomes.'

Table D. Evidence on harms of osteoporosis drug continuation versus discontinuation*

<table>
<thead>
<tr>
<th>Comparison # Studies by Design Treatment Duration</th>
<th>Participant Characteristics</th>
<th>Harms</th>
<th>Relative and Absolute Risk Differences (95% CI)</th>
<th>Strength of Evidence† (Justification)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alendronate continuation vs. discontinuation (AL x 10 yr vs. AL x 5 yr followed by PBO x 5 yr) 1 RCT</td>
<td>1,099 PM women previously received alendronate 5 yr for osteopenia or osteoporosis (T-score &lt;-1.6)</td>
<td>SAE</td>
<td>Stated as no difference, but no data provided</td>
<td>Insufficient (no data)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Subtrochanteric or femoral shaft fracture DX with rare x-ray review (n=3 cases)</td>
<td>HR=1.33 [0.12, 14.67]; ARR= -0.1 [-0.5, 0.7]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ONJ not defined (n=0 cases)</td>
<td>No cases in either group</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A7 vs. A5/P2: RR= 1.05 [0.57, 1.96]; ARR=1 [-7, 8]</td>
<td>A7 vs. A5/P2: Insufficient (h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A10 vs. A7/P3: RR= 1.21 [0.75, 1.96]; ARR=5 [-7, 16]</td>
<td>A7/P3: Insufficient (RB, IM)</td>
</tr>
</tbody>
</table>
### Table D. Evidence on harms of osteoporosis drug continuation versus discontinuation*

<table>
<thead>
<tr>
<th>Comparison # Studies by Design Treatment Duration</th>
<th>Participant +Characteristics</th>
<th>Harms</th>
<th>Relative and Absolute Risk Differences (95% CI)</th>
<th>Strength of Evidence† (Justification)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bisphosphonate continuation vs. discontinuation (Continued BP 3.5 yr [mean] [persistent group] or 4.1 yr [mean] [nonpersistent group] vs. BP holiday 3.1 yr [mean]) 1 retrospective cohort observational study63</td>
<td>39,502 women aged &gt;45 yr with &gt;3 yr of prior &gt;50% adherent BP use (99% alendronate)</td>
<td>“AFF” (not defined) (n=47 cases)</td>
<td>Higher risk with bisphosphonate (alendronate) continuation: Pooled continuation groups 0.15% (44/28005) vs. discontinuation 0.03% (3/11497) OR=6.03 [1.87, 19.42]; ARR=0.13 [0.08, 0.19]</td>
<td>Low (RB, IM, LE)</td>
</tr>
<tr>
<td>Zoledronate continuation vs. discontinuation (Z x 2 yr vs. Z x 1 yr followed by PBO x 1 yr) 1 RCT61</td>
<td>379 PM women with osteopenia</td>
<td>SAE</td>
<td>No difference: RR=0.91 [0.50, 1.67]; ARR= -1 [-7, 5]</td>
<td>Low (h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ONJ (n=0 cases)</td>
<td>No cases occurred</td>
<td>Insufficient (h-IM)</td>
</tr>
<tr>
<td>Zoledronate continuation vs. discontinuation (Z x 6 yr vs. Z x 3 yr followed by PBO x 3 yr) 1 RCT58</td>
<td>1,233 PM women previously received zoledronic acid 3 yr for osteoporosis by BMD or RVF</td>
<td>SAE</td>
<td>No difference: RR=1.14 [0.96, 1.36]; ARR=4 [-1, 9]</td>
<td>Low (IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AFF not defined (n=0 cases)</td>
<td>No cases occurred</td>
<td>Insufficient (h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ONJ (exposed jaw bone &gt;6 wks) (n=1 case)</td>
<td>One case occurred (in continuation group)</td>
<td>Insufficient (h-IM)</td>
</tr>
<tr>
<td>Zoledronate continuation vs. discontinuation (Z x 9 yr vs. Z x 6 yr followed by PBO x 3 yr) 1 RCT59</td>
<td>190 PM women previously received zoledronic acid 6 yr for osteoporosis by BMD or RVF</td>
<td>SAE</td>
<td>No difference: RR=0.86 [0.54, 1.36]; ARR= -3 [-16, 9]</td>
<td>Low (IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AFF with radiologic features (n=0 cases)</td>
<td>No cases occurred</td>
<td>Insufficient (h-IM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ONJ (exposed jaw bone &gt;6 weeks) (n=0 cases)</td>
<td>No cases occurred</td>
<td>Insufficient (h-IM)</td>
</tr>
</tbody>
</table>
## Table D. Evidence on harms of osteoporosis drug continuation versus discontinuation* (continued)

<table>
<thead>
<tr>
<th>Comparison # Studies by Design Treatment Duration</th>
<th>Participant +Characteristics</th>
<th>Harms</th>
<th>Relative and Absolute Risk Differences (95% CI)</th>
<th>Strength of Evidence† (Justification)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denosumab continuation vs. discontinuation (D x 4 yr vs. D x 2 yr followed by PBO x 2 yr) 1 RCT¹⁸</td>
<td>314 PM women with osteopenia or osteoporosis by BMD</td>
<td>SAE</td>
<td>No numerical data</td>
<td>Insufficient (no data)</td>
</tr>
</tbody>
</table>

Abbreviations: AFF=atypical femoral fracture; AL=alendronate; ARR=absolute risk reduction; BMD=bone mineral density; BP=bisphosphonate; CI=confidence intervals; D=denosumab; DX=diagnosis; h-IM=highly imprecise; HR=hazard ratio; IM=imprecise; LE=large effects; ONJ=osteonecrosis of the jaw; OR=odds ratio; PBO=placebo; PM=postmenopausal; RCT=randomized controlled trial; RB=medium risk of bias; RR=risk ratio; SAE=serious adverse event; Z=zoledronate

*Discontinuation ≥1 year after prior treatment ≥1 year.
†Definitions of terms for strength of evidence grades and domains ratings are detailed in the section of the main report titled, 'Strength of Evidence for Major Comparisons and Outcomes.'

## Discussion

In long-term placebo-controlled trials, alendronate for 4 years reduced incident radiographic vertebral and nonvertebral fractures in women with osteoporosis, and zoledronate for 6 years reduced vertebral and nonvertebral fractures in women with osteopenia or osteoporosis. Observational studies suggested that long-term treatment with bisphosphonates as a class increased risk of AFF, ST/FSF, and ONJ, though these adverse events were rare. In women with osteoporosis, long-term raloxifene for 4 years reduced incident vertebral fractures, but not hip or nonvertebral fractures; long-term raloxifene also increased risk of deep vein thrombosis and pulmonary embolism. In women with unknown osteoporosis or osteopenia status, oral hormone therapy for 5-7 years reduced incident clinical and hip fractures compared with placebo, but increased risk of cardiovascular disease and cognitive impairment. Trials also showed that continuation of zoledronate or alendronate after 3-5 years of prior treatment versus discontinuation reduced some vertebral fracture outcomes but not others, did not reduce nonvertebral fractures, and observational data suggested that continuation of bisphosphonates as a class may increase risk of ST/FSF compared with discontinuation.

Whereas long-term treatment with alendronate reduced risk of incident clinical fractures compared with placebo in women with osteoporosis, it did not reduce fracture risk in women with osteopenia. Otherwise, risk of fracture with long-term alendronate versus placebo did not vary by history of prior fracture, World Health Organization Fracture Risk Assessment Tool (FRAX*) score, or pretreatment levels of bone turnover markers. Risk of incident fracture between long-term raloxifene and placebo did not vary as a function of age, baseline BMD, or history of prior fracture. Reduction in incident clinical fracture with oral hormone therapy compared with placebo appeared possibly greater in women aged 60-79 years than in younger women, though similar results were not found for hip fracture,
suggesting possible chance findings. We found no information about possible modifiers of fracture risk with long-term zoledronate treatment.

Our findings have several clinical implications. In women with osteoporosis, indications for long-term raloxifene may be limited, because it only reduces vertebral fractures, while both long-term alendronate and zoledronate also reduced nonvertebral fractures. While the effects of long-term alendronate and zoledronate appear roughly similar in older women with osteoporosis, only long-term zoledronate also reduced nonvertebral fractures in women with osteopenia. It is unclear if these possibly discrepant findings are explained in part by differences in study populations (e.g., the zoledronate population was older). Unfortunately, there are no eligible long-term trials that directly compare alendronate and zoledronate in older women with osteopenia. While oral hormone therapies for 5-7 years lowered both clinical and hip fractures in women not selected to be at high fracture risk, and might be expected to have larger effects in those with osteoporosis, because fracture benefits were offset by risk of serious harms, these agents are not likely to be a viable option for long-term osteoporosis treatment. However, it is unknown whether a lower dose or different route of administration of hormone therapy would have a more favorable balance of fracture benefits to harms. In patients who have completed 3-5 years of bisphosphonate treatment, continued alendronate or zoledronate versus discontinuation each reduced one of two measures of incident vertebral fracture, but did not reduce nonvertebral fractures. Observational data suggested that long-term bisphosphonates increase risk of AFF, that risk likely increases with longer duration of treatment, and that these events are rare. Estimating the relative balance between benefits and harms, for every 1,000 women with osteoporosis treated with alendronate for 4 years or with osteopenia or osteoporosis treated with zoledronate for 6 years compared with placebo, approximately 50 to 70 more will avoid an incident clinical fracture, while an additional 2 will experience a ST/FSF.

Since only a minority of ST/FSF meet AFF criteria, the absolute number of additional AFF would be expected to be smaller. Analogously, for every 1,000 women previously treated for osteopenia or osteoporosis with 3 to 5 years of alendronate or zoledronate who continue bisphosphonate treatment another 3 to 5 years, compared with discontinuation, approximately 30 more will avoid an incident vertebral fracture, while an additional 1 will experience a ST/FSF. However, the inconsistency of the vertebral fracture results, the uncertainty around the outcome risk estimates, and the fact that relatively few ST/FSF meet AFF criteria, suggest that the ratio of these fracture benefits to AFF harms with bisphosphonate continuation could be either substantially larger or smaller. Evidence appeared less robust for ONJ, but suggested long-term bisphosphonate treatment also may increase risk of this outcome. Data from eligible studies did not identify clear patient, bone or drug characteristics that modify likelihood of fracture benefits or harms with long-term or continuing osteoporosis drug treatment.

**Limitations**

The available data limit this review in several ways. First, there were few unique trials of long-term osteoporosis drug treatment or of drug discontinuation, and only one trial that included a treatment arm involving osteoporosis drug discontinuation and subsequent osteoporosis drug resumption. We often identified only one trial for a given treatment comparison. Second, only two trials were designed with incident fracture as the primary outcome. Consequently, many studies had few incident clinical fractures, especially for hip fractures, and statistical power often was low to precisely estimate differences in their risk between treatment interventions. Third, all trials were conducted in generally healthy, usually white, postmenopausal women, limiting their generalizability. Further, most of the trials were conducted in populations selected for osteoporosis or osteopenia by BMD or radiographic vertebral fracture criteria. Generalizability of results to
populations who have other reasons for heightened fracture risk is unknown. Fourth, observational studies investigating the association between treatment and risk of AFF or ONJ had marked methodologic differences that likely affected the specificity of these outcomes and the associated risk estimates. Major differences included the definitions of the cases (e.g., whether or not fractures were defined using American Society for Bone and Mineral Research [ASBMR] radiographic AFF features) and noncase controls, drug therapy exposure and control groups, and adjustment for possible confounding. Fifth, reporting on harms was sparse and inconsistent between studies, limiting confidence around harms risk estimates, and raising concerns about possible reporting bias. Sixth, few studies reported information about possible effect modifiers of drug treatment outcomes. These analyses were almost entirely post hoc, often did not test for interactions between potential effect modifiers, treatment assignment, and treatment outcomes, and did not test for multiple testing, raising the likelihood of type 1 errors. Seventh, there were no eligible long-term fracture trials for several U.S. Food and Drug Administration (FDA) approved osteoporosis drugs, including risedronate, ibandronate, teriparatide, and abaloparatide; long-term fracture data for denosumab came from only one small trial in which the pooling of fracture data from different intervention groups made interpretation impossible. There also were no eligible trials of sequential treatment, such as with an anabolic followed by an anti-resorptive, or denosumab followed by bisphosphonate. Finally, there were no usable data comparing different durations of osteoporosis drug holidays.

Research Needs

Future trials of long-term osteoporosis drug treatment and osteoporosis drug continuation versus discontinuation should be designed with adequate statistical power to assess risks of clinical fracture endpoints, including hip fractures, the fracture type with the greatest risk of morbidity and mortality. Broader trial samples that include men, nonwhite women, more adults with comorbidities, and adults aged 80 years and older are needed to improve generalizability. Future long-term trials should evaluate sequential osteoporosis drug treatment, including comparisons of anabolic therapy followed by antiresorptive therapy, and denosumab followed by bisphosphonate therapy, with both compared with continuous long-term antiresorptive therapy. Trials should compare continuous long-term osteoporosis drug treatment to different osteoporosis drug holiday durations, with or without restarting osteoporosis drug therapy, and possibly with repeating cycles of osteoporosis drug therapy alternating with drug holidays. Future studies should systematically collect, analyze and report harms data. Randomized trials will continue to have limited statistical power to estimate the risk of rare treatment harms such as AFF and ONJ, so observational studies will be essential for examining these outcomes. These observational studies should use consensus case definitions, standard non-case and exposure controls, cohort designs to estimate incidence rates, and adequate statistical adjustment to reduce the effects of confounding by indication and selection bias. Observational studies may provide insights about the benefits and harms of drug holidays of different durations and about patient and treatment characteristics that predict which patients are likely to benefit or be harmed by treatment continuation versus discontinuation. Future trials and observational studies should pre-specify analyses to investigate possible effect modifiers of benefits and harms of long-term osteoporosis drug treatment and drug holiday outcomes. Among other factors, these should include age, and BMD and bone markers before and during long-term treatment or drug holidays. Patient-level data from osteoporosis drug trials on the associations of early treatment changes in BMD and bone turnover markers with risk of incident fractures may improve understanding of the potential use and limitations of these measures as surrogates for incident fracture.
Conclusions

Only alendronate, zoledronate, and oral hormone therapy reduced nonvertebral fractures with long-term treatment. However, for all these agents, these fracture benefits were limited to mostly older, postmenopausal women. They were further limited to women with osteopenia or osteoporosis for zoledronate, and to women with osteoporosis for alendronate. Absolute reductions in clinical fractures with long-term bisphosphonates appeared far greater than absolute increases in risk of AFF and ONJ with these treatments. However, reductions in hip fracture with long-term oral hormone therapy appear offset by risk of serious harms. In patients with prior osteoporosis drug treatment, continued treatment appeared to reduce vertebral fractures but not nonvertebral fractures, and may increase risk of AFF. While fracture benefits of continued osteoporosis drug treatment versus drug holiday numerically appeared to outweigh these risks, the more limited morbidity prevented and greater uncertainty about the outcome measures and risk estimates require further investigation to better inform clinical decisions about continuing treatment. This research should include examination of how these benefits and risks vary as a function of patient, bone, and drug treatment characteristics (e.g., age, sex, comorbidity, pre-drug holiday BMD, duration of prior osteoporosis drug treatment). Future modeling studies also may incorporate probabilities of experiencing fracture-related morbidity to help patients weigh trade-offs of treatment more easily.

References


70. Black D. Change in BMD as a Surrogate for Fracture Risk Reduction in Osteoporosis Trials: Results from Pooled, Individual-level Patient Data from the FNIH Bone Quality Project [abstract]. American Society for Bone and Mineral Research Annual Meeting, September 28-October 1, Montréal, Québec, Canada. 2018.

**Full Report**