



# Comparing medication adherence in patients receiving bisphosphonates for preventing fragility fractures: a comprehensive systematic review and network meta-analysis

A. Bastounis<sup>1</sup> · T. Langley<sup>1</sup> · S. Davis<sup>2</sup> · Z. Paskins<sup>3,4</sup> · N. Gittoes<sup>5,6</sup> · J. Leonardi-Bee<sup>1</sup> · O. Sahota<sup>7</sup>

Received: 30 September 2021 / Accepted: 14 February 2022  
© The Author(s) 2022

## Abstract

**Background** Bisphosphonates are effective in preventing fragility fractures; however, high rates of adherence are needed to preserve clinical benefits.

**Objective** To investigate persistence and compliance to oral and intravenous bisphosphonates (alendronate, ibandronate, risedronate, and zoledronate).

**Methods** Searches of 12 databases, unpublished sources, and trial registries were conducted, covering the period from 2000 to April 2021. Screening, data extraction, and risk of bias assessment (Cochrane Collaboration risk-of-bias tool 1.0 & ROBINS-I) were independently undertaken by two study authors. Randomised controlled trials (RCTs) and observational studies that used prescription claim databases or hospital medical records to examine patients' adherence were included. Network meta-analyses (NMA) embedded within a Bayesian framework were conducted, investigating users' likelihood in discontinuing bisphosphonate treatment. Where meta-analysis was not possible, data were synthesised using the vote-counting synthesis method.

**Results** Fifty-nine RCTs and 43 observational studies were identified, resulting in a total population of 2,656,659 participants. Data from 59 RCTs and 24 observational studies were used to populate NMAs. Zoledronate users were the least likely to discontinue their treatment  $HR=0.73$  (95%CrI: 0.61, 0.88). Higher rates of compliance were observed in those receiving intravenous treatments. The paucity of data and the heterogeneity in the reported medication possession ratio thresholds precluded a NMA of compliance data.

**Conclusions** Users of intravenously administered bisphosphonates were found to be the most adherent to treatment among bisphosphonates' users. Patterns of adherence will permit the more precise estimation of clinical and cost-effectiveness of bisphosphonates.

**Trial registration** PROSPERO 2020 CRD42020177166

**Keywords** Adherence · Bisphosphonates · Compliance · Discontinuation · Network meta-analysis

## Abbreviations

ALN	Alendronate	IBN	Ibandronate
BP	Bisphosphonate	iv	Intravenous
CrI	Credible interval	MPR	Medication possession ratio
$d_{Ik}$	Relative effect	OP	Osteoporotic
$D_{res}$	Total residual deviance	OR	Odds ratio
DIC	Deviance Information Criterion	$pD$	Effective number of parameters
HR	Hazard ratio	PDC	Proportion of days covered
		RIS	Risedronate
		SD	Standard deviation
		SE	Standard error
		SUCRA	Surface under the cumulative ranking
		ZOL	Zoledronate
		$\beta$	Regression coefficient
		$\kappa$	Cohen's kappa

Jo Leonardi-Bee and Opinder Sahota are joint seniors.

✉ A. Bastounis  
Anastasios.Bastounis1@nottingham.ac.uk

Extended author information available on the last page of the article

$\sigma$	Between-study standard deviation
$\tau$	Within-class standard deviation

## Introduction

Bisphosphonates (BPs) are the most commonly prescribed treatment for osteoporosis [1] and are broadly regarded as the optimal treatment for adults with osteoporotic fractures [2, 3]. Adherence to BPs is crucial to realise clinical benefits and reduce the risk of fractures. Recent evidence, however, suggests that adherence to oral and parenteral BPs are sub-optimal and they tend to decrease over time [4, 5]. Given that duration of BP treatment is affected by both individual and contextual factors [3, 6], as well as the high heterogeneity observed in adherence rates among BPs, there is a need to undertake a comparative evaluation of adherence among BP users. In this review, adherence is used as an umbrella term which encompasses the following terms [7, 8]: (i) initiation, (ii) implementation, and (iii) discontinuation of treatment. Initiation of treatment refers to the time when people start a prescribed medication (i.e. receive the first dose of a prescribed medication), implementation refers to the level of compliance to the dosing regimen of a prescribed medication from the first to the last dose, and discontinuation refers to the time when people stop taking their prescribed medication [9]. In this review, the term ‘compliance’ has been used as synonymous to ‘implementation’, while the term ‘persistence’ has been used for denoting ‘(non)-discontinuation to treatment’.

This is a systematic review and network meta-analysis which seeks to synthesise the evidence on treatment discontinuation, persistence, and compliance among oral and parenteral BP users. In this review, four BPs were considered: alendronate 10 mg/daily and 70 mg/weekly (ALN), ibandronate 150 mg/monthly (IBN-oral), and ibandronate 3 mg/quarterly intravenously administered (IBN-iv), risedronate 5 mg/daily or 35 mg/weekly (RIS), and zoledronate 5 mg/annually (ZOL). A comparative evaluation of users’ adherence will enable better estimation of BPs’ clinical effectiveness which in turn will inform cost-effectiveness modelling more accurately and facilitate clinical decision-making. This review aims to provide estimates regarding users’ probability to adhere in BPs’ treatment, exploring patterns of discontinuation, persistence, and compliance among people with different clinical profiles.

## Methods

This study was reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [10] and the PRISMA Extension Statement for

Reporting of Systematic Reviews Incorporating Network Meta-analyses of Health Care Interventions checklist [11]. A prospective protocol for the systematic review has been previously published in PROSPERO [CRD42020177166] [12].

## Eligibility criteria

Eligible participants were women and men aged  $\geq 65$  or  $\geq 75$  years and women aged  $\leq 64$  years and men aged  $\leq 74$  years in the presence of risk factors [2] (i.e. previous fragility fracture, current use or frequent recent use of oral or systemic glucocorticoids, a history of falls, a family history of hip fracture, other causes of secondary osteoporosis, BMI lower than 18.5 kg/m<sup>2</sup>, smoking or an alcohol intake of more than 14 units per week in women or more than 21 units per week in men). Studies that recruited mixed populations in terms of gender and other population characteristics were eligible for inclusion. The following four BP medications within their licensed doses [2] were eligible for inclusion: ALN, IBN-iv and IBN-oral, RIS, and ZOL. Relevant comparators included eligible treatments (compared with each other), placebo, or other non-active treatments. Eligible study designs were RCTs, non-randomised parallel comparative studies and retrospective observational studies. Other comparative observational designs (e.g. case-control and cross-sectional studies) were excluded. The outcome of interest were persistence and compliance, quantified either as continuous (e.g., absolute numbers or rates) or discrete measures (e.g., absolute number of participants being persistent/compliant based on pre-specified thresholds). In RCTs, persistence was indirectly inferred by assessing the total number of participants who dropped out at 12 and 24 months; in observational studies, persistence was inferred by assessing the total number of participants who discontinued their treatment based on treatment refill-gaps, using data from claim databases or medical records. In observational studies, compliance was indirectly measured by assessing ‘treatment continuity’ and using percentages/absolute numbers of medication possession ratios (MPR) and number of users with MPRs over a pre-specified threshold. Reports published as abstracts or conference presentations were excluded where insufficient details were reported. Randomised controlled trials which were judged otherwise eligible but did not report outcome data per treatment arm or reported zero dropouts for both arms were also excluded. Studies which reported the outcomes of interest for BP groups collapsed or studies reporting comparisons based solely on the frequency of administration (e.g. daily versus weekly) were also excluded.

## Search strategy and information sources

A comprehensive search strategy was developed to identify eligible studies (Appendix 1). The search strategy comprised the following main elements: searching of electronic databases (including unpublished data and trial registries) and scrutiny of bibliographies of retrieved papers. The following databases were searched:

- MEDLINE(R) In-Process & Other Non-Indexed Citations and MEDLINE(R) (Ovid and PubMed);
- EMBASE (Ovid);
- Cochrane Database of Systematic Reviews (Wiley Interscience);
- Cochrane Central Register of Controlled Trials (CENTRAL) (Wiley Interscience);
- Cumulative Index to Nursing and Allied Health Literature (CINAHL, EBSCO);
- Database of Abstract of Reviews of Effects (Wiley Online Library);
- Health Technology Assessment Database (CRD Database);
- NHS Economic Evaluation Database (CRD Database);
- OpenGrey;
- Science Citation Index (ISI Web of Knowledge);
- Conference Proceedings Citation Index—Science (Web of Science);
- ClinicalTrials.gov(<https://clinicaltrials.gov/>)

Searches covered the period from January 2000 to 5 April 2020. Updated searches in the same databases and trial registries were conducted from 2019 to 25–26 March 2021. All potentially relevant citations were downloaded to Endnote X8 Reference Manager bibliographic software (version 8.0; Clarivate Analytics, Philadelphia, PA, USA).

## Study selection, data collection process, and data items

Selected studies were imported into Rayyan online software [13]. Two independent reviewers screened studies for relevance based on titles/abstracts and later full texts (AB & TL) with disagreements resolved through discussion or by consulting a third reviewer (JLB). Two independent reviewers (AB & TL) conducted full-text screening with a high-level of agreement ( $\kappa=0.84$ ). A pilot-tested data extraction form was used to extract relevant data. One reviewer (AB) extracted data with a second reviewer (TL) independently checking at least 50% of the extracted records. Data extracted consisted of the following categories: (i) descriptive statistics (e.g. number recruited and participants' characteristics), (ii) moderators of action (e.g. glucocorticoids use, patients with osteoporosis, history of fractures/fractures at baseline, medication related

to the incidence of secondary osteoporosis), (iii) treatments' characteristics (e.g. drug-type, administration mode, concomitant treatments), (iv) statistics and relevant data on adherence expressed either as binary (e.g. number of participants who dropped out from RCTs, number of users who discontinued with BP treatment, and number of users with varying compliance levels based on pre-specified thresholds) or continuous outcomes (e.g. mean/range MPR, mean number of infusions/tablet counts, proportion of days covered (PDC) percentage, mean duration of BP treatment).

## Geometry of networks

Treatment-placebo and treatment-active comparisons were visually displayed and network plots were created for all outcomes included in the analyses. For persistence in observational studies (i.e. discontinuation), nodes indicate the different treatments included in the analysis and the thickness of edges connecting the nodes indicates the number of studies informing each comparison (thicker lines indicate more populated comparisons). For persistence in RCTs (i.e. dropouts), nodes indicate the different treatments included in the analysis, the node size indicates the number of studies included in each node and the thickness of the lines indicates the overall sample size informing each comparison (thicker edges indicate more populated pairwise comparisons).

## Risk of bias within individual studies

The methodological quality of the included RCTs was independently assessed at the study level by two reviewers (AB & JLB), using the Cochrane Collaboration risk of bias tool 1.0 [14]. The methodological quality of the included observational studies was independently assessed at the study level by one reviewer (AB) with a second reviewer (JLB) independently checking 50% of the included studies. The assessment of methodological quality in observational studies was undertaken, using the Cochrane Collaboration Risk Of Bias In Non-randomised Studies of Interventions tool (ROBINS-I) [15]. Risk-of-bias plots were created by using the 'robvis' tool [16].

## Summary measures and methods of analysis

### Persistence

In RCTs, the number of dropouts at 12- and 24-month was reported in a binary form (number of participants who dropped out subtracted from the total number of participants per arm). The data generation process was assumed to follow a binomial likelihood while NMAs were modelled using the logit function [17]. Log odds ratios (OR) were estimated from the median and corresponding 95% credibility intervals (CrI) from the 2.5th

and 97.5th centiles of the posterior distribution. In retrospective observational studies, discontinuation was reported in a binary form (number of participants who discontinued the treatment as this is indicated by pre-specified refill-gaps). Given the absence of control conditions in retrospective observational cohorts, ALN was used as the reference treatment. The data generation process was assumed to follow a binomial likelihood. To account for different trial durations, an underlying Poisson process was assumed for each trial arm. The probabilities of any of the aforementioned binary outcomes were considered non-linear functions of event rates, so we modelled the NMAs for the binary outcomes using the complementary log–log link function [18]. Log hazard ratios (HR) were estimated from the median and corresponding 95% credibility intervals (CrI) from the 2.5th and 97.5th centiles of the posterior distribution. Treatment ranking probabilities and surface under the cumulative rankings (SUCRA) are also reported [19]. For studies including ZOL users, meaningful (> 12-month) follow-up assessments were selected and included in the NMA. In case, there were follow-up assessment at 12-month only, ZOL arms were excluded from the NMA.

Standard, independent random (treatment)-effects models were fitted for evaluating users' comparative probability of persistence assessing the total number of dropouts and total number of discontinuations in RCTs and retrospective observational studies respectively [17]. Conventional reference prior distributions were used: (i) trial-specific baseline,  $\mu_1 \sim N(0, 100^2)$ , (ii) treatment effects relative to reference treatment,  $d_{1k} \sim N(0, 100^2)$ , and (iii) between-study SD of treatment effects,  $\tau \sim U(0, 2)$ . All analyses were conducted using OpenBUGS (MRC Biostatistics Unit, Cambridge, UK) [20] and R Studio (R version 4.0.3) [21], using the 'gemtc' and 'rjags' packages. Convergence to the target posterior distributions was assessed using the Gelman–Rubin statistic for three independent chains with different initial values. For all outcomes, results were based on three independent chains of initial values and 60,000 iterations after a burn-in of 40,000 iterations. Most of NMAs exhibited moderate correlation between successive iterations of the Markov chain, so were thinned by retaining every 5<sup>th</sup> sample.

### Vote-counting synthesis on persistence and compliance

For those observational studies which were not included in the discontinuation NMA, effect sizes of discontinuation or persistence were summarised using the vote-counting synthesis method based on the direction of effects [22]. Similarly, data on compliance drawn from both RCTs and observational studies were summarised based on the vote-counting synthesis method. Findings from both syntheses are presented using cross-study visual displays [23].

### Assessment of inconsistency

Consistency of evidence for NMAs on persistence using dropout data from RCTs was assessed, using the node-splitting method [24] in RStudio (R version 4.0.3). Differences between direct and indirect evidence in all network loops were calculated with  $p$  values < 0.05 indicating the presence of significant inconsistency. Due to the multiple arms reported per study in retrospective observational studies, a formal assessment of inconsistency was not performed.

### Credibility of the findings

Credibility of findings on persistence was assessed in RCTs only by following the CINeMA approach [25], where the credibility of findings is accounted for by the assessment of: (i) within-study bias, (ii) reporting bias, (iii) indirectness, (iv) imprecision, (v) heterogeneity, and (vi) incoherence. Conventional levels of OR ranging from 0.8 to 1.25 were used to indicate clinical significance. CINeMA's freely available web application [26] was used to assess credibility of findings.

### Additional analysis

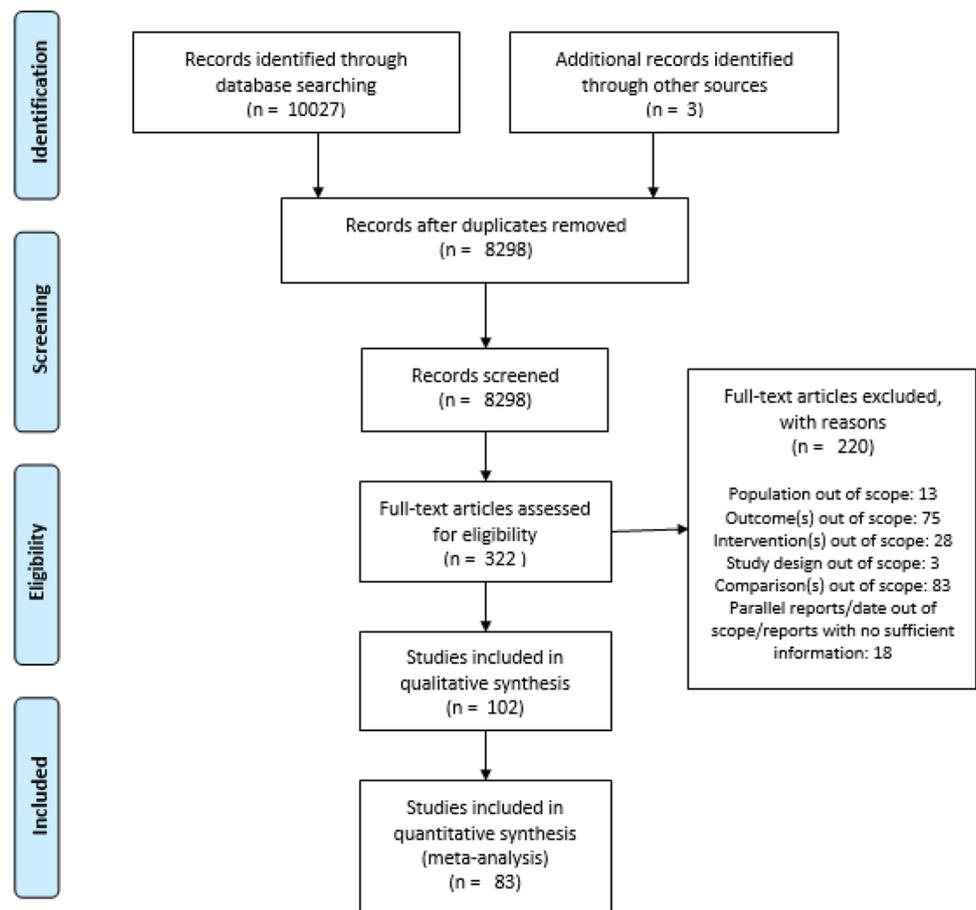
Sensitivity analysis was conducted for measures of the number of dropouts in RCT designs only. Studies with an overall high risk of bias were excluded in the sensitivity analyses. Heterogeneity in treatment effects was explored by considering potential treatment effect modifiers. A set of subgroup meta-regressions were conducted on the dropouts outcome, testing the effects of the following three covariates: (i) proportion of patients  $\geq 75\%$  with osteoporosis, (ii) proportion of patients  $\geq 75\%$  at increased fracture risk, and (iii) mode of administration (oral versus intravenous). In all subgroup analyses, we assumed a common interaction effect that applies to the relative effects of all the treatments relative to the reference treatment [27].

## Results

### Study selection

A PRISMA flow diagram shows the selection of papers for inclusion and exclusion (Fig. 1). A total of 10,030 articles were retrieved, of which 1,729 were duplicates. Overall, 7,976 studies were excluded following title and abstract screening, and 220 were excluded following the full-text screen. A reference list of studies included in this review is reported in the appendices (Appendix 12). Data were extracted from 59 RCTs drawn from 69 published reports [1–69] and 43 observational studies drawn from 45

**Fig. 1** PRISMA flow diagram of the selected studies (the number of records retrieved in each database are reported in Appendix 1)



published reports [70–114], resulting in a total population of 2,656,659 participants. Overall, for persistence, 16,577 participants were included in the NMAs of RCTs and 985,484 BP users were included in the NMA of retrospective observational studies.

### Networks' structures and geometry

Two network plots comparing BP effects on the absolute numbers of dropouts were created (Appendix 3). Data on dropouts at 12 months provided six closed loops of evidence. Overall, 30 two-arm and one three-arm studies were included in the analysis, resulting in a total of 10,419 participants. The most studied treatment was ALN ( $n_{\text{studies}} = 16$ ), while PLB was used as a comparator in 24 studies. Data on dropouts at 24 months provided four closed loops of evidence. Overall, 21 two-arm and one three-arm studies were included in the analysis, resulting in a total of 6,158 participants. The most studied treatment was ZOL ( $n_{\text{studies}} = 10$ ), while PLB was used as a comparator in 19 studies. One network was created for discontinuation data drawn from observational studies. Overall, eight two-arm, 12 three-arm, and four five-arm studies were included in the analysis. The most commonly

treatments were ALN and RIS, with each contributing 23 arms in the analysis.

### Studies characteristics and risk of bias within individual studies

Overall, 59 trials and 43 observational studies were included in this review (Appendix 2). Of the included trials, 31 trials were conducted in North America or in multiple countries [1,3,6,8,9,10,12,13,16,18,20–23,36–41,43–45,48–50,54,55,61,63,65]. Overall, 38 trials exclusively targeted female participants [1–3,6,8,12,13,15,16,20–24,27–30,32,33,35,37–41,46,47,49,55,57–59,61,64,65,67,68], while in 21 trials, most of the participants fulfilled the criteria of osteoporosis [3,6,8,12,13,15,16,20,27,28,33,37,41,42,52,57,60,61,63,64,67]. In total, the overall risk of bias was high in 18 trials [12,13,15,16,27,28,30,34,46,54,55,58,59,63–65,67,69] (Appendix 6). The majority of observational studies adopted a retrospective design, while three of them adopted a prospective comparative design [88,98,99]. Twenty observational studies were conducted in Europe [71,73,74,80,84,87,88,90,92,94,96,98,101,102,104–106,109,113,114] and 18 studies were conducted in the USA [70,76–79,81–83,86,89,95,97,100,103,108,110–112]. In 14 studies, the majority of participants

fulfilled the criteria of osteoporosis [72,80,82,86,88,89,97,98,100,105,106,110,112,113]. Overall, three studies received a moderate risk of bias rating [88,97,99], two received a critical risk of bias rating [80,111], and the rest received a serious risk of bias rating.

### Synthesis of results on persistence: measured using dropouts in RCTs

A NMA was used to compare the effects of ALN, RIS, ZOL, and IBN-oral relative to PLB on the total number of dropouts at 12 months. Overall, data were available from 30 two-arm and one three-arm RCTs. The network provided nine direct treatment comparisons. Each of the direct comparisons between ALN versus ZOL and RIS versus IBN-oral were informed by one small study. Eight contrasts were checked for inconsistency between direct and indirect evidence. None of the comparisons showed significant evidence of inconsistency, as assessed using Bayesian  $p$  values ( $p > 0.05$ ) (Appendix 8). The model fitted the data relatively well (difference  $< 3$ ), with a total residual deviance of 64.8 being close to the number of data points included in the analysis, which was 63 ( $DIC = 369$ ). The between-study SD was estimated to be 0.16 (95%CrI: 0.009, 0.41), implying mild heterogeneity in treatment effects between RCTs. Users of ZOL and RIS were less likely to dropout compared to PLB users with none of these effects being statistically significant (Table 1). The lowest likelihood of dropping out was detected in ZOL

users  $OR = 0.73$  (95%CrI: 0.51, 1.05; probability: 0.88; SUCRA: 0.95) (Appendix 4).

A NMA was used to compare the effects of ALN, RIS, ZOL, and IBN-oral relative to PLB on the total number of dropouts at 24 months. Overall, data were available from 21 two-arm and one three-arm RCTs. The network provided eight direct treatment comparisons. Each of the direct comparisons between ALN versus IBN, ALN versus RIS, and RIS versus IBN-oral were informed by one small study. Three contrasts were checked for inconsistency between direct and indirect evidence. None of the comparisons showed significant evidence of inconsistency, as assessed using Bayesian  $p$  values ( $p > 0.05$ ) (Appendix 8). The model fitted the data well with a total residual deviance of 45.51 being close to the number of data points included in the analysis, which was 45 ( $DIC = 276.7$ ). The between-study SD was estimated to be 0.34 (95%CrI: 0.09, 0.64), implying mild to moderate heterogeneity in treatment effects between RCTs but with reasonable uncertainty. Users of ALN, RIS, and IBN-oral were less likely to dropout compared to PLB users with none of these effects being statistically significant (Table 1). The lowest likelihood of dropping out was detected in IBN-oral users  $OR = 0.72$  (95%CrI: 0.31, 1.66; probability: 0.54; SUCRA: 0.72) (Appendix 4).

**Table 1** League table presenting network meta-analysis estimates (lower triangle) and direct estimates (upper triangle) regarding BPs effectiveness on persistence as measured using dropouts data from RCTs and discontinued treatment data from observational studies. From the left to the right: (i) number of participants who dropped out from RCTs at 12 months, (ii) number of participants who dropped out

from RCTs at 24 months, (iii) number of participants who discontinued BPs treatment in observational studies. Posterior ORs (95%CrI) are reported in persistence (drop-out) NMAs of RCTs and posterior median HRs (95%CrI) are reported in persistence (discontinuation) NMA of retrospective observational studies.

ZOL		0.81 (0.47, 1.39)				IBN-or			-		ZOL		-		
0.76 (0.53, 1.11)	RIS	0.97 (0.66, 1.5)				0.86 (0.37, 2.05)	RIS		0.84 (0.55, 1.3)		0.86 (0.68, 1.09)	IBN-iv	-		
0.73 (0.51, 1.05)	0.95 (0.72, 1.26)	Placebo				0.84 (0.36, 2)	0.97 (0.56, 1.74)	ALN	0.78 (0.41, 1.52)		0.73 (0.61, 0.88)	0.84 (0.68, 1.06)	ALN		
0.69 (0.49, 1.02)	0.9 (0.69, 1.2)	0.94 (0.75, 1.2)	ALN			0.72 (0.33, 1.66)	0.84 (0.57, 1.24)	0.86 (0.53, 1.37)	Placebo		0.73 (0.6, 0.88)	0.84 (0.67, 1.06)	0.99 (0.87, 1.13)	IBN-or	
0.64 (0.39, 1.02)	0.83 (0.54, 1.23)	0.87 (0.6, 1.24)	0.92 (0.63, 1.27)	IBN-or		0.67 (0.28, 1.68)	0.78 (0.47, 1.31)	0.8 (0.47, 1.33)	0.92 (0.66, 1.32)	ZOL	0.67 (0.56, 0.8)	0.77 (0.62, 0.97)	0.91 (0.82, 1.01)	0.91 (0.8, 1.04)	RIS

**Note.** ALN = Alendronate; IBN-iv: Ibandronate 3mg intravenous; IBN-or: Ibandronate 150mg oral; RIS: Risedronate; ZOL: Zoledronate.

Treatments are reported in order of relative ranking for efficacy. Comparisons between treatments should be read from left to right, and their odds ratio/hazard ratio is in the cell in common between the column-defining treatment and the row-defining treatment. Odds ratios (OR) and hazard ratios (HR)  $< 1$  favour the column-defining treatment for the network estimates and the row-defining treatment for the direct estimates.

## Synthesis of results on persistence: measured using discontinuation of treatment data from observational studies

A NMA was used to compare the effects of RIS, ZOL, IBN-oral, and IBN-iv relative to ALN on the absolute number of people who discontinued their BP treatments. Overall, data were available from 24 retrospective observational studies. The model fitted the data well with a total residual deviance of 73.27 being close to the number of data points included in the analysis, which was 73 ( $DIC=762.5$ ). The between-study SD was estimated to be 0.24 (95%CrI: 0.19, 0.31), implying mild heterogeneity in treatment effects between observational studies with reasonably uncertainty. Users of ZOL and IBN-iv were less likely to discontinue compared to ALN with the effects of the former being statistically significant (Table 1). The lowest likelihood for discontinuation was detected in ZOL users  $HR=0.73$  (95%CrI: 0.61, 0.88; probability: 0.88; SUCRA: 0.97) (Appendix 4). Heterogeneity of effects was explored by undertaking a post hoc meta-regression on the absolute number of discontinuers using refill-gap as a moderator variable. Although slightly decreased in magnitude, the direction of effects remained the same. The model fit remained almost the same with a total residual deviance of 72.63 ( $DIC: 755.9$ ). The between-study SD was estimated to be 0.29 (95%CrI: 0.20, 0.42), implying mild heterogeneity in treatment effects between RCTs with reasonable uncertainty. Higher medication effects on discontinuation were detected in participants with longer refill gap thresholds, although the results were not statistically significant  $\beta = -0.23$  (95%CI:  $-0.72, 0.21$ ).

### Additional analysis—persistence using data from RCTs

Heterogeneity of effects was explored by undertaking sensitivity analysis on persistence in RCTs (i.e. number of dropouts) using risk of bias assessment as a moderator variable (Appendix 5). For persistence using the absolute number of dropouts at 12 months, data were available from 20 two-arm trials. The model had a good fit with the data with a total residual deviance of 41.25 being close to the number of data points included in the analysis, which was 40 ( $DIC=237.5$ ). The between-study SD was estimated to be 0.2 (95%CrI: 0.008, 0.57), implying mild heterogeneity in treatment effects between RCTs with reasonably uncertainty. The direction of the findings remained the same compared to the main analysis with the only exception being for users of IBN-oral. None of the treatment effects was significantly different compared to PLB ( $p > 0.05$ ). Users of ZOL were the least likely to dropout compared to participants of PLB  $OR=0.88$  (95%CrI: 0.54, 1.42). For the total number of dropouts at 24 months, data were available from

18 two-arm trials. The model had a good fit with the data with a total residual deviance of 37.16 being close to the number of data points included in the analysis, which was 36 ( $DIC=224.7$ ). The between-study SD was estimated to be 0.43 (95%CrI: 0.13, 0.83), implying moderate heterogeneity in treatment effects between RCTs. The direction of the findings remained the same compared to the main analysis, although their magnitude was slightly decreased. None of the treatment effects was significantly different compared to PLB ( $p > 0.05$ ). Users of oral ibandronate were found to be the least likely to dropout compared to participants on PLB,  $OR=0.57$  (95%CrI: 0.09, 3.18). Heterogeneity of effects of persistence was explored by undertaking separate meta-regression analyses on the absolute number of dropouts at 12 and 24 months using osteoporosis diagnosis and history of or prevalent fractures at baseline as effect modifiers. None of the tested effect modifiers was found to significantly interact with the treatment effects while in none of the cases, the model fit was improved compared to the main NMAs (Appendix 5).

### Vote-counting synthesis on persistence

Overall, 12 observational studies were included in the vote-counting synthesis on persistence with eight studies providing data for ALN versus RIS [72,78,80,84,106,107,111,113], two for RIS versus IBN-oral [86,111], one for RIS versus PLB [99], one for ALN versus IBN-oral [111], one for ALN versus IBN-iv [88], and one for ZOL versus IBN-iv [77] comparisons (references are reported in Appendix 11). There was mixed evidence regarding the comparison between ALN and RIS with four studies favouring ALN users [72,80,107,111], while four studies favoured RIS users [78,84,106,113]. Data expressed in years showed comparable persistence rates between ALN and RIS users [71,79,90,98,114]. In a post hoc sensitivity analysis of ALN versus RIS comparison, we restricted our synthesis only on the weekly administration for both BPs. From a total of five studies, in four studies, ALN users tend to be more persistent than RIS ones with these effects being statistically significant [74,83,107,114]. Data from two studies [86,111] showed that RIS and ALN users tend to be more persistent compared to IBN-oral users, with these effects being statistically significant. One study showed that IBN-iv users are more persistent through time compared to ALN users [88], while ZOL users were found to be more persistent than IBN-iv users [77].

### Vote-counting synthesis on compliance

Overall, 29 studies were included in the vote-counting synthesis of compliance data. From 11 RCTs, five RCTs provided data on the comparison of PLB versus ALN

[14,23,60,63,69]. In four trials, PLB participants were found to be more compliant, although these effects were not statistically significant [23,60,63,69]. Three trials provided data on the comparison of PLB versus RIS, with all of them favouring RIS participants; however, none of these effects was statistically significant [10,23,65]. One trial provided data on the comparison of PLB versus ZOL, with PLB participants being more compliant and this effect was statistically significant [3]. One trial provided data on the comparison between PLB and IBN-oral, with PLB participants being more compliant although these effects were not statistically significant [43]. One trial provided data for the comparison between ALN and ZOL, with the ZOL participants being more compliant, although the effects were not statistically significant [44]. The only three-arm trial provided data for IBN-oral, RIS, and ALN, with IBN-oral participants being the most compliant, although these effects were not statistically significant [46].

Overall, 18 observational studies included in the vote-counting synthesis on compliance with 14 studies providing data on the comparison of ALN versus RIS. In eight studies, ALN users were found to be more compliant than RIS ones with effect sizes in six studies being statistically significant [76,79,81–83,94]. In six studies, RIS users were found to be more compliant than ALN users with effect sizes drawn from two studies being statistically significant [84,103]. In two studies [89,114], compliance was expressed as percentage ranges of MPR. In one study [89], comparable mean MPRs were observed between the two BPs, while higher compliance rates were observed in RIS users in the third study [114]. Eight studies provided data on the comparison between ALN and IBN-oral and six of them were included in the vote-counting synthesis. In four studies [76,82,83,101], ALN users were found to be more compliant than IBN-oral users with three of these providing statistically significant effect sizes [82,83,101]. In two studies [81,94], IBN-oral users were found to be more compliant while in one study, the observed effect size was statistically significant [94]. In two studies [89,108], higher mean MPR rates were observed in the IBN-oral users. In one study, which compared IBN-oral versus RIS users, participants in the latter were found to be more compliant with the effects being statistically significant [85,86]. Two studies provided data on the comparison of ALN versus IBN-iv with mixed evidence [81,88]. Two studies provided data on the comparison of IBN-oral versus IBN-iv with participants in the latter being more compliant with statistically significant effects [97,114]. In one study, ZOL users were found to be more compliant compared to ALN users with the effects being statistically significant [81]. Direct comparison between ZOL and IBN-iv users found that users of the former were more compliant in terms of the mean number of infusions received [77].

## Discussion

This is a systematic review with network meta-analyses and vote-counting synthesis, assessing the probability of adherence to BP treatment administered for preventing fragility fractures. For persistence, results from the NMA from RCTs showed that ZOL users may be less likely to dropout from trials at 12 months, although these effects were marginally non-significant. Results from the NMA using data from the observational studies showed that ZOL and IBN-iv users were less likely to discontinue their treatment over time, with ZOL users being statistically significantly more persistent compared to oral BPs users. Data drawn from the vote-counting synthesis were in line with the results of NMAs, where ZOL and IBN-iv users were more likely to persist with their treatment, with ZOL users being more persistent compared to their IBN-iv counterparts. Users of ALN and RIS showed comparable persistence rates; however, when we restricted our analysis to weekly administration, ALN users were found to be more likely to persist to treatment over time. Due to the paucity of data and the heterogeneity in reporting compliance data, we were unable to perform NMAs, but synthesis based on vote counting found that compliance to ZOL is greater within 24 months after the initiation of their treatment. Users of IBN-iv were found to be more compliant compared to the IBN-oral users. Users of ALN were found to be more compliant than RIS users, while mixed evidence were observed in the comparison between ALN and IBN-oral users.

These findings have important implications for clinical practice and future research. In general, persistence to BP treatment was found to decrease after 12 months, stressing the need to address adherence barriers according to BP treatment and people with different clinical profiles. Nevertheless, ZOL users were found to be less likely to discontinue their treatment over time and they showed higher compliance rates. These findings are partly in line with the results from the drop-out NMA at 12 months. Without ignoring the interplay of individual and contextual factors, which affect participation and adherence in clinical trials [28, 29], we can assume that most ZOL users are likely to receive at least two infusions before discontinuing their treatment. The use of ZOL has been generally recommended for at least three years [30] and, although reduced adherence rates have been observed in ZOL users after the first year [31], simpler drug regimens can improve adherence rates [32]. Results of vote-counting synthesis on oral BPs were partly in line with the NMA results. Alendronate and RIS users showed comparable persistence rates; however, ALN users were found to be more compliant than their RIS counterparts. When we restricted our synthesis

to weekly administration of both BPs, weekly ALN users were found to be more persistent to treatment compared to RIS weekly users. Given that ALN is the most widely prescribed medication, clinical decision-making should consider, alongside its clinical effectiveness, ways in which ALN users could be assisted to receive medication properly and remain on treatment long-term.

## Strengths and limitations

This systematic review has several strengths. First, this review includes a robust search strategy, covering a wide range of databases, trial registries, and grey literature. Second, this systematic review employed gold-standard methods in conducting, reporting, and assessing the credibility of findings. Third, this systematic review included both RCTs and observational studies, adopting a combined approach to synthesise data. Inevitably, this review has also some limitations. First, participants' persistence on BP treatments in RCTs was assessed by using the total number of dropouts as a proxy measure. Given that this was the only way to capture discontinuers in RCTs, dropout NMA findings should be interpreted under this limitation. Second, persistence in the NMA of observational studies was assessed as the absolute number of discontinuers per BP treatment based on varying refill-gaps and without accounting for the censored follow-up time. Third, due to the scarcity of data on males, a subgroup analysis of persistence rates between males and females was not conducted. Fourth, compliance was indirectly assessed by measuring treatment continuity based on different measures. Although a NMA would be more informative, vote-counting synthesis is well-suited in the presence of incomplete and highly-heterogeneous data [23] in both observational studies and RCTs. Fifth, this review did not assess the comparative effectiveness of bisphosphonates against monoclonal antibody and anabolic medications. Sixth, the paucity of data precluded the subgroup analysis between participants receiving bisphosphonates for primary prevention and those receiving bisphosphonates for secondary prevention purposes.

## Conclusions

Adherence was higher in intravenously administered BP users. Clinical decision-making could be facilitated by taking into account adherence patterns in BP users who are at increased risk of fractures.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00198-022-06350-w>.

**Funding** This work was supported by the National Institute for Health Research (NIHR), Health Technology Assessment (HTA) Programme [NIHR127550]. Zoe Paskins is funded by the NIHR, Clinician Scientist Award (CS-2018–18-ST2-010)/NIHR Academy.

## Declarations

**Competing interests** The authors declare no competing interests.

**Disclaimer** This report presents independent research commissioned by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, MRC, CCF, NETSCC, the Health Technology Assessment (HTA) programme or the Department of Health.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

## References

1. Compston JE, McClung MR, Leslie WD (2019) Osteoporosis. *The Lancet* 393:364–376
2. National Institute for Health and Care Excellence (2020) Osteoporosis: assessing the risk of fragility fracture (NICE Clinical Guideline [CG146]) Available from: <https://www.nice.org.uk/guidance/cg146/chapter/2-Research-recommendations> [Assessed at 12 July 2021]
3. Qaseem A, Forciea MA, McLean RM, Denberg TD (2017) Treatment of low bone density or osteoporosis to prevent fractures in men and women: a clinical practice guideline update from the American college of physicians. *Ann Intern Med* 166(11):818–839
4. Fatoye F, Smith P, Gebrye T, Yeowell G (2019) Real-world persistence and adherence with oral bisphosphonates for osteoporosis: a systematic review. *BMJ Open* 9(4):e027049
5. Koller G, Goetz V, Vandermeer B, Homik J, McAlister FA, Kendler D, Ye C (2020) Persistence and adherence to parenteral osteoporosis therapies: a systematic review. *Osteoporosis International* 1–0
6. Yeam CT, Chia S, Tan HC, Kwan YH, Fong W, Seng JJ (2018) A systematic review of factors affecting medication adherence among patients with osteoporosis. *Osteoporos Int* 29(12):2623–2637
7. Vrijens B, De Geest S, Hughes DA, Przemyslaw K, Demonceau J, Ruppert T, Dobbels F, Fargher E, Morrison V, Lewek P, Matyjaszczyk M (2012) A new taxonomy for describing and defining adherence to medications. *Br J Clin Pharmacol* 73(5):691–705

8. De Geest S, Zullig LL, Dunbar-Jacob J, Helmy R, Hughes DA, Wilson IB, Vrijens B (2018) ESPACOMP medication adherence reporting guideline (EMERGE). *Ann Intern Med* 169(1):30–35
9. Hiligsmann M, Cornelissen D, Vrijens B, Abrahamsen B, Al-Daghri N, Biver E, Brandi ML, Bruyère O, Burlet N, Cooper C, Cortet B (2019) Determinants, consequences and potential solutions to poor adherence to anti-osteoporosis treatment: results of an expert group meeting organized by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO) and the International Osteoporosis Foundation (IOF). *Osteoporos Int* 30(11):2155–2165
10. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R (2021) PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 372
11. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C et al (2015) The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 162(11):777–784
12. Bastounis A, Leonardi-Bee J, Langley T, Paskins Z, Davies S, Sahota O (2020) Comparing treatment adherence in patients receiving alendronate and alternative bisphosphonates for preventing fragility fractures: a comprehensive systematic review and meta-analysis. PROSPERO CRD42020177166. Available from: [https://www.crd.york.ac.uk/PROSPERO/display\\_record.php?RecordID=177166](https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=177166)
13. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A (2020) Rayyan — a web and mobile app for systematic reviews. *Syst Rev* 5:210
14. Higgins JT, Altman D, Group CSM, Group CBM (2011) Chapter 8: Assessing risk of bias in included studies. *Cochrane Handbook for Systematic Reviews of Interventions*. In *Cochrane handbook for systematic reviews of interventions*. The Cochrane Collaboration and John Wiley & Sons Ltd; Chichester, UK
15. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, Henry D, Altman DG, Ansari MT, Boutron I, Carpenter JR (2016) ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 355
16. McGuinness LA, Higgins JP (2021) Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Res Synth Methods* 12(1):55–61
17. Dias S, Ades AE, Welton NJ, Jansen JP, Sutton AJ (2018) Chapter 4. Generalised Linear Models. In *Network Meta-Analysis for Decision-Making*. Oxford, UK: John Wiley & Sons Ltd
18. Dias S, Welton NJ, Sutton AJ, Ades AE (2012) NICE DSU technical support document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. (Technical Support Document in Evidence Synthesis; No. TSD2). National Institute for Health and Clinical Excellence
19. Salanti G, Ades AE, Ioannidis JP (2011) Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 64(2):163–171
20. Lunn D, Spiegelhalter D, Thomas A, Best N (2009) The BUGS project: evolution, critique and future directions. *Stat Med* 28(25):3049–3067
21. RStudio Team (2020) RStudio: Integrated Development for R. RStudio, PBC, Boston, MA
22. McKenzie, JE & Brennan, SE (2019) Chapter 12: Synthesizing and presenting findings using other methods. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (eds) *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons. Available from: <https://training.cochrane.org/handbook/current>.
23. Thomson HJ, Thomas S (2013) The effect direction plot: visual display of non-standardised effects across multiple outcome domains. *Res Synth Methods* 4(1):95–101
24. van Valkenhoef G, Dias S, Ades AE, Welton NJ (2016) Automated generation of node-splitting models for assessment of inconsistency in network meta-analysis. *Res Synth Methods* 7(1):80–93
25. Nikolakopoulou A, Higgins JP, Papakonstantinou T, Chaimani A, Del Giovane C, Egger M, Salanti G (2020) CINeMA: An approach for assessing confidence in the results of a network meta-analysis. *PLoS medicine* 17(4):e1003082
26. Papakonstantinou T, Nikolakopoulou A, Higgins JP, Egger M, Salanti G (2020) CINeMA: Software for semiautomated assessment of the confidence in the results of network meta-analysis. *Campbell System Rev* 16(1):e1080
27. Dias S, Sutton AJ, Welton NJ, Ades AE (2013) Evidence synthesis for decision making 3: heterogeneity—subgroups, meta-regression, bias, and bias-adjustment. *Med Decis Making* 33(5):618–640
28. Noirmain C, Gil-Wey B, Pichon I, Brindel P, Haller G (2020) Factors associated with patient willingness to participate in anaesthesia clinical trials: a vignette-based cross-sectional study. *BMC Med Res Methodol* 20(1):1
29. Avis NE, Smith KW, Link CL, Hortobagyi GN, Rivera E (2006) Factors associated with participation in breast cancer treatment clinical trials. *J Clin Oncol* 24(12):1860–1867
30. Compston J, Cooper A, Cooper C, Gittoes N, Gregson C, Harvey N, Hope S, Kanis JA, McCloskey EV, Poole KE, Reid DM (2017) UK clinical guideline for the prevention and treatment of osteoporosis. *Arch Osteoporos* 12(1):43
31. Curtis JR, Yun H, Matthews R, Saag KG, Delzell E (2012) Adherence with intravenous zoledronate and intravenous ibandronate in the United States Medicare population. *Arthritis Care Res* 64(7):1054–1060
32. Selak V, Elley CR, Bullen C, Crengle S, Wadham A, Rafter N, Parag V, Harwood M, Doughty RN, Arroll B, Milne RJ (2014) Effect of fixed dose combination treatment on adherence and risk factor control among patients at high risk of cardiovascular disease: randomised controlled trial in primary care. *BMJ* 27:348

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

A. Bastounis<sup>1</sup>  · T. Langley<sup>1</sup> · S. Davis<sup>2</sup> · Z. Paskins<sup>3,4</sup> · N. Gittoes<sup>5,6</sup> · J. Leonardi-Bee<sup>1</sup> · O. Sahota<sup>7</sup>

T. Langley  
tessa.langley@nottingham.ac.uk

S. Davis  
s.davis@sheffield.ac.uk

Z. Paskins  
z.paskins@keele.ac.uk

N. Gittoes  
Neil.Gittoes@uhb.nhs.uk

J. Leonardi-Bee  
jo.leonardi-bee@nottingham.ac.uk

O. Sahota  
opinder.sahota@nuh.nhs.uk

<sup>2</sup> School of Health and Related Research, Regent Court (ScHARR), University of Sheffield, Sheffield S1 4DA, UK

<sup>3</sup> School of Medicine, Keele University, Keele ST5 5BG, UK

<sup>4</sup> Haywood Academic Rheumatology Centre, Midlands Partnership NHS Foundation Trust, Stoke-on-Trent, Staffordshire, UK

<sup>5</sup> Centre for Endocrinology, Diabetes and Metabolism (CEDAM), University of Birmingham, Birmingham, UK

<sup>6</sup> Queen Elizabeth Hospital, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2TH, UK

<sup>7</sup> Queens Medical Centre (QMC), University of Nottingham, Nottingham University Hospitals NHS Trust, Nottingham NG7 2UH, UK

<sup>1</sup> Division of Epidemiology & Public Health, School of Medicine, University of Nottingham, City Hospital, Nottingham NG5 1PB, UK