Anabolic Therapies for Osteoporosis in Postmenopausal Women: Effectiveness and Value

Response to Public Comments on Draft Evidence Report

June 16, 2017

Contents

Comment	Response/Integration
Amgen	
See note at beginning of full submission on FDA approval and romosozumab.	We have removed romosozumab from the network meta-analysis, ICER rating of comparative clinical effectiveness, and all cost analyses.
ICER's osteoporosis assessment has serious methodological flaws that compromise its results, which inappropriately imply overall poor value of bone-forming agents. For example, ICER selected an inappropriate comparator for this assessment despite extensive feedback on this issue. ICER's comparative clinical effectiveness is based on a literature review that does not include one of the comparators, zoledronic acid (ZA), in the search strategy (draft report tables A2-A4). Moreover, ICER's base case cost-effectiveness model utilizes clinically unsound efficacy assumptions and data inputs, and does not reflect the uncertainty associated with efficacy estimates and their impact on the results and conclusions. Summarized below are the critical issues and recommendations, and how to address them based on an understanding of economic evaluation; clinical practice; the biology of osteoporosis; and of patients suffering its consequences.	Comparative effectiveness analyses often compare drugs across classes. For example, antihypertensive drugs and diabetes drugs are often directly compared despite having differing mechanisms of action when they are used for the same indication. The first line therapy for the average woman with osteoporosis is an oral bisphosphonate. Parenteral agents like zoledronic acid and the anabolic drugs are reserved for women with particularly high risk for fracture and those who cannot tolerate oral agents. During the initial comment period on the draft scope, we received feedback from pharmaceutical manufacturers, professional societies, and clinician experts that zoledronic acid was the appropriate comparator and we changed our comparator from alendronate to zoledronic acid to reflect that input.
1. Comparing bone-forming agents to bisphosphonates is an inappropriate way to estimate the value of bone-forming agents. Despite early feedback from multiple stakeholders, ICER continues to base their value assessment of bone-forming agents (teriparatide, abaloparatide, and romosozumab) on a comparison to a bisphosphonate. ICER selected ZA as the comparator, with the rationale that this agent is used in patients at high risk for fracture. However, this comparison is fraught with limitations. Bisphosphonates, including ZA, are a different class of agents that slow bone loss rather than building new bone, and are generally used in a treatment context that differs from bone-forming agents.	See prior response.
Bone-forming agents are viewed as a distinct class of therapy by the medical community. Although ICER correctly notes treatment recommendations of a T-score ≤ -2.5 or 10-year fracture risk based on FRAX (hip fracture risk of ≥ 3% or major osteoporosis-related fracture risk of ≥ 20%), patients who receive the bone-forming agent, teriparatide, tend to be at a much higher fracture risk relative to patients treated with antiresorptive agents. Real world evidence shows that patients receiving teriparatide were significantly older, had more comorbidities and fracture-related hospitalizations and	We agree that patients who receive zoledronic acid and the anabolic agents tend to be at higher risk for fracture compared to those patients treated with other therapies for osteoporosis. However, there is no clear definition of the risk level at which it is appropriate to initiate therapy with anabolic drugs. We hope that the discussion during the CTAF meeting can define clear criteria based on risk factors (bone density, age, prior

substantially higher baseline fracture rates. In these higher-risk patients, bone-forming agents can improve impaired bone mass and structure allowing for more rapid offset of fracture risk. Subsequent sequencing to antiresorptive agents may help maintain or augment gains in new bone and continue fracture reduction over the long-term. Prior fracture history, lower BMD, and other co-morbidities are features reflecting higher fracture risk. Amgen is conducting research that will further identify patients who are at high risk of a near-term fracture and can provide additional information on this.	fracture history, corticosteroid use, etc.) or other criteria to define the population of women at high enough risk to warrant initiation of parenteral therapy.
ICER compares bone-forming agents to a bisphosphonate requiring making a comparison across different classes of agents, generally used in different treatment contexts, in different patients and over different timeframes. This indicates a lack of recognition of patients' heterogeneity in their needs and preferences and it seems more a misleading price-centric comparison than one informing a relevant decision. Furthermore, ICER also compares active treatments to no treatment, which again represents an unrealistic scenario where bone-forming agents may be considered and yet does not compare bone-forming agents to each other, which would be a more useful exercise of value assessment.	As noted above, it is common practice to compare drugs across classes when they share a common indication. As practicing clinicians, we recognize the heterogeneity of patient characteristics and preferences and how that is at the core of shared decision-making about therapy. The goal of the drive towards personalized medicine is to identify individual patient characteristics that define the best therapy for that patient. We hope that those characteristics can be defined at the CTAF meeting. All the active agents were compared to each other in the network meta-analysis and in the cost analyses. For example, Tables 5 and 7 in the revised report as well as Tables 18, 23, and 25. We have added a sentence in the updated summary section of the evidence review highlighting the insufficiency of the evidence to distinguish between teriparatide and abaloparatide.
Recommendation: Value assessments should compare newer therapies to the most relevant comparator being used in the same context, with the same therapeutic objective in the same population. In this case, a comparison across bone-forming agents would be most appropriate.	Multiple stakeholders recommended that we use zoledronic acid as the baseline comparator, and this is supported by clinical guidelines.
2. ICER's assessment is based on clinically unfounded efficacy assumptions. (1). Available hip fracture data (e.g., romosozumab's HR 0.54 vs. placebo at 12 months) are not used for any product due to some not having appropriate data (i.e., abaloparatide) and non-vertebral fracture data are used to model hip fractures for all products (2). Time-dependent treatment effects are not considered despite existing evidence of the rapid onset of bone-forming agents (1-2 years), particularly romozosumab (1 year), in contrast to 3-5 years of ZA and bisphosphonates in general.	We have removed romosozumab from all comparative analyses, but have always included the HR for romosozumab in the report (Page 23 and Tables E5 and E6) The existing evidence suggests that the relative reduction in fractures starts early for zoledronic acid and the anabolic agents – prior to large changes in bone mineral density. Please review the Kaplan-Meier curves for

	vertebral and particularly non-vertebral fractures in each of the clinical trials for clear demonstration of this effect. There are not clear time-dependent effects on fracture efficacy. In fact, the early fracture benefit with bisphosphonates despite minimal change in BMD has often been cited as a surprising finding.
Product-Specific Hip Fracture Estimates The ICER base case model uses nonvertebral fracture estimates in place of hip fracture estimates for all products evaluated. This could be considered appropriate in the case of abaloparatide since hip fracture estimates could not be accurately calculated given only two hip fractures were observed (both in the placebo arm) in the ACTIVE trial10. However, using nonvertebral data instead of hip fracture data for romosozumab is inappropriate as hip fractures are reported from the FRAME study: HR 0.54 (0.22 – 1.35) for romosozumab vs. placebo at 12 months and 0.50 (0.24 – 1.04) for romosozumab/denosumab vs. placebo/denosumab at 24 months. ICER only tests this flawed assumption in a sensitivity analysis resulting in almost double the estimated health benefit and a change in result for romosozumab from over $4 million dollars per QALY to less than $193,000 per QALY (draft report tables 16 and 24).	We agree that hip fracture estimates are unstable, because the teriparatide and abaloparatide trials were underpowered. That is why we did not report the NMA results in the revised report. We've removed romosozumab from the NMA and cost model, but look forward to more data about romosozumab in the future. Based on feedback on the draft report, we decided to use the hip fracture results from HORIZON as the estimate for zoledronic acid in the cost model. Since the trials of romosozumab and zoledronic acid showed a greater reduction in hip fractures than for other non-vertebral fractures, we have estimated a similar reduction in hip fractures for abaloparatide and teriparatide.
Time-Dependent Efficacy ICER assumes an immediate, full effect of ZA, which over-estimates the value of ZA. Clinical trials have reported effects at time points that are not always aligned with each other; while cross-study comparisons require considering heterogeneity in patient populations studied, the time frame of efficacy assessments across studies should be reflected in ICER's modeling. The clinical trial data ICER is considering, in combination with an understanding of the mechanism of action of each therapy, strongly suggest a faster effect attributable to bone-forming agents (1-2 years) and romosozumab in particular (1 year) in contrast with a slower, more gradual effect with bisphosphonates such as ZA, particularly for non-vertebral fracture.	The Kaplan-Meier curves in the Horizon trial do not suggest time-dependent efficacy. The HORIZON trial was powered to demonstrate a reduction in hip fractures as well as vertebral fractures at 3 years. Furthermore, the reduction in vertebral fractures was highly significant after 1 year and likely sooner ($p<0.001$). The length of the individual trials is immaterial.
It is also important to note that romosozumab is penalized in the ICER assessment for offering a 1 year treatment option, with rapid results (at 1 year), since it results in only 7 years of treatment for the sequence including romosozumab compared to 8 years for the sequences including teriparatide or abaloparatide (2 years treatment). This stems from the questionable assumption of a fixed 6 year sequenced treatment	Actually, it is zoledronic acid that is penalized as it has only 6 years of full efficacy in the model while abaloparatide and teriparatide have 8 years of full efficacy. As noted above, romosozumab has been removed from the model.

with ZA following each bone-forming agent, instead of a non-sequenced comparison or the use of the same total time frame across products (i.e. all treatment sequenced for X years).	
Recommendation: (1). ICER should use existing hip fracture data and replace with non-vertebral fracture data only for those treatments lacking robust data (e.g., abaloparatide), (2). ICER should incorporate time-dependent efficacy data into the model to capture the rapid effect of bone-forming agents, particularly romosozumab.	As noted above, we have included the hip fracture results for zoledronic acid and imputed them for abaloparatide and teriparatide.
3. ICER underestimates fracture costs and overall disease burden, including mortality. (1). Short and long-term fracture costs (the primary direct medical cost) are underestimated by ICER by using cost data from as far back as 2001 and 1989 respectively, (2). Fracture-related impact on death is inadequately captured.	We have updated the acute fracture costs to more recent estimates by Bonafede et al.
ICER's model utilizes fracture and post fracture cost inputs from as far back as 2001 and 1989 respectively, with just an adjustment for inflation that could not possibly account for the changes in care and the use of new technology that has occurred in the last 25 years. This represents a gross underestimation of the financial burden of osteoporosis even when compared to estimates from 200714 with differences of up to $10,000 dollars per fracture, or about 50% of their cost, observed.	Please see above.
An equally concerning issue identified in ICER's assessment is their reference of Tosteson et al 2007 in the claim that "excess mortality only occurred after hip fractures." Tosteson does not make that claim.15 The article focuses on mortality associated with hip fractures, and states that vertebral and nonvertebral fractures were too difficult to identify from retrospective patient charts and were thus not considered. In a literature search, we identified multiple references providing evidence that mortality increases after other fracture types such as vertebral fracture.	We revised the report to say that a review of studies reporting excess mortality following fractures showed that all but one study did not control for comorbidities. The study that did control for underlying health status found that excess mortality occurred after hip fractures (vertebral and non-vertebral fractures were not considered) at a rate roughly 50% lower than studies that adjusted for age and gender only. We therefore applied fracture-related excess mortality to hip fractures only.
Underestimating the burden of osteoporosis does a disservice to patients and physicians by undervaluing the impact of fracture-related mortality and costs, and ultimately the value of the bone-forming agents that have demonstrated their efficacy in preventing fractures. The incomplete picture painted by ICER could perpetuate under treatment of an already undertreated patient group and disease in general with often quoted treatment rates of 20% or less even in high risk elderly post-fracture patients.	We agree that there is substantial under-diagnosis and under-treatment of osteoporosis and hope that our assessment helps to highlight this important public health issue.
Recommendation: ICER should use up-to-date short and long-term cost estimates for fractures based on a systematic review of the literature. ICER should also account for the downstream disease burden of fractures in terms of their impact on	Please see above.

mortality as inputs into their model, to better capture the value of preventing such catastrophic events for patients.	
4. ICER uses unrealistic base case assumptions that do not reflect clinical practice. (1). ICER assumes 100% persistence for ZA despite their acknowledgment of real world evidence indicating that up to 60% of US patients discontinue ZA after 1 injection. (2). The assumption of a rate of decline of the effect over 10-years post-ZA appears unsubstantiated as it is based on data on residual effects on the bone and not on long-term fracture protection data over 10 years.	We explored multiple adherence scenarios including one where we "turn off" zoledronic acid (and accompanying efficacy influence) after the first year, effectively mimicking a situation in which a patient stops using ZA the first injection. However, this scenario, as well as other (lower) adherence and treatment effect decline scenarios did not produce a cost-effective result for the anabolics.
ICER assumes 100% persistence for ZA.; however, recent peer-reviewed publications on real world use of osteoporosis therapies indicate 30-60% of US patients discontinue ZA after 1 injection. ICER's report acknowledges the issue citing a 59% discontinuation of ZA by two years and 67% for teriparatide, and yet assumes 100% persistence, for six years in the case of ZA. Importantly, compromised persistence for ZA may be related to the high incidence of infusion reactions that occur with ZA.10-13 In addition, the assumption of an additional 10-years offset of effect for ZA is based on bone mineral density data of much shorter duration, which show only residual bone mineral density effects on the bone (not long-term fracture protection over 10 years). With the combined assumptions of 100% persistence and an additional 10-years offset effect for ZA, ICER's assessment inappropriately overestimates the real-world benefit of ZA.	There are insufficient data real-world data measured in a consistent fashion for the three primary drugs in the updated review to fairly assess adherence rates. This is primarily because abaloparatide has just been approved by the FDA and has no real-world data. Real world data cited in the report found that two-year adherence was similarly poor for both teriparatide and zoledronic acid. Since abaloparatide appears to have more local reactions than teriparatide and must be given by daily SC injections, adherence may be worse, but we have no data. To be fair to all drugs considered, we have modeled 100% adherence, recognizing that this is an idealized, and not real-world comparison.
Finally, ICER focuses on the time on sequenced therapies (i.e., on ZA), which confounds the estimation of value of the bone-forming agents being assessed; time on ZA accounts for 80% of the total treatment period in ICER's assessment.	The effect of this is that the full benefits of anabolics compared to zoledronic acid are extended over a long period of time, i.e. over the entire duration of subsequent ZA use. Nonetheless, the results show that even under very favorable conditions, the ICERs do not approach cost-effectiveness.
Recommendation: ICER should simulate real world estimates of persistence of each therapy over time and assume credible ranges for the decline of effect over time.	Multiple scenarios have been analyzed simulating a broad range of persistence and treatment effect decline over time.
5. ICER's model is unstable as demonstrated by the extremely large volatility of its results. In ICER's model, variation in one model input changes the results by millions of dollars per QALY. This is a sign of enormous uncertainty and lack of robustness of the model. However, ICER chose to focus the sensitivity analysis on factors with little impact on results such as utility (40% of ICER's one-way sensitivity analysis) and reaches strong and definitive conclusions that seem disconnected from the underlying uncertainty.	The sensitivity analysis included all model parameters. The figures display the 10 most influential parameters.
Variations of one single input in ICER's model cause changes on results by millions of dollars per QALY. In the case illustrated above (issue #2), when the use of non-vertebral fracture rates	All parameters were jointly varied in probabilistic sensitivity analysis (PSA). We have described the range and statistical

to model hip fractures is reversed, the results are approximately 15 times or $4M/QALY better for romosozumab. However, ICER does not make an appropriate use of probabilistic sensitivity analyses to examine the joint uncertainty in parameters thus putting too much emphasis on point estimates that are greatly uncertain. This results in overly strong conclusions disconnected from the high uncertainty around key parameters and assumptions made.	distributions used for each model parameter in the report. Please see the PSA results in Appendix F.
Correcting the above-mentioned additional issues results in romosozumab being cost-effective according to generally accepted willingness-to-pay thresholds.	Romosozumab has been removed from the cost-effectiveness analysis of the report.
<u>Amgen Modeling</u> Amgen, in collaboration with external experts, have replicated ICER's cost-effectiveness model, despite the scarcity of details provided, and also created a de-novo model based on published models. The former was used to estimate the extent of the impact of the assumptions and data input choices made by ICER in the results, which helped confirm the issues illustrated above. The latter was used to simulate relevant comparisons using clinically relevant inputs and assumptions and demonstrates that romosozumab would provide good value for patients, healthcare systems and society as a whole, and will be subject of upcoming publications.	We have added additional details on our modeling methods to the report.
The correction of the flaws in the ICER assessment is strongly recommended to ensure an appropriate valuation of bone-forming agents for osteoporosis in postmenopausal women in need of rapid bone formation. To provide full transparency, ICER should make their model more transparent and accessible.	We have addressed a number of concerns and added additional detail in the latest version of the report.
Recommendation: ICER should choose clinically sound base case assumptions and conduct a robust assessment of uncertainty around data inputs and assumptions, and utilize the results to appropriately inform conclusions of the assessment as per established good practice in economic evaluation.	Please see above.
Eli Lilly	
(1) The definitions for each fracture site across studies are not consistent in the NMA.	We agree that this adds uncertainty to the results.
a. For vertebral fractures, the approach recommended by FDA is to assess lateral spine radiographs using a combination of quantitative morphometry (QM) and semi-quantitative (SQ) assessment, and this approach was used in the zoledronic acid Horizon Trial (Black DM 2007) and in most other osteoporosis studies. The abaloparatide trial used a SQ with SQ confirmation approach (Miller 2016), which is considered similar (Harry Genant, personal communication), and the radiographs were assessed in blinded fashion so that the vertebral fracture data in the abaloparatide study should not be subject to bias. However, the method initially used in the teriparatide Fracture Prevention Trial (FPT, Neer 2001) used a single SQ reading, a less rigorous definition of fractures, and this methodology	The vertebral fracture measurements in Prevrhal were published 8 years after the primary results of the trial and were not pre-specified, although additional analyses based on alternative definitions for morphometric fractures were anticipated in the analysis plan of the trial. The method used does not match that of the other trials, but is closer than the original approach. Despite these concerns, we have elected to use the Prevrhal estimates as the primary inputs to the NMA and the cost-models for the reasons noted in the report. The results using the original results

includes putative fractures which would not be confirmed during a confirmation step, introducing "noise" and reducing biological signal. To be consistent, the teriparatide data from the FPT using the QM+SQ method (Prevrhal 2009) should be included in the NMA. Some important methodological points about the Prevrhal et al. analysis include:	presented in Neer 2001 are presented in the Appendix.
i. The original Neer publication from the Fracture Prevention Trial reported single SQ readings performed in blinded fashion by radiologists under the supervision of Dr. Genant. However, The Fracture Prevention Trial protocol included text recognizing that other definitions of vertebral fracture might be employed to assess the radiographs.	We have described both in the final report.
ii. The quantitative morphometry was performed by a trained and validated central reader blinded to group assignment using in-house (Department of Radiology and Biomedical Imaging, University of California San Francisco, San Francisco, CA, USA) software under the supervision of Dr Prevrhal.	Noted.
iii. Working with Dr. Dennis Black, a statistical analysis plan was approved prior to the completion of the QM assessments. The statistical analysis plan described the definition of fracture, defined how missing data would be handled, and specified all aspects of the statistical analysis.	See above
b. Non-vertebral fragility fracture is the standard endpoint in most osteoporosis studies, and excludes fingers, toes, face, skull, and traumatic or pathological fractures (Krege and Wan 2012). While this is the correct endpoint, the assessment of whether fractures are due to fragility must be performed in blinded fashion to avoid bias. ICER should not compare unblinded, nonvertebral fragility fracture data for teriparatide from the abaloparatide ACTIVE trial to blinded data from other studies. Although ICER did run a sensitivity analysis excluding these open label data, the base case should exclude the unblinded open-label teriparatide data from the abaloparatide study.	In the abaloparatide ACTIVE trial, the fractures were adjudicated by a central committee blinded to treatment status. Furthermore, fractures are an objective "hard" outcome, which have been shown to be much less subject to bias arising from lack of blinding in trials than subjective outcomes such as pain. Thus, we have not changed our primary analysis and have included the open-label teriparatide non-vertebral fracture results.
(2) Although the Draft Evidence Report relied on the traditional PICOTS format, the patient populations of the 3 anabolic studies used in the NMA were widely heterogeneous in terms of prior vertebral fracture (100% [FPT], 24% [ACTIVE], and 18% [FRAME]); and mean BMD T-scores at the total hip (-2.6 [FPT], -1.9 [ACTIVE], and -2.5 [FRAME]). The higher incidence of reported fractures in the control group of the FPT further indicates that the patient populations included in the teriparatide study were at 2-4 times higher risk (see Table 7, p.26), and thus not comparable to patients included in the other trials pooled for the NMA.	We agree that the trials have different inclusion and exclusion criteria. However, all the trials included post-menopausal women with osteoporosis and the ages of the women in the trials were remarkably similar. Thus, the study samples were more similar than different. Furthermore, differences in sample characteristics only impact the validity of an NMA if there is effect modification of one or more of the interventions by characteristics that differ between the study populations. Analyses for each of these agents did not identify effect modification by prior vertebral fracture or baseline risk for fracture (FRAX

	score). Thus, this concern has no bearing on the validity of our NMA.
(3) The NMA used a fixed effect model and assessed goodness of fit and heterogeneity using deviance information criterion (DIC) and residual deviance (resdev). A fixed effect model (as used by the authors) assumes that there is a single true effect of the intervention which is common across all studies. However, given the noted heterogeneity between the baseline characteristics of patient populations, the fact that each of the three included studies examined different interventions, as well as the wide range of reported treatment effects, it is highly unlikely that a fixed effects model would be appropriate. Thus, a random effects model should be considered for the NMA. Additionally, as the authors did not report out the results of their model fit parameters (DIC or resdev) it is impossible to assess whether the model and subsequent results appropriately characterize the combined and relative effects of the intervention.	As noted above, there is no evidence for effect modification for any of the agents. Thus, a fixed effects model is appropriate. We have included the random effects model results in the Appendix as well as the DIC and resdev statistic. Note: the point estimates for the random effect model results are essentially identical to those of the fixed effect models, but the credible intervals for the random effect models are too wide to be plausible. Furthermore, they support our primary conclusion: the data do not support significant differences in fracture outcomes between the 3 agents.
[2] Teriparatide's Real-World Evidence (RWE). The Draft Evidence Report does not appear to take into consideration the large body of RWE on teriparatide's safety and effectiveness. In Section 5, Other Benefits or Disadvantages, the report concludes there are no differences between drugs in terms of their impacts on "individual patients, caregivers, the delivery system, other patients, or the public that would not have been considered as part of the evidence on comparative clinical effectiveness." The lack of RWE for abaloparatide or romosozumab should not be a justification for ignoring teriparatide's RWE. Lilly strongly recommends that ICER take into account the long history of real-world safety evidence, its real-world fracture evidence, and other real-world outcomes. The RWE on fracture effectiveness supports the findings from the FPT, and extends those results in the case of hip fractures, despite the numerous challenges and limitations associated with heterogeneous patient populations, suboptimal adherence to therapy, wide variations in clinical practice, and with incomplete information on clinical risk factors.	We have added a section on the observational evidence supporting the efficacy of teriparatide and highlighted this as a strength in the evidence base for teriparatide and zoledronic acid that is lacking for abaloparatide.
(1) Non-vertebral fracture effectiveness. a. NV fracture relative risk reductions from large prospective observational studies range from 45% to 38% (45%-38% Langdahl B, 2009 [EFOS study]; 43% Silverman S 2012 [DANCE study, Mo. 18-24 vs. Mo. 0-6]	We have included these results.
(2) Hip fracture effectiveness a. Hip fracture relative risk reductions range from 56% to 45% (Silverman S 2017 [pooled observational study data from EFOS, ExFOS, DANCE, JFOS (56% reduction in hip fracture events]; Burge RT .2017 [retrospective claims database analysis on	We have included these results.

teriparatide's hip fracture efficacy (OR = 0.55; 95% CI, 0.42, 0.74) and was based on 149 hip fracture events].	
(3) Clinical vertebral fracture effectiveness. a. Clinical vertebral fracture relative risk reductions range from 73% to 40% from retrospective claims database studies (73%-61%, Yu S 2012; 40%, Burge RT 2017); and an estimate of 62% from EFOS (Langdahl B, 2009).	We have included these results.
These fracture reduction effectiveness estimates, and particularly for hip fracture where data have been lacking, could be included in sensitivity analyses in the model.	Thank you. We have included an estimate for hip fracture reduction in the primary analysis that was imputed as an incremental reduction in fractures beyond the NMA estimate for non-vertebral fractures. The estimate from this imputation (RR 0.61) is similar to that reported in the one published observational study (RR 0.55).
Other RWE results from EFOS include improvements in back pain (decrease in bed days due to back pain; and decrease in back pain; Fahrleitner-Pammer et al 2011; decrease in frequency and severity in back pain (Aloumanis 2011); decrease in limitations of activities (Aloumanis 2011); improved mobility (Aloumanis 2011); and decreased pain and discomfort (Aloumanis 2011). In a U.S. claims database study, reductions in fragility fracture risk for teriparatide patients compared to matched non-teriparatide controls was seen as early as 6 months and continued up to 24 months [Boytsov N 2015]. In addition, fracture-related hospitalizations were 30% to 45% lower among teriparatide patients with borderline statistical significance during 12 and 18 months of follow-up and became statistically significant at 24 months. Fracture-related ER visits were 67%, 69%, 62% and 59% lower over 6, 12, 18 and 24 months of follow-up, respectively, among teriparatide patients vs. a matched non-teriparatide cohort.	We have summarized some of these results in the final report.
[3] VERO head-to-head trial (teriparatide vs. risedronate) The VERtebral Fracture Treatment Comparisons in Osteoporotic Women (VERO) trial (NCT01709110) compares teriparatide to risedronate 35mg once weekly. The study was a randomized, double blind, and double dummy active comparator study, and the primary endpoint was the proportion of patients with new morphometric vertebral fractures at 24 months. The topline data from this study were disclosed at the WCO-IOF in March 2017 (Kendler D et al.). After 2 years, fewer patients had new vertebral fractures in the teriparatide group compared to risedronate (5.4% vs 12.0%, $p<0.0001$; RR = 0.44 [0.29; 0.68]), and after 1 year (3.1% vs 6.0%, $p<0.05$). The relative risk for teriparatide vs. risedronate for other fracture endpoints included: moderate/severe vertebral fractures 0.42 (0.27; 0.65); multiple vertebral fractures 0.16 (0.04; 0.74); and clinical	We have summarized the VERO trial results in an unpublished trials section of the final report.

fractures (vertebral + non-vertebral) 0.48 (0.32; 0.74). Lilly recommends that ICER include this important H2H study in its assessment.	
[4] Analytic base case. 1. The cost-effectiveness model compares the three anabolic therapies to IV zolendronic acid bisphosphonate (BP) in the base case. A more realistic base case would consider actual real-world place in therapy for teriparatide (and newer injectable therapies), whereby substantial access barriers exist in the form of Prior Authorizations that often require lower BMD, previous fractures, and prior BP use. Following teriparatide usage, treatment with an antiresorptive therapy is recommended to help maintain the gains in bone mass from teriparatide and low rate of fracture. Available data show that antiresorptive agents increase BMD after teriparatide cessation (see for example, Prince R 2005, Leder BZ 2015). Importantly, the fracture rate after stopping teriparatide treatment remains low (Prince R 2005, Silverman S 2013, Fahrleitner-Pammer A 2011). Therefore, the base case should compare anabolics followed by an antiresorptive therapy as a sequence to each other and no treatment, while in a secondary analysis comparisons to BPs could be as conducted.	Multiple stakeholders recommended that we use zoledronic acid as the baseline comparator, and this is supported by clinical guidelines.
2. In the cost-effectiveness analysis (CEA), the authors "assumed the facture risk was similar to that observed in the clinical trials of the anabolic agents" and used a single baseline risk across the entire CEA. This is inappropriate for the reasons outlined above, that the baseline characteristics and risks for fracture were significantly different for patients in the Fracture Prevention Trial compared to the other 3 included studies. By pooling the annual fracture probabilities from the pooled placebo arms across the three studies, the authors may have biased the results to favor the effects of trials that included lower risk patients. The higher risk of fracture in the placebo arm of the Fracture Prevention Trial may increase the reported effect for trials that included lower risk patients.	We examined multiple risk groups using multiple scenarios and made favorable assumptions for anabolic therapies versus zoledronic acid. However, the anabolic regimens remained outside of commonly-cited cost-effectiveness ratios in most cases.
[5] Health utilities for non-clinical vertebral fractures. The Draft Evidence Report applies health utility decrements for clinical vertebral fractures (comprising 35% of all vertebral fractures), and no utility decreases for non-clinical vertebral fractures. However, non-clinical vertebral fractures have been associated with utility decreases, though at about one-third the impact from clinical vertebral fractures (Hiligsmann 2008; Kanis JA 2004; Cockerill W 2004), and should be used in the model to calculate QALYs.	We have added a scenario analysis that explores the addition of morphometric vertebral fracture disutility. This scenario showed little difference compared to the base case results, as most of the differences in QALYs were canceled out among the comparators.
ICER Table 4. The morphometric vertebral fracture data should use the data reported by Prevrhal et al. 2009	We have incorporated the Prevhral study results
Table E4. Under the "measurements equal and valid column", the Neer study used a different assessment. Instead, the	See above

Prevrhal 2009 study should be used in this table and in the NMA.	
Table E5. Nonvertebral fragility fractures were 6% in placebo, and 3% in the teriparatide 20 mcg/day group.	Thank you. Corrected to 5.5% and 2.6%.
Table E6. The vertebral fracture data for teriparatide should use the data reported in Prevrhal 2009 to be consistent with the SQ with QM confirmation from the zoledronic acid study, and the SQ with SQ confirmation from the abaloparatide study.	We have added the results from Prevrhal 2009.
Table E8. It is not believable that abaloparatide has a 95% reduction in hip fracture, when the data are based on 2 fractures in placebo vs. 0 on abaloparatide. There should not be a ranking of the drugs.	We agree that these results are not believable, but included them for completeness as our initial intent was to look at hip fractures as well.
Table E9. The data for teriparatide are from single SQ readings (reported in Neer 2001). A better endpoint is QM plus SQ confirmation, which is reported in Prevrhal et al. 2009.	The Prevrhal 2009 data are used.
In Table F1. Detailed Results Per Regimen, the results from the probabilistic sensitivity analysis (PSA) are given. It would be helpful to readers to supplement this table with scatterplots.	Cost-effectiveness acceptability curves are the more appropriate representation of PSA, however the results remained outside of commonly-cited cost-effectiveness thresholds in the vast majority of simulations, and we opted to omit this from the report.
Description of abaloparatide. On page 6, abaloparatide is described as "Abaloparatide is a new PTH analog, approved by the FDA on 4/28/17, and is similar to teriparatide." The precise description, as contained in the TYMLOS label, should be used in order to correctly provide these important differences between molecules: "TYMLOS injection for subcutaneous administration contains abaloparatide, a synthetic 34 amino acid peptide. Abaloparatide is an analog of human parathyroid hormone related peptide, PTHrP(1-34). It has 41% homology to hPTH(1-34) (human parathyroid hormone 1-34) and 76% homology to hPTHrP(1-34) (human parathyroid hormone-related peptide 1-34)."	We have clarified that abaloparatide is an analog of PTHrP and not PTH.
Harms. The TYMLOS (abaloparatide) label is now available: During the first month of the trial, injection site reactions were assessed daily one-hour after injection. TYMLOS had a higher incidence than placebo of injection site redness (58% vs. 28%), edema (10% vs. 3%) and pain (9% vs. 7%). Severe redness, severe edema, and severe pain were reported in 2.9%, 0.4%, and 0.4% of the TYMLOS-treated patients.	Thank you.
Of the patients receiving TYMLOS for 18 months, 49% (300/610) developed anti-abaloparatide antibodies; of these, 68% (201/297) developed neutralizing antibodies to abaloparatide. Of the patients with anti-abaloparatide antibodies tested for cross-reactivity, 2.3% (7/298) developed cross-reactivity to PTHrP, 43% (3/7) developed neutralizing antibodies to PTHrP, and 0% (0/298) developed cross-reactive antibodies to PTH. Antibody formation did not appear to have any clinically significant impact on safety or efficacy endpoints,	We have added a sentence about this in the harms section.

including bone mineral density (BMD) response, fracture reduction, immune-related hypersensitivity or allergic reactions, or other adverse events. Most of the patients with anti- abaloparatide antibodies during treatment with TYMLOS, 85% (256/300), had follow-up antibody measurements six months after completion of TYMLOS therapy. Among these patients, 56% (143/256) remained antibody positive."	
Also, abaloparatide was reported to cause tachycardia in the ACTIVE clinical trial; increasing heart rate by 15 beats/minute (TYMLOS package insert).	Thank you.
Radius	
Radius Health believes that any meaningful value framework must recognize the distinctions between drugs and drug classes, or risk being deeply flawed and worse will raise access barriers to the very women it is intending to help. ICER's draft evidence report is misaligned with real-world clinical practice and osteoporosis treatment guidelines. These guidelines differentiate between anabolic agents for their bone building mechanisms and their efficacy benefit of the reduction of both vertebral and non-vertebral fractures in patients at high risk for fracture and the maintenance role of bisphosphonate agents (Camacho et al., 2016). TYMLOS™ (Abaloparatide-SC injection) is the first new anabolic agent available to postmenopausal women with osteoporosis in nearly 15 years. TYMLOS is a human parathyroid hormone related peptide [PTHrP (1-34)] analog indicated for the treatment of postmenopausal women with osteoporosis at high risk for fracture defined as a history of osteoporotic fracture, multiple risk factors for fracture, or patients who have failed or are intolerant to other available osteoporosis therapy. In postmenopausal women with osteoporosis, TYMLOS reduces the risk of vertebral fractures and nonvertebral fractures (TYMLOS Prescribing Information, April 2017).	As noted above, comparative effectiveness analyses often compare drugs across classes. For example, antihypertensive drugs and diabetes drugs are often directly compared despite having differing mechanisms of action when they are used for the same indication. The anabolic agents and zoledronic acid are all parenteral agents indicated for severe post-menopausal osteoporosis and patients intolerant of oral therapy – hence it is appropriate to compare them.
Current Crisis in Osteoporosis Management: Osteoporosis remains significantly under treated Many osteoporosis patients at risk of fractures remain untreated. Recent evaluation of Medicare data suggests that the plateauing of age-adjusted temporal reduction in hip fractures may be associated with the decline in testing and treatment of osteoporosis (Khosla and Shane, 2016). Unfortunately, patients discharged with a hip fracture from hospitals today remain under-treated compared to those discharged with other major events (i.e., myocardial infarction). The undertreatment may also be due in part to negative media reports associated with the risk of rare but serious adverse events with bisphosphonates including osteonecrosis of the jaw (ONJ) and atypical femur fractures (AFF) (Yood et al., 2008). In 2010, the FDA issued a global warning regarding these safety risks early in treatment, which may have contributed to a 50%	We agree with the observation that many patients with osteoporosis have not been identified nor treated. We have highlighted some of the studies documenting this issue in our review including concerns about AFF and ONJ, though as you are aware, these are rare events and the benefits of treatment outweigh the risks. Furthermore, no matter what treatment is used initially, anti-resorptive therapy is required (either as the primary therapy or to preserve the benefits of anabolic therapy). The concerns about ONJ and AFF apply equally to all treatment approaches in our model and thus cancel themselves out.

decline in use of these agents (FDA 2010 warning-FDA website). To overlook these causes of morbidity and resultant lack of adherence in today's treatment paradigm as part of any cost effectiveness model would be an omission affecting the validity of the model's outcomes.	
A recent evaluation of medical and pharmacy claims data from a large, geographically diverse cohort of private commercial and Medicare Advantage plans with no prior history or treatment of osteoporosis who experienced a new hip fracture (n=8,349) further documents the gap between evidence-based guidelines and reality. Of women who experienced a hip fracture, only 17.1% and 23.1% had evidence of osteoporosis assessment and/or treatment within 6 or 12 months of their fracture respectively (Gillespie and Morin, 2017). Hip fractures are considered "non-vertebral" and as discussed earlier, nonvertebral fractures in total represent the clear majority of all fractures (Burge et al., 2007).	We agree. Thank you.
Early therapy of patients at high-risk of fragility fractures is key in reducing osteoporosis morbidity, mortality and associated costs The importance of early intervention has been consistently supported in several studies. The 12-month period after the first osteoporotic fracture has been noted as the critical year, a key high-risk period requiring interventions to improve patient outcomes. Prior fracture history is the highest predictor of future fracture risk (Weaver et al., 2016). The rate of repeat fracture within 1 year of the initial fracture based on real-world data varies between 4%-9% and is dependent on the fracture site (Song et al., 2011).	We recognize that a fracture increases the risk of a subsequent fracture and account for this in the model in a scenario analysis.
Although there is a high prevalence of vertebral fractures (27%) nonvertebral fractures represent 73% of all fractures and 94% of related costs (Burge et al., 2007). They include wrist (19%), hip (14%), pelvic (7%) and other fractures (humerus, clavicle, and hand/fingers-33%). Looker and colleagues recently provided the first nationally representative estimates of FRAX-based 10-year probability of major osteoporotic fracture (hip, spine, proximal humerus, or distal forearm) for adults aged 50 and over using the 2013-14 NHANES survey. The 10-year probability of major osteoporotic fractures varied from 2.9% for 50-59 age group to 27% for 80+ age group (Looker et al., 2017).	Thank you. Our model includes both an increasing risk for fractures (vertebral and non-vertebral) with aging and with prior fracture.
Hip fracture has a significant downstream impact including associated health and economic consequences, which need consideration. Of women over age 50 who sustain a hip fracture approximately 25% of women die in the year following the fracture, 50% never walk independently again and 20% require permanent nursing home placement (U.S. Department of Health and Human Services, 2004). According to a recent Bone Health Index Survey by the National Osteoporosis	We agree that hip fractures have substantial morbidity, but they are also a marker for fragility / poor health. When co-morbidities are accounted for, the excess mortality, for instance, is lower than is commonly reported from naïve analyses of the data.

Foundation (NOF), loss of independence (42%) and lost mobility (25%) ranked as the leading concerns about aging for osteoporosis patients as well as their caregivers' uncertainty about their ability to manage their patient care (50%) (National Osteoporosis Foundation 2016). The burden and cost of disease associated with distal radial fracture has also been significantly underestimated. Patients with distal radial fracture have a much greater risk for subsequent hip fracture within 1 year (HR=3.45) (Litwic et al., 2014). The risk is the greatest in the first month after the distal radial fracture (Chen et al., 2013). However, many postmenopausal women, especially those in their 50s, as well as their treating physicians, fail to recognize that the fragility fractures (e.g., wrist) could be a sentinel event or warning sign for osteoporotic disease progression. For these patients that need immediate fracture prevention alternative treatment options that are specifically designed to build bone are suggested. The use of commonly prescribed bisphosphonates, that only slow bone loss and do not improve or build bone, are simply not enough.	We have highlighted patients concerns about loss of independence and mobility in several sections of the report.
The total cost of care is also significantly higher for those experiencing a subsequent fracture compared to those without a history of prior fracture for both Medicare ($34,327 vs. $20,790; p<0.001) and for the commercial health plan enrollees ($39,501 vs. $19,131; p<0.001) (Weaver et al., 2016). In a recent study of US managed care enrollees, the -subsequent fracture was estimated to increase medical costs by $47,351, $43,238, and $23,852 for commercial patients with prior hip, clinical vertebral, and non-hip/-nonvertebral (NHNV) fractures and $18,645, $19,702, and $19,697 for Medicare patients respectively. The AACE/ACE guidelines acknowledge the importance of the inclusion of an anabolic therapy for treatment of patients at high risk of fracture, including those with a prior fracture history (Camacho et al., 2016). It is also recognized that the use of anabolic therapies to build bone as early as one month, and not just enhance existing bone mineral density, will have a positive impact on reducing the humanistic and economic burden of subsequent fractures.	We have updated the model with more recent fracture costs that are more in line with this. However, our model indicates that fracture costs play a small role in the cost-effectiveness equation compared to the difference in QALYs that is conferred by the relative risk estimates.
Limited access to effective therapies for patients at high-risk of fractures will prolong the poor health and economic outcomes Recent position papers from medical societies include calls to action to: (1) emphasize the importance of early diagnosis and early treatment, (2) highlight the value of shared decision making and customizing treatment in consideration of benefits and risks of individual therapies and patients; and (3) suggest additional approaches to identification and treatment of high risk patients where current healthcare pathways may not be sufficient.	Thank you. As noted above, we agree that under-screening and under-treatment of osteoporosis is an important public health concern.

The ASBMR working group suggests a goal-directed treatment for osteoporosis where treatment decision is guided to maximize patient's ability to achieve goal. Osteoporosis treatment goals need to parallel indications for initiating treatment and logical treatment goals are BMD levels above and fracture risk levels below those for which treatment is usually recommended. This Working Group interim report supports the potential value of goal-directed treatment and sets out several principles to guide this approach to selecting and monitoring treatments. Some of these principles such as considering a more potent initial treatment in those with high risk of fracture and continuation or intensification of treatment when a vertebral fracture occurs on therapy could be put into practice now (Cummings et al., 2017).	We have read the position paper by Dr. Cummings et al, but there is a lack of clinical trial data supporting the approach. As you state, it has "potential value." We look forward to learning more about the actual value from future studies.
The American Association of Clinical Endocrinologists and American College of Endocrinology (AACE/ACE) recognize the importance of anabolic or bone-building agents. The AACE published new guidelines on the treatment of osteoporosis last September (Camacho et al., 2016). It is now recommended that an anabolic agent be used as a first-line treatment for patients at high risk for osteoporotic fracture. And in fact, there is new evidence that sequence of therapies matters, suggesting the use of a bone-building agent (anabolic) followed by an antiresorptive agent, such as a bisphosphonate or denosumab wherever possible to improve bone density and decrease fracture risk in these patients (Cosman et al., 2017).	We read the AACE/ACE guidelines as we drafted our scope for this review and have had regular input from the organization throughout the process. There is evidence from studies of bone mineral density supporting anabolic agents, but this remains controversial because of the lack of fracture data supporting this hypothesis (see the letters in response to *Cosman 2017*, for instance *Grey et al 2017 PMID 28294409*).
Finally, the National Committee for Quality Assurance (NCQA) recognized that health care providers often neglect treating patients with osteoporosis including high risk patients and has called for the need to focus on secondary fracture prevention and closing the care gap for testing and treatment for high risk patients (National Committee for Quality Assurance, 2016).	As noted above, we agree that this is an important public health issue, though it is more complex than simple neglect on the part of health care providers.
We reiterate our position that ICER should focus on a patient-centered approach that clearly delineates the distinction of patients at high risk for fractures in need of immediate fracture prevention, as well as the need for quality care for that specific patient population that takes into consideration total cost of care, and not limit the analysis to only direct product unit costs without current and comprehensive direct fracture costs and indirect treatment and intolerance costs.	ICER is fully committed to patient centered care. We hope that the public discussion at the meeting will point to studies defining patient characteristics that identify individuals for whom anabolic therapy represents a good value.
Of interest is the patient consultation and feedback to ICER in the May 3rd report noting that insurance often requires that they fail an oral therapy before authorizing an injectable therapy. Bisphosphonates are often recommended as first line use; however, these agents slow the loss of existing bone but do not build new bone. Osteoporosis is not one disease, and no one treatment will work for everyone. Those who make new bone too slowly need another option, particularly during the first critical year post the initial fracture. We suggest ICER's	First, the ICER report is not a practice guidelines. We are assessing the comparative clinical and economic value of new therapies. That said, it is important to note that we have modeled the anabolic agents as first line therapy, rather than subsequent to bisphosphonate therapy. This approach biases our analysis in favor of anabolic agents. Furthermore, the outcome that matters to

recommendation consider evidence-based guidelines and to reduce the administrative burden on clinicians and patients supporting early access to targeted therapy for high-risk patients.	patients is fracture prevention, not change in bone mineral density. Changes in bone mineral density are informative, but not definitive.
Appropriately compare like agents as they are not designed to do the same things. Therefore: -Agents that build bone and demonstrate early fracture reduction, within 2 years, should be compared with others that do the same. -Antiresorptives that slow resorption of existing bone, such as zoledronic acid and denosumab, should be compared with other antiresorptive. -The guidelines do make a distinction between drug classes in consideration of "patients' risk of fracture, prior disease, and treatment history" and so ICER should equally take these differences in drug classes into consideration in their model.	Initially, we considered comparing the anabolic drugs to alendronate, but received feedback from multiple clinical experts (endocrinologists, rheumatologists), multiple pharmaceutical companies, patient organizations, and specialty societies that zoledronic acid would be the most appropriate competitor. It is common practice to perform comparative effectiveness reviews of drugs from different classes that share a common clinical indication. That is what we have done in this review.
Use WAC instead of net price: -Use correct price of TYMLOS vs. the other approved agents at a WAC basis. -Consistent with the BIO response to ICER on the "National Call for Proposed Improvements to its Value Assessment Framework" we previously suggested using the WAC which can be easily verified rather than the variable, estimated and unsubstantiated net prices of the prescription drugs in the value assessment methodology. -Moreover, use of a net price fails to take into consideration the impact on patient cost-sharing obligations between the agents and the corresponding discontinuation of treatment due to affordability issues. Since manufacturer discounts are not directly passed on to patients, a reduced WAC is the only direct way a manufacturer can lower out of pocket cost for Medicare D patients fostering greater adherence and associated outcomes.	The overwhelming majority of comments we have received on pricing considerations have focused on our previous use of WAC pricing to determine cost-effectiveness estimates, as this ignores the reality of discounting and rebating. In contrast, our switch to estimated net prices (based on SSR Health's use of publicly disclosed net sales data from manufacturers) has generated an overwhelmingly positive response. Ultimately, this is an issue of semantics, as our value-based price benchmarks are generated based on cost-effectiveness thresholds (not WAC or net prices), and the discounts from WAC required to achieve these thresholds are clearly presented in the report. Manufacturers have multiple options available to approach these benchmarks, including increasing discounts or rebates, entering value-based contracts, or even reducing WAC pricing.
Use the studied treatment duration for TYMLOS: -TYMLOS was approved based on 18 months of treatment, not 24 months of treatment, substantiated by the ACTIVE and ACTIVExtend trials.	We have received expert input that, like teriparatide, abaloparatide is likely to be prescribed for 24 months of treatment, consistent with its label.
Incorporate impact of the first demonstrated sequential therapy approach for TYMLOS: -TYMLOS approval includes data from two trials, 18 months of using TYMLOS (ACTIVE) (Miller et al., 2016) to demonstrate relative risk reduction of vertebral and non-vertebral fractures followed by the first six months of the use of a bisphosphonate	The ACTIVExtend trial results formed part of the evidence base for our decision to maintain the benefits of anabolic therapy with ongoing bisphosphonate therapy. The HRs for vertebral and non-vertebral fractures were nearly identical for the ACTIVE trial and the ACTIVExtend trial as shown in Figures 2 and 3

(alendronate) to "build and extend" gains in BMD (ACTIVExtend) (Cosman et al., 2017). -ACTIVExtend is an important sequential treatment data set to inform physicians and patients how to treat postmenopausal women with osteoporosis and a high risk of fracture. The ICER model should take into consideration this demonstrated treatment paradigm with both its efficacy and safety results.	of the paper. For example, there was an 86% reduction in vertebral fractures during ACTIVE and an 87% reduction in the ACTIVE+ACTIVExtend analysis. There are concerns about selection bias given that 1645 patients were randomized to the two arms evaluated in ACTIVExtend but only 1139 provided data for ACTIVExtend (31% lost to follow-up).
When available, utilize data from comparative trials to accurately compare like agents (e.g. anabolics) rather than using cross-study comparisons which have inherent limitations due to study design, inclusion / exclusion criteria, etc.	As noted above the trial populations and study designs are quite similar and there is no evidence of effect modification by prior fracture history or risk for fracture for example for abaloparatide as described in Dr. Cosman's 2017 paper (Cosman F, Hattersley G, Hu MY, Williams GC, Fitzpatrick LA, Black DM. Effects of Abaloparatide-SC on Fractures and Bone Mineral Density in Subgroups of Postmenopausal Women With Osteoporosis and Varying Baseline Risk Factors. *Journal of bone and mineral research: the official journal of the American Society for Bone and Mineral Research.* 2017;32(1):17-23.)
Use both event and incremental costs of subsequent fractures as well as the burden to the patients based on peer-reviewed published data.	Please see above re: fracture costs.
Weight each fracture site appropriately, based on prevalence and associated cost so as to not overestimate the impact of hip and underestimate the impact of other fractures.	To weight each fracture, we summed 1) the number of each fracture type (hip, vertebral, and non-vertebral) from the trials, as well as 2) the fracture types' associated follow-up time in person-years, then calculated annualized rates of each fracture type. To estimate the relative risk of another fracture we applied the ratio of hip to nonvertebral fracture relative risks reported in the HORIZON trial (see Table 11 in report for additional hip fracture relative risk explanation). The latter was done to not overestimate the risk of hip fractures or underestimate others. The costs of a hip or other fracture is then applied only to the proportion of patients experiencing this event.
Utilize real-world evidence (third party data) to estimate adherence rates as the ICER assumption of 100% is inconsistent with real-world evidence (Yang et al., 2016; Earnshaw et al., 2016; Modi et al., 2016).	We explored multiple adherence scenarios including one where we "turn off" zoledronic acid (and accompanying efficacy influence) after the first year, effectively mimicking a situation in which a patient stops using ZA the first injection. However, this scenario, as well

	as other (lower) adherence and treatment effect decline scenarios did not produce a cost-effective result for the anabolics.
Model serious adverse events (Table 6 in ICER report) to accurately reflect the safety of each of the agents. It is important to not only consider the efficacy of each agent but also their safety profiles. For example: TYMLOS and teriparatide each have boxed warnings for osteosarcoma, a rare but serious adverse event found in preclinical studies in rats, resulting in a cumulative use of no more than two years in a patients' lifetime. There have been no incidences of osteosarcoma in the human trials conducted for TYMLOS or for teriparatide (Andrews et al., 2012). Antiresorptives, such as bisphosphonates and denosumab, have rare but serious adverse events of AFF and ONJ. In addition to AFF and ONJ risk, denosumab also carries in its label a warning for multiple vertebral fractures (MVF) following the discontinuation of denosumab treatment, with new vertebral fractures occurring as early as 7 months (on average 19 months) after the last dose. The costs associated with this known risk should be included in the ICER analysis. Any agents with a REMS would also have published data that informs the real-world incidence of any safety events as well. Any published clinical or real-world data that demonstrates the impact of discontinuation on the sustainability of fracture risk reduction should also be taken into consideration in the model (Yang et al., 2016, Earnshaw et al., 2016; Modi et al., 2016). Underestimating safety could risk ignoring the burden this places on patients and any potential hospitalization or resource utilization costs. This must be taken seriously.	Anabolic regimens as well as zoledronic acid exhibited similar serious adverse event rates compared to placebo and each other in their respective trials. These small event rate differences are unlikely to impact cost-effectiveness results.
Agents that have not been approved should not be considered in the model.	Romosozumab is no longer considered in the model.
Accurately estimate treatment uptake: -ICER must take into consideration current treatment guidelines, levels of payer access, and access restrictions. -As suggested to ICER through BIO previously "In addition to using historical data, estimation of new treatment uptake can also consider evidence-based treatment guidelines especially where there is a treatment paradigm change as well as any other quality of care measures that may impact prescribing habits."	ICER no longer attempts to estimate the uptake of a new intervention as part of its potential budget impact analysis. Rather, ICER presents information that can allow stakeholders to ascertain the potential budget impact of a new treatment according to a wide range of assumptions on price and uptake.
Please also refer to the ASMBR "Call to Action" which cites that new evidence is emerging that the 30-year downward trend in hip fractures in the U.S. has hit a plateau in the last few years, indicating that the field as a whole must take action to aggressively reduce fracture risk in the US aging population. Many experts are now acknowledging that there is a crisis caused by the declining rate of testing, diagnosis and treatment of high- risk patients. Allowing these patients to go untested	As noted above, we acknowledge the importance of identification and treatment of osteoporosis in the US.

and untreated frequently leads to debilitating fractures that cause disability, loss of independence and even death. In fact, 25% of women over the age of 50 who sustain a hip fracture die in the year following the fracture, 50% never walk independently again and 20% require permanent nursing home placement (U.S. Department of Health and Human Services, 2004). According to a recent Bone Health Index Survey by the NOF, loss of independence (42%) and lost mobility (25%) ranked as the leading concerns about aging for osteoporosis patients as well as their caregivers' uncertainty about their ability to manage their patient care (50%) (National Osteoporosis Foundation, 2016).	
As one ages, the bone building (or formation) part of the process is often unable to keep up with the bone loss (or resorption) part of the process. In women, estrogen plays a role in regulating the bone formation and resorption process. Women start losing estrogen at menopause, which is accelerated over the initial period during their postmenopausal phase, and contributes to women beginning to lose more bone than they are replacing or building. Left untreated, osteoporosis can lead to bone deterioration throughout the body, leaving patients vulnerable to osteoporotic fracture. Today, the hospitalizations for osteoporotic fractures are higher than that of stroke or heart attack or breast cancer (Singer et al., 2015). Additionally, osteoporotic fractures account for more hospitalizations and associated costs than cardiovascular disease or breast cancer. Nonvertebral fractures represent the clear majority of osteoporotic fractures as well as the associated costs (Burge et al., 2017). Once a patient has experienced a fracture, the risk of another fracture is highest in the first year, and the patient is 3 times more likely to have another fracture. The risk remains high for the subsequent years (Harvey et al., 2016). Fractures due to osteoporosis are estimated to cost $25 billion per year by 2025. It is counter to clinical evidence to recommend limiting treatment options to only antiresorptives and not acknowledge the clear clinical data supporting the use of anabolic therapy for postmenopausal women with osteoporosis to reduce their high risk for future fracture.	As you know, bone formation and resorption are tightly linked and as we age the balance tips towards greater resorption than formation, leading to gradual loss of bone. Either slowing resorption or increasing formation tips the balance in the other direction. The ultimate goal of therapy is to prevent fractures. Both anti-resorptive and anabolic therapies prevent fractures. Published data to date do not clearly demonstrate that one approach is more effective than the other in any defined subgroup of patients, though BMD data do suggest that starting with an anabolic agent may be more effective. As noted above, we have incorporated the increased risk for fracture following an index fracture in our model. We agree that osteoporotic fractures are common and have substantial impacts on patient quality of life, hospital utilization, and costs in the US and have incorporated all those elements into our model.
Aimed Alliance	
To prevent bone deterioration and fractures, and improve overall quality of life, individuals with osteoporosis must have access to effective treatment options. However, the Draft Report may limit those options.	

QALYs are Discriminatory The use of quality-adjusted life-years ("QALYs") to develop a rigid price cap is inconsistent with American values and public policy. Congress added language to the Patient Protection and Affordable Care Act that prohibited the Patient-Centered Outcomes Research Institute ("PCORI") from using QALYs as a threshold for determining coverage, reimbursement, or incentives in the Medicare program. The ban reflected a long-standing concern that the approach would lead to health care rationing as well as age- and health status-based discrimination, unfairly favoring healthier and younger populations. This is especially problematic when applied to a condition, such as osteoporosis, which disproportionately affects those who are 50 years of age and older. QALYs put a price tag on the value of a human life that merely reflects an individual's diagnosis. They treat individuals' lives and health as a commodity and ignore the patients' and practitioners' individualized concept of the value of treatment. Therefore, the QALY should not be used to set a threshold for a large population of individuals with one-of-a-kind life narratives across a complicated health care system. Instead, Aimed Alliance urges ICER to consider other methods of valuation, including life years gained, as ICER did in its rheumatoid arthritis report, to measure the benefits of osteoporosis medications.	We feel these statements misunderstand the purpose and practice of using QALYs in cost-effectiveness analyses.
While patient advocacy groups have noted that step therapy and prior authorization are major barriers to accessing anabolic treatment, suggestions that anabolic therapies are not cost-effective would serve to bolster third-party payers' use of such policies in direct conflict with patients' best interest. As a result, patient access to live-saving treatments could decrease. Instead, recommendations must be made that prioritize access to such treatments for individuals for whom they are clinically indicated.	We appreciate this input.
Patient and Practitioner Perspective Patients must have a meaningful role in the discussion of value given that they are directly impacted by a report that seeks to define the effectiveness and value of their treatment options. Therefore, accounting for how patients define the value of their treatment options should be critical to ICER's analysis.	We agree on the critical role of patient input.
While the Draft Report notes that loss of independence and loss of mobility are the top two concerns among patients, it is unclear how these two factors were calculated into the cost-benefit analysis. Moreover, as ICER notes, available studies and clinical trials do not report outcomes most meaningful to patients, including living independently, the ability to perform the activities of daily living, social engagement, quality of life, reduced fear and anxiety about the disease and treatment, and	Loss of independence, mobility and the other factors mentioned are captured in the utility estimates, which are used to calculate the QALY, and hence they are an integral part of the model outcomes. These utility estimates are based on additional studies that specifically report on the health-related

safety from adverse drug effects. Therefore, adequate studies must be conducted on these important factors before an accurate cost-benefit analysis can be conducted.	quality of life impact/utility associated with the disease, its treatment and with fractures.
Additionally, the current committee lacks health care practitioners and patients with bone health background. It is unclear whether a bone health physician was consulted in drafting the report. And while patient groups may have been consulted, no one with experience in the bone health field will be voting on the final outcome of this report.	Bone health experts contributed to the report. The CTAF members have broad expertise in interpreting evidence and reviewing policy.
Improper Comparison The Draft Report compares anabolic therapies to zoledronic acid. Yet, anabolic therapies are designed to build new bones whereas zoledronic acid slows down degeneration. Also, as the Draft Report acknowledges, anabolic therapies are taken for a period of one to two years, whereas zoledronic acid is recommended for longer periods, with a treatment holiday after three years for individuals with low to moderate risk osteoporosis and after six years for individuals at higher risk.	As noted above, clinical experts in osteoporosis, specialty societies, patient groups and pharmaceutical companies all provided feedback recommending zoledronic acid as the appropriate comparator.
These medications work differently, have different results, and are taken for different periods of time. Moreover, as ICER acknowledges, there was only one head-to-head study of these drugs, and therefore, insufficient data to make such a comparison. As such, anabolic therapies should not be compared to zoledronic acid.	They do work differently, but the HORIZON trial demonstrates almost immediate efficacy of the drug in preventing fractures, despite minimal effects on bone density in the first few months of therapy. Similar observations apply to the anabolic drugs.
Significance of Hip Fractures The Draft Report downplays the frequency of hip fractures. Every year, over 300,000 individuals 65 years of age and older are hospitalized for hip fractures. This is not an insignificant number given the severity of such fractures. Hip fractures result in chronic pain, reduced mobility, disability, loss of independence, and death. Within one year of a hip fracture, mortality rates are between 20 and 24 percent, 40 percent of individuals are unable to walk independently, and 60 percent require. As a result of these losses, 33 percent are completely dependent or in a nursing home in the year following a hip fracture. Yet, the Draft Report does not seem to take into account costs associated with assisted living. Therefore, the final report should adequately assess the impact of hip fractures, including their indirect costs.	We recognize the key nature of hip fracture outcomes and include the costs, disability, and impacts on quality of life of hip fractures in our cost model. It is unfortunate that none of the clinical trials of anabolic therapy were powered to examine hip fractures as an outcome. We have biased our analyses in favor of anabolic agents by assuming that they will have a reduction in hip fractures that is greater than their reduction in non-vertebral fractures given the complete lack of evidence of benefit in preventing hip fractures. The lack or randomized trial evidence on hip fractures is why the 2017 ACP guidelines do not recommend anabolic agents as first line therapy for patients with osteoporotic fractures or osteoporosis. If we took a strict methodologic approach, we would assume no benefit for the anabolic agents in preventing hip fractures and the drugs would have significantly greater costs per QALY.
American Bone Health	
As a community-based, consumer organization, American Bone Health works with consumers to improve awareness of osteoporosis and educate them on what to do to prevent bone	We agree as noted above and have highlighted the under-screening and under-treatment of osteoporosis in the report.

loss and fractures. During our last national awareness screening event in July 2017, we found that 55% of the participants of Medicare age had not had a bone density test (a covered benefit under Medicare) and only 24% of individuals at high risk for fractures were on a treatment for osteoporosis. This gap in diagnosis and treatment leads us to serious concerns about the unintended consequences that may result from the ICER report.	
First, for patients at high risk of fracture, the benefits of osteoporosis treatments, and the favorable benefit/risk ratio, are clearly demonstrated in clinical trial data with large groups of patients. New therapeutic options allow greater flexibility for patients; however, determining the best treatment option is an individual decision best left in the hands of doctors. Patients with certain clinical profiles, including eg, low-turnover osteoporosis, steroid-induced osteoporosis, or adult-onset hypophosphatasia (HPP) may/will benefit from an anabolic agent (and patients with HPP and osteoporosis should never be treated with a bisphosphonate), even if the ICER report does not deem it to be a cost effective option based on the comparative data.	We fully agree that patients with osteoporosis should be treated. Our report is not a clinical guideline making treatment recommendations for individual patients. We are providing comparative effectiveness and cost information to help patients, clinicians, payers, insurers, and guideline authors make informed decisions about the choices that are made. Our analysis is not looking at the special subgroups of patients with HPP, steroid-induced osteoporosis, or low turnover osteoporosis.
Second, placing an economic "score card" on the available treatment options will likely be seen by insurance companies as guidelines for limiting formularies. This will effectively reduce the ability of physicians to prescribe the most appropriate treatment for their individual patients and continue the practice of allocating the best options only to those individuals who have the resources to pay for them.	As with clinical effectiveness, cost of medical therapies is important to all stakeholders. The goal of ICER's work is sustainable access to high-quality care for all patients.
As an example of the continued inequity of care in osteoporosis management, in the last two months, our facility has seen three patients from the local clinic population with displaced hip fractures who were unaware of their fracture. Thankfully, these three patients had access to a bone density test through their county insurance plan. Still, it is quite disturbing that it took a preventive screening to discover a serious, potentially deadly fracture. These women should not only have access to screening, but access to the best treatment options to prevent further fractures. [Photos included in full submission]	We agree that access to care is a significant issue in our country and commend your efforts to support access to care to all people in our country. Under-diagnosis and under-treatment of osteoporosis is an important public health issue.
Finally, the ICER analysis assumed that "a new drug or device that would take market share from one or more drugs, and calculate the blended budget impact associated with displacing use of existing therapies with the new intervention." This way of thinking continues to undermine the crisis that we have with the under diagnosis and under treatment of patients at high risk for fracture.	While underdiagnosis and undertreatment are concerns in this and many other clinical areas, our budget impact analysis simply attempts to estimate what the dollar impact of a new intervention would look like under current conditions (which unfortunately include underdiagnosis and undertreatment).
Global Healthy Living Foundation	
As we have stated in previous draft report public comments, GHLF remains very concerned about the approach ICER takes when evaluating "value". Our organization represents patients	Individual patient values should always be taken into account when clinicians are making decisions with patients. This sort of individual

suffering from chronic diseases, including arthritis and osteoporosis, and we find the lack of the patient perspective and the use of the antiquated QALY measurement incredibly troubling. We believe value means something different to every patient and that treatment decisions should be personal and made between patients and their doctors. We are also concerned by the lack of inclusion of a bone expert among the two clinical reviewers and as a result question the accuracy and validity of this report and perceived conflict of interest created by the source of funding for this report.	decision making is not the goal of ICER's work. Bone experts contributed to the report. ICER is funded primarily by non-profit organizations and its funding sources are shown on its website.
We have found that the overwhelming majority of ICER's reports favor the cheapest cost drug, do not take a long-term cost or outcomes view, and shoehorn analytics into an uncomfortably odd set of comparators. For example, ICER chose the intravenous bisphosphonate zoledronic acid, an established treatment protocol, but one that has obvious patient-centered flaws that make it an inappropriate comparator.	This is not always true. For instance, our report on new therapies for hepatitis C found that they represented a good value (cost/QALY < $50,000) despite their high price ($1000-$1200 per pill). Our models take a lifetime horizon and account for negotiated prices for drugs that are lower than the WAC price in order to attempt to appropriately value therapies that have significant long-term benefits that may not be adequately captured in a 2-5 year time horizon. We initially proposed using alendronate as the comparator, but received feedback from patient groups, expert clinicians, pharmaceutical companies and specialty societies that zoledronic acid was the more appropriate comparator. That said, the NMA and cost models also allow for direct comparisons between all of the drugs considered – in this case abaloparatide and teriparatide. We have removed romosozumab from those analyses, because of the new data and the delay in FDA consideration of the drug.
Compliance is assumed in the draft evidence report as 100 percent. This is not credible. Although the IV comparator used appears to have better compliance because it is given once yearly vs. the oral protocol, there is no credible way to assume 100 percent compliance as ICER's calculations do.	We explored multiple adherence scenarios including one where we "turn off" zoledronic acid (and accompanying efficacy influence) after the first year, effectively mimicking a situation in which a patient stops using ZA the first injection. However, this scenario, as well as other (lower) adherence and treatment effect decline scenarios did not produce a cost-effective result for the anabolics.
The comparator is not the best drug for people at high risk for fracture. The three drugs evaluated are. They, at varying levels, restore bone mass quickly and are designed to be used	As noted above, the three drugs can all be compared to each other as part of our analysis. We have removed romosozumab

to prevent fractures in this high-risk group. Time/value is an appropriate ratio to consider, we believe. If, while on a conventional bisphosphonate, a person suffers a fracture when one of the drugs in the study would have created bone mass quickly and prevented the fracture, what is the value of the conventional bisphosphonate? What is ICER's acceptable fracture rate while on bisphosphonates and where is the calculation that weighs the cost of these fractures vs. the cost of preventing them? These questions, and other more specific issues, such as whether ICER assumed the benefits of bisphosphonate at treatment initiation vs. when those benefits actually occur.	from the final report because of the new trial results and delay in FDA consideration of the drug. Zoledronic acid is parenteral as are the anabolic agents and it has been studied and is recommended for use in high risk populations (for instance see the AACE/ACE 2016 guidelines). As described above, evidence from the clinical trials of anti-resorptive therapies, like the HORIZON trial, find that the fracture reduction benefit begins soon after treatment.
What kind of fractures is ICER valuing? It appears to us that hip fractures have been chosen as the weighted favorite. However, vertebral fractures are more common. We are also unsure whether the costs are amortized to 2016 dollars.	The costs and utilities associated with hip, vertebral (clinical and morphometric) and other fractures are weighted by their probability of occurring. Costs are expressed in 2016 dollars.
Every day, patients look to our organization for help because they do not have access to their medications. Our fear is that insurance companies will cite ICER, and their flawed methodology, when making coverage decisions, further limiting the already poor access to new, innovative, and life changing therapies. While we recognize that actions need to be taken to address the high cost of medications and are appreciative of ICER's transparency of their funding sources, we believe their ties to the insurance industry impedes their ability to create a neutral framework.	We are glad that GHLF recognizes that the high costs of medications need to be addressed.
American College of Rheumatology	
The ACR supports efforts to define cost-effective and clinically appropriate strategies to manage patients with OP and therefore we appreciate ICER's work in addressing this important topic. We would, however, like to share some concerns about the methods used to generate the report and point out the potential for serious unintended negative consequences for OP patients.	Thank you for providing detailed comments on ICER's draft analysis.
ICER, in this draft report, relies on clinical trial data as the highest form of evidence. Such studies are entirely appropriate for the purpose of demonstrating efficacy and identifying prominent safety signals. However, because of strict inclusion criteria, the patient populations in clinical trials often fail to adequately represent the complexities of real-world patients. Specialists, including rheumatologists, often take care of such patients and manage severe OP in post-menopausal women who also suffer from complex comorbid conditions and have multiple risk factors for progression of OP and fractures (including glucocorticoid use and rheumatoid arthritis). Clinical trial data must therefore be supplemented with data from	We have included additional real-world data on teriparatide in the Evidence Report. There are no real-world data on abaloparatide, nor on romosozumab. In general, patients in clinical trials are more adherent and have better outcomes than the general population. Thus, our approach is likely to bias the findings in favor of the drugs. Furthermore, two of the drugs only have clinical trial data, so we are comparing all four drugs (now three given the deferral of comparisons with romosozumab) on a level playing field.

"real-world" patient populations in order to arrive at conclusions that have validity and broad applicability in the clinical setting. In this context, we think the preference for clinical trial data predisposes the ICER report to underestimate the value of anabolic therapies.	
This shortcoming is reflected in the voting questions that were given to the panelists. The questions define high risk of fracture as being both a fragility fracture and T score <= -2.5. First, the presence of a fragility fracture of the spine or hip is sufficient in the absence of a low T score to define high risk. Furthermore, the definition fails to take into account other conditions as mentioned above that increase the risk of fracture. Thus, the reports takes a very narrow focus on populations in clinical trials while missing the diversity and complexity of patients that rheumatologists see in practice every day.	Based on feedback, we have updated the definition of the patient population under consideration in the voting questions to match the labeled indications for the anabolic agents, which is a lower risk population than we originally proposed. We think that it is more likely that these drugs will be reserved for a population at higher risk than the overall population of patients at high enough risk to warrant treatment for osteoporosis. We hope that a patient population with exceptionally high risk can be defined in the meeting.
Clinical trial data also frequently fail to capture long-term outcomes and, in the case of OP, a drug's long-term effects on bone architecture. In this way too, the methodology of the ICER report may predispose it to underestimate the benefits of anabolic therapies. Additional risk stratification and sensitivity analysis may have revealed scenarios that better reflect current and medically appropriate use of anabolic therapies in clinical practice.	Our approach models the benefits and harms of the drugs and disease process over a lifetime. Given the paucity of long-term data it is inevitable that we underestimate some benefits and harms and overestimate others. We are using the best data available based on input from the pharmaceutical companies, specialty societies, patient advocacy groups, and others.
It should also be noted in the report that anabolic therapies are faster acting and can stabilize a patient more quickly than bisphosphonates – a characteristic of anabolic agents that is of vital importance to patients with severe OP. Furthermore, the model does not consider the importance of published data (Leder B 2015 Lancet 386:1147) that suggest that a patient's response to anabolic agents may be blunted by prior therapy with an anti-resorptive. For this reason, as well as their more rapid onset of action, a clinician and patient may appropriately choose an anabolic agent as first-line therapy in high-risk scenarios.	The Kaplan Meier curves from the Horizon trial demonstrate almost immediate benefit in terms of fracture prevention with zoledronic acid. There are no published data demonstrating more rapid benefits with anabolic agents. As noted above, there is a debate among osteoporosis specialists about the evidence supporting anabolic agents as first line therapy (Cosman 2017 and responses). This is a promising hypothesis that is supported by BMD data, but awaits confirmation in clinical trials powered to assess fractures. It is also disturbing that we lack data supporting a benefit for anabolic therapy in reducing hip fractures.
Finally, we note that only two clinical experts were utilized as reviewers for this effort. We do not believe that this number is adequate for an analysis of this scale and complexity and would like to suggest that the report would have benefited from input from additional experts and practitioners involved in patient care in a variety of settings.	We worked closely with teams of experts from each of the manufacturers (Lilly, Amgen, Radius) as well as a diverse group of experts assembled by the National Bone Health Alliance.

American Society for Bone and Mineral Research	
The report has the potential to serve as a "wake up call" to address the current crisis in the under-treatment of osteoporosis by examining the value of new therapies. However, we have concerns about the timing of the report, given the current landscape, and the availability of evidence upon which analyses can be conducted at this time.	It should be a wake-up call for under-diagnosis as well as under-treatment.
Timing of Report and Paucity of Data The ICER report is a comparative analysis of the effectiveness and cost effectiveness of two new anabolic osteoporosis therapies (abaloparatide and romosozumab) as well as an established one (teriparatide) compared to the antiresorptive medication, zoledronic acid. Unfortunately, as the report's authors point out, there is currently very limited comparative efficacy data on which to draw conclusions. Instead, the report relies primarily on a limited number of placebo controlled trials and utilizes a network meta-analysis approach requiring numerous assumptions. We feel that this is problematic for several reasons. First, while there are limited published comparative efficacy data that include a fracture efficacy endpoint, there are several studies that do compare anabolic therapies directly to antiresorptives with validated surrogate endpoints such as bone mineral density (BMD) that could have provided additional data to consider. Moreover, given that there are currently 2 completed comparative efficacy trials – VERtebral Fracture Treatment Comparisons in Osteoporotic Women (VERO) and ARCH (Active-contRolled FraCture Study in Postmenopausal Women with Osteoporosis at High Risk of Fracture) – that, when published, will likely provide some of the most pertinent data on which to base any conclusions, it seems prudent that ICER delay such a report until after this data becomes available.	1. The outcomes that matter to patients are fractures, not change in BMD. 2. We have added a summary of the available data on the unpublished VERO and ARCH trials. 3. There will always be new studies about to be presented or published. As many of the comments above highlight, patients and providers are clamoring for access to the new agents. Given the recent FDA approval of abaloparatide, the timing seems appropriate. 4. Given the unexpected harms observed in the ARCH trial that have delayed FDA consideration of romosozumab, we have removed romosozumab from the NMA, economic models, and voting questions.
Moreover, the recent announcement regarding potential safety concerns with romosozumab that will delay its FDA approval also supports the impression that the current report may be significantly premature.	As noted above, we have removed romosozumab from the comparative effectiveness and economic models as well as the voting questions given new data available after the release of our draft report.
Long Term Benefits of Anabolic Therapies and Importance of Drug Sequence The current analysis in the draft report, which assesses benefits for only up to 5 years, does not take into account the potential for long-term benefits of anabolic therapies. The analysis also does not take into account that virtually all patients who are treated with anabolic agents are treated with antiresorptives at some point in their treatment course as well. Thus an analysis that assumes a single course of anabolic therapy is of limited clinical relevance. Finally, the report does not appear to recognize the growing body of evidence that the sequence in which anabolic and antiresorptive therapies are administered	This is incorrect on multiple points. First, the analysis assesses both benefits and harm over a lifetime horizon to fully capture the long-term benefits of therapy. Second, the model assesses anabolic therapy (2 years of either teriparatide or abaloparatide) followed by 6 years of zoledronic acid. Third, it has long been recognized that BMD response to therapy is a poor marker of the

has a profound effect on resultant bone strength. To illustrate this point, it was recently reported in a randomized controlled trial that 2 years of teriparatide followed by 2 years of the antiresorptive, denosumab, increased femoral neck BMD by over 8% in postmenopausal women whereas the same drugs given in the opposite order resulted in a femoral neck BMD increase of less than 5%, which was statistically significantly lower.	impact on fracture reduction. For instance, the Kaplan-Meier curves in the HORIZON trial demonstrate almost immediate benefit in terms of fracture reduction (vertebral, non-vertebral, and hip) despite very modest short-term effects on bone density. Similar findings are apparent for the anabolic agents.
Voting Questions The questions for the panel appear to be overly-simplified and do not adequately address the long-term value of anabolic agents in the context of a patient's long-term treatment course with multiple agents. Questions should be added to address the value of this approach.	The value questions are grounded in economic assessments based on a lifetime horizon – there can be no longer term assessment of value.
Expert Review The process of peer review is an underlying fundamental of the scientific endeavor. This report should not be published without that type of rigorous external peer review that we all abide by. There were only two clinical expert reviewers for this report and we believe that it would benefit from having additional input from physicians from multiple specialties, including those with extensive real-world experience in the treatment of osteoporotic patients.	The draft scope, initial results, and draft report have been reviewed by each of the involved pharmaceutical companies and we have had at least 3-4 phone calls with a team of experts from each of the companies (Amgen, Lilly, Radius). The NBHA put together a panel of more than 10 experts in all aspects of osteoporosis who provided input on our review. The many other stakeholders represented in this summary of comments received have had input at multiple steps in the development of this report. Finally, we look forward to additional input at the public meeting from all interested stakeholders.
Bendcare, LLC	
The question we must ask is: what conditions must be met for ICER to conclude that the evidence is inadequate to address comparative clinical effectiveness and value at this time?	ICER reports frequently conclude that evidence is insufficient. However, clinicians and patients must currently make decisions between approved therapies, and so it is not helpful to refuse to address comparative effectiveness and value using the best evidence currently available, while noting that uncertainties exist.
We urge you to reconsider making any recommendation based on the methodologies you employed to evaluate evidence on effectiveness and value for anabolic therapies for osteoporosis in postmenopausal women. The only empirically supportable recommendation you can make is that further studies are needed. It would help to include details on proposed study designs.	As mentioned, clinicians and patients are currently making these decisions with regard to treatment.
First, there are only three studies which meet your evidence requirements. Each of them compares a single drug to placebo. No studies compare drugs to one another. "Comparative Effectiveness" of any drug to any other drug is generated as a probabilistic projection based on statistical features extracted	A statement that evidence is not adequate to show that drug A is different from drug B is not the same as stating that evidence shows that drug A is the same as drug B.

from the three drug vs. placebo studies. The cited three drug trials also have significant sample-selection constraints on the clinical and demographic variability of patients studied. Is ICER seriously recommending anything to clinicians in the utter absence of germane evidence for the specific clinical decision to select a single agent to treat a far more diverse patient population? Consider the form this might take: (ICER to doctor): "Abaloparatide, Teriparatide and Romosozumab are clinically not different." (doctor to ICER): "How do you know?" (ICER to doctor): "Probabilistic projections from three drug vs. placebo trials." (doctor to ICER): "Sounds like you don't know enough to make any recommendations." (ICER to doctor): "Yeah, but we have a great methodology!"	
Second, much effort was devoted to the study of payor osteoporosis policies. However, payor policies can be—and often are—changed at a moment's notice. Moreover, payors frequently aggressively incentivise clinical decision making in a fashion that minimizes short-term expense, with little or no regard to improving long-term patient outcomes. Physicians in our group frequently report that payors deny approval, despite conditions being met for prior authorization. Not infrequently, prior authorization is received and then payment is clawed-back later. How payor policies actually manifest themselves in a population of real-world patients is dramatically more important. If those policies are seldom followed, with clinical management more frequently denied or delayed, they become decoupled from clinical relevance when considering expense of a particular drug.	ICER reports include this information for background and to assist in policy discussions at its public meetings.
Third, expense calculations for bad outcomes associated with osteoporosis fractures were assumed to be normally distributed. However, those of us that have spent decades with population-wide clinical encounter data know that disease-specific expense is highly skewed. Specifically, the moments of the distribution that best characterize disease-specific expense distributions are first: skewness, second: variance, least: mean. The evidence used by ICER may be inadequate to characterize mean expense. They are grossly inadequate to characterize skewness.	We used a log-normal distribution for cost data and the variance and mean as reported in the underlying studies. We also confirmed that the model results are not sensitive to using a gamma distribution.
For these reasons, ICER will be doing a serious disservice by making any recommendations based on the evidence cited. The only defensible recommendation is that further research is needed. The recommended research must include, at a minimum, head-to-head comparisons of Abaloparatide, Teriparatide, Romosozumab, and bisphosphonate. An example of one such study is the STRUCTURE trial, which compares Teriparatide to Romosozumab. Unfortuantely, the results from this study are not yet fully published. We are not aware of any	With newer therapies, whenever a report is published new evidence will become available in the future. Clinicians and patients still need to make treatment decisions now.

Abaloparatide vs. Teriparatide, Abaloparatide vs. Romosozumab, and all three vs. bisphosphonate.	
Furthermore, ICER must develop criteria to specifically identify topics for which evidence is inadequate to make any recommendations. These criteria should include: (1) "Evidence Adequate," (2) "Evidence Marginal," and (3) "Evidence Inadequate." As a validation step, these proposed ICER criteria must find the current topic to be "Evidence Inadequate" as of June 1, 2017.	See above.
Coalition of State Rheumatology Organizations	
General Draft Report Concerns In January, the Institute for Clinical and Economic Review (ICER) released an initial report on the coming assessment of comparative clinical effectiveness and value of three anabolic treatments for osteoporosis, one of which has been on the market for over 14 years (teriparatide), one that was only very recently approved (abaloparatide) and a medication that has not yet been approved (romosozumab). While the concept of comparative effectiveness is a rational approach for measuring value, the CSRO feels that this process must use comparisons with a more appropriate method.	As noted above, the PDUFA dates suggested that the FDA was likely to approve two new anabolic agents in the first half of 2017. It seemed an appropriate time to evaluate the anabolic agents as a class and feedback from stakeholders (pharma, clinical experts, specialty societies, and patient advocacy organizations) suggested that zoledronic acid was the most appropriate comparator.
ICER is comparing a group of bone-forming agents to a bisphosphonate – zoledronic acid, which is an antiresorptive agent. Its use can be limited by renal impairment which is common in the population that has osteoporosis. Additionally, in clinical practice these anabolic agents are primarily used to target the highest risk patient populations. The anti-resorptive drugs have a different mechanism of action, different onset of action and different lengths of treatment compared to anabolic therapies. The document states that these comparisons were valid as all of these drugs are approved for patients who are at high risk of fracture. Another drug which is also approved for treating patients at high risk of fracture, denosumab, was not included in this comparison for reasons that are not stated clearly. The CSRO also questions the definition of high risk of fracture as a T-score of -2.5 or lower AND a history of a fracture. In the clinical trials of these drugs, the definition of high risk of fracture does not include both of these criteria.	You are correct. Zoledronic acid is also recommended for the highest risk group (see AACE/ACE guidelines for example). We have changed the explanation for high risk in the voting questions to match the FDA indications for teriparatide and abaloparatide. We welcome input at the public meeting to establish a clear definition of "the highest risk patient populations." It would be helpful to identify the group of patients in whom first line anabolic therapy is indicated. For the group of patients with significant renal impairment, we agree that denosumab would be a more appropriate comparator, but this subgroup, while important, is less policy relevant than the larger group of patients with osteoporosis.
The CSRO also questions why morphometric fractures were not included as an outcome as morphometric fractures were the primary endpoints of all of the clinical trials for osteoporosis agents. Morphometric fractures are associated with increased morbidity including pulmonary disorders related to kyphosis, increased risk of subsequent fracture both vertebral and non-vertebral and increased fall risk, all of which have related costs that were not included in this model. Other costs that were not considered in this model include but are not limited to the cost of surgical repair of fractures, the cost of vertebral	We performed two NMAs – the first is of morphometric fractures, which were the primary endpoint of most of the pivotal trials. In the model we have added a scenario analysis that explores the addition of morphometric vertebral fracture disutility. This scenario showed little difference compared to the base case results, as most of

augmentation procedures, the cost of rehabilitation post-fracture, the cost of medications used to treat pain associated with fractures and the cost of treating comorbid conditions associated with these fractures such as pneumonia, pulmonary emboli, deep vein thromboses and bleeding post-hip fracture.	the differences in QALYs were canceled out among the comparators. The costs estimates used are taken from previous analyses that have included all relevant costs associated with a fracture including nursing home stay, physician visits for adverse events, etc. In addition, the model results are not sensitive to changes in these costs as its outcomes are mainly driven by the differences in relative fracture risks.
E. Michael Lewiecki, MD	
1. Clinical evidence a) Limitations and uncertainties. There is additional evidence on comparative effectiveness, not included in the ICER report, including a study showing superiority of teriparatide compared with risedronate in reducing vertebral fracture risk. Evidence is continuing to emerge for abaloparatide, a recently approved anabolic agent. Romosozumab, a sclerostin inhibitor, has been shown to have a dual-effect on bone remodeling, stimulating bone formation and inhibiting bone resorption, and therefore is not an anabolic agent in the same sense as teriparatide and abaloparatide. Recent developments regarding romosozumab may result in additional analyses of past studies and delayed FDA review. The fast moving pace of new data with these agents and many uncertainties suggest that it may be premature to reach any conclusions regarding comparative effectiveness.	We have included a summary of the unpublished trial comparing teriparatide to risedronate in the final report. We agree that romosozumab has a unique mechanism of action. We have delayed full consideration of romosozumab in light of the delay in FDA consideration of the drug.
b) Anabolic effects on bone structure. It should also be noted that anabolic treatments for osteoporosis have beneficial effects on bone structure and bone strength that are not fully captured by bone density tests with DXA and cannot be measured by standard available clinical tools.	We agree. The full effects should be captured in clinical trials powered for fracture outcomes. We look forward to additional head-to-head trials powered for hip fractures as well as vertebral and non-vertebral fractures.
c) Order of treatment. There are accumulating data that the sequence of treatment (e.g., anabolic before antiresorptive, anabolic after antiresorptive, anabolic combined with antiresorptive) may have important implications with clinical effectiveness.	We agree that there are suggestive data from BMD and bone turnover markers, but as you note in your prior comment, these measures do not fully capture the benefits and harms of the drugs. Grey et al, in response to Cosman et al 2017, offered a cogent explanation of why the evidence is not yet sufficient to conclude that anabolic therapy should be the

	standard first line therapy for patients with osteoporosis. *Grey et al 2017 PMID 28294409*
d) Clinical trials vs. clinical practice. Patients in clinical practice are often not the same as subjects who participated in pivotal fracture trials. Differences in comorbidities and preferences play an important role in the individualization of treatment decisions that should not be overly constrained by regulatory issues.	These issues are particularly important to bring forward during the policy round table discussion of the public meeting.
2. Long-term value for money a) Uncertainties of pricing. Teriparatide is likely to be available in generic form within the next several years and presumably have a different pricing structure than the current brand name product. The recently announced retail price of abaloparatide is considerably lower than brand name teriparatide. Expected pricing for romosozumab, an investigational agent, is unknown.	Prices of drugs change over time, and generally not in predictable ways.
b) Cost-effectiveness modeling. Any assessment of cost-effectiveness, if a valid assessment can be done at all, must consider pricing uncertainties of existing, new, and emerging anabolic agents, the sequence of anabolic and antiresorptive therapies, and the use of antiresorptive agents other than zoledronic acid, such as denosumab.	The model considers treatment sequencing of anabolic and antiresorptive agents and clinical experts have confirmed zoledronic acid to be an appropriate agent. We considered including denosumab in our initial scope, but were received feedback that we should focus on the anabolic agents and that the appropriate comparator for the anabolic agents was zoledronic acid, not denosumab. We agree that prices of drugs change over time, and generally not in predictable ways, and therefore we subjected prices to sensitivity analysis. In addition, we provide a "value-based" price so readers can assess whether they believe a future price may be close(r) to the value-base price or not.
3. Questions for deliberation a) The selection of zoledronic acid as the only choice for antiresorptive therapy does not reflect real-world clinical practice, where other agents, particularly denosumab, are often used in high risk patients.	Please see our comments above. We initially included denosumab in our scope and had alendronate as a comparator. Input from the pharmaceutical companies, patient advocacy organizations, specialty societies, and clinical experts recommended limiting the assessment to anabolic agents with zoledronic acid as the comparator.
b) The definition of high risk solely according to T-score fails to include many high risk patients with T-scores better than -2.5.	We have attempted to clear up this misunderstanding. Patients who warrant treatment for osteoporosis include those with T-scores less than -2.5 AND those with fragility fractures who have T-scores greater than -2.5. By high risk, we intended to mean exceptionally high risk: those who merit initial treatment with drugs other than the typically recommended oral bisphosphonates. For the

	voting questions, we have changed the language to the FDA indication for abaloparatide and teriparatide.
c) Comparison of "net health benefit" with a limited choice of options is an artificial constraint that is not representative of clinical practice, where many patients have been on multiple osteoporosis medications at different times and in combinations and sequences that could have variable effects on bone strength and fracture risk.	For clarity and simplicity, we are comparing first line anabolic therapy to first line zoledronic acid. This biases the results in favor of anabolic therapy: as many have commented above, it appears that anabolic therapies may not be as effective if used after patients have been treated with anti-resorptive therapy. There are inadequate data to model the impact of anabolic therapy followed by anti-resorptive therapy on patients with more complex prior treatment histories.
In summary, the numerous uncertainties and limitations of the data should lead the ICER report to be cautious in reaching conclusions of comparative effectiveness and cost-effectiveness. From what is known of the effects of anabolic agents on bone structure and bone strength, their use in proper sequence with highly effective antiresorptive therapy provides the best treatment option for appropriately selected high risk patients.	We agree that the data on the value of anabolic therapies are remarkably limited, particularly for hip fractures and for determining the appropriate sequencing of therapy. However, two of the drugs are approved and clinicians and patients must make decisions about therapy based on the current state of the evidence. Given the recent approval of abaloparatide, this is a particularly apt time for a comparative effectiveness review. We look forward to your input on the definition of "appropriately selected high risk patients."
National Bone Health Alliance	
The NBHA and its' members are concerned that complicated patients with osteoporosis, patients who are excluded from randomized clinical trials, are not adequately considered in the report. We understand ICER's intent is to have a process that is strongly based on the highest level of evidence; however, for patients at highest risk for fracture, using and even starting with an anabolic therapy may sometimes be the best choice and based on as yet unpublished clinical trial data, substantial observational data, and an understanding of bone physiology and the mechanisms of action of these drugs. We have detailed the evidence supporting our strong recommendation on this below.	We have greatly benefited from the input of the NBHA and its members throughout this process and look forward to your additional input at the public meeting.
Limited Data The report does not include more recent data such as The VERO study which was presented at WCO this year. VERO compared teriparatide to risedronate over two years with 680 patients per group in a double blind double dummy trial. Patients in the teriparatide arm had fewer vertebral fractures (5.4% vs 12%, $p<0.001$) at two years. Equally important the difference was seen at one year as well (3.1% vs 6.0%, $p<0.05$) showing the	We have included a summary of the VERO trial in the updated report.

rapid onset of action of this anabolic therapy. Anabolics are more effective than antiresorptive therapy in terms of fracture reduction and have a more rapid onset. We understand that the VERO study has not yet been peer reviewed, but because there are so few studies, we believe it should be noted.	
Timing of Report We understand the impetus to review anabolic therapies was the potential addition of two new anabolic agents in the marketplace Abaloparatide has just been approved, but review of romosozumab is now delayed. It seems too early to be able to adequately assess the role of these anabolics in the treatment armamentarium.	As noted above, romosozumab has been removed from the NMA, economic analyses, and voting questions. Since abaloparatide is now available for clinical use, a comparative effectiveness review seems particularly timely.
With the delay in the review of romosozumab, the discussion regarding this agent should be minimized. Furthermore, there are two large studies, VERO and ARCH, which when published, may help us better understand the role of anabolic agents.	We have reduced the role of romosozumab in the review and added a summary of the VERO and ARCH studies.
Modeling The model compares one or two years of anabolic therapy to three years of an antiresorptive therapy (the HORIZON trial). It is not fair to compare therapies of different duration. One should compare equal duration of exposure, which would include one or two years exposure to an antiresorptive. It is understandable that there would be fewer hip fractures in a smaller trial of shorter duration. The registration trial for teriparatide was cut short because of safety concerns, although further surveillance has not demonstrated increased risk for osteosarcoma. Nonetheless, the lack of longer term studies with teriparatide make comparisons with anti-resorptive therapy studies that much more difficult.	The model compares 6 years of zoledronic acid to 2 years of anabolics followed by 6 years of zoledronic acid. See figure 4 in the report. This is considered fair as it best reflects the use of these agents as recommended in clinical guidelines and use in clinical practice. The lack of longer term studies should not preclude a comparative (cost-)effectiveness analysis now. With newer therapies, whenever a report is published new evidence will become available in the future. Clinicians and patients still need to make treatment decisions now.
The authors of the report used nonvertebral fracture data to model for hip fracture reduction. Typically nonvertebral fracture reduction is lower as compared to hip fracture reduction with an osteoporosis therapy. A sensitivity analysis should be done to account for greater hip fracture reduction with an anabolic. We have both observational data in almost 9000 patients (Silverman presented at WCO 2017) and claims data (Burge) which suggest hip fracture reduction as high as 56% with an anabolic, teriparatide. This becomes important since, of all fractures, incident hip fracture is associated with the greatest loss of health utility (define – death, independence, QOL).	We have modeled reductions in hip fracture rates with the anabolic agents that are greater than those observed for non-vertebral fractures despite the complete lack of randomized trial data on the efficacy of abaloparatide and teriparatide on hip fractures. This is a significant bias in our model in favor of anabolic therapy. The lack of clinical trial data on hip fractures is the reason that the 2017 ACP Guidelines did not list any of the anabolic agents as first line therapy for the treatment of osteoporosis. We hope the the NBHA will support efforts to perform clinical trials with sufficient power to demonstrate a reduction in hip fractures for new drugs given your observation that “of all fractures, incident hip fracture is associated with the greatest loss of health utility (define – death, independence, QOL).”

The model assumes no “ramp up” for use of anabolics. Several studies have shown that prior treatment with an antiresorptive may blunt the BMD response of anabolics. Clinicians seeing a patient at very high risk for fracture might want to start with an anabolic first rather than an antiresorptive to realize the greatest bone density gains.	You are correct. The lack of a ramp up period biases the results in favor of the anabolics. Essentially, we modeled the anabolic as first line therapy as promoted by some experts in the field (F. Cosman et al, 2017). We hope at the meeting you can help define what constitutes “a patient at very high risk for fracture” who might benefit from parenteral rather than oral first line therapy.
The clinical trial data do not account for the differential effects on bone architecture seen with anabolic vs antiresorptive therapy. Some patients have poor bone quality as well as quantity at baseline and anabolics such as teriparatide have been shown to improve bone quality more than antiresorptive therapy, and this may then result in greater bone strength.	That is correct. The outcomes that matter to patients are fracture outcomes. Changes in BMD and microarchitecture are only important if they translate into a difference in fracture outcomes. These intermediate outcomes are important in the rapid identification of the most promising drugs and doses, but must be confirmed with trials demonstrating improved efficacy at preventing fractures.
There is a need for further risk stratification and sensitivity analyses.	It is unclear if this applies to the network meta-analysis and comparative effectiveness review or to the cost model. The paucity of trials and the lack of trial data stratified by risk makes this impossible in the evidence review section. More detail about the suggestions would be appreciated. The model includes deterministic, probabilistic and multiple scenario analyses considering various risk groups and other variables.
The basic model structure and assumptions used to complete the model-based cost-effectiveness analyses (CEA) should be more clearly and consistently stated. For a reviewer interested in the actual model structure and underpinnings, the report contains insufficient detail and could not possibly be reproduced from the information provided. Further information on model structure should be provided in appendices.	We have added several explanations and provided more detail on the CEA to clarify our approach and improve consistency in the report.
An important area for further development relates to model validation. Model validation typically includes comparisons of modeled outcomes vs. epidemiological or other source data.	Where data availability allows, this has been done. However, as data on anabolics are particularly scarce, limiting the possibility to do external validation, we have also reported how our results compare to other published models (cross validation).
The implementation of the fracture hierarchy as described in the draft report is problematic. As implemented at the time the draft was released the assumption effectively rendered the assertion that all tracked fractures were meaningless. This was	It is unclear what is meant by “all tracked fractures were meaningless”. From a post-fracture state, patients can transition to a worse fracture state only (or death). The

discussed with the modeling team who says it is being/has been changed, but this requires further scrutiny and review of model validation results.	hierarchy for fracture severity is hip > vertebral > other. Patients may also have a morphometric vertebral fracture but we assumed they do not change health states, due to the negligible cost and QALY impacts of morphological vertebral fractures; we explored a potential QALY loss for these patients in a scenario analysis.
Transparency The document needs to help the reader understand why ICER chose to review these osteoporosis medications and explain all sources of ICER funding.	The primary reason for choosing these agents at this time was because of the expected FDA consideration and approval of both abaloparatide and romosozumab in late Spring 2017. Teriparatide was included as the only other anabolic agent. The comparator was chosen based on extensive discussions with and feedback from the manufacturers of the anabolic therapies, specialty organizations such as the NBHA and the AACE, patient organizations, and discussions with specialists including endocrinologists, rheumatologists, and directors of osteoporosis specialty clinics. All sources of ICER funding are described in detail on their website: https://icer-review.org/about/support/ Note: funding is not accepted from manufacturers or private insurers to perform reviews of specific technologies.
Voting Questions The voting questions are dichotomous and ask the panelists to vote on the net health benefit for postmenopausal women with osteoporosis as defined by a T score <=-2.5 and a fragility fracture, which is only a subset of the high risk population.	The voting questions have been reframed to specifically ask about the population covered by the FDA indication.
The definition of high risk should be reexamined. We accept that patients with a fragility fracture of the hip or spine have osteoporosis independent of their bone density (Siris 2016). Studies have shown that almost half of patients with osteoporotic fracture have a BMD above -2.5, indicating that a T-score of <= -2.5 should not be the sole defining factor for risk. High risk may also be defined by clinical risk factors such as glucocorticoid exposure or patients with multiple fractures. Net Health Benefit ignores the urgency to treat in some patients. Patients who have had a prior fracture and those with multiple fractures have a substantially increased risk for future fracture, and the risk for subsequent fracture is greatest in the 2 years following the initial fracture. These patients may thus need the faster action of an anabolic, e.g. the 11.8% increase in lumbar spine BMD seen with teriparatide at just 18 mos.	We agree that patients with a fragility fracture as you describe warrant treatment for osteoporosis and should be considered to have osteoporosis as well as those patients with a T-score ≤ -2.5. Our language was not clear. We would like to identify the population at high enough risk to warrant treatment with a parenteral drug rather than the oral agents that are typically recommended by guidelines as the primary therapy for the typical patient with osteoporosis. We hope that population can be better defined during the public meeting.

	As noted above, the fracture efficacy of zoledronic acid starts almost immediately following treatment (please see the Kaplan Meier curves for incident fractures in the HORIZON trial), despite a longer time horizon for change in BMD. What matters to patients is reducing their risk for fractures, not changing their BMD.
These voting questions ask panelists to vote only about a subset of the population at risk for further osteoporotic fracture	The voting questions have been reframed to specifically ask about the population covered by the FDA indication.
The votes using the current voting questions should not be interpreted as providing any clinical guidance to clinicians or payers.	The votes of the CTAF and the policy recommendations from the roundtable discussion are intended to provide context to all stakeholders.
United Rheumatology	
We applaud ICER's initiative to bring multiple stakeholders together, not in the interest of restricting access to critical and effective treatments, but to determine how to best address the challenge of rising drug costs. UR does, however, have fundamental concerns regarding ICER's analysis as well as specific questions regarding the modeling assumptions that were employed.	
United Rheumatology disagrees with ICER's attempt to define patients at high risk for fracture as those who have had a prior fragility fracture AND have evidence of osteoporosis on a DXA study. Although alluded to in ICER's draft Evidence Report, this exact definition is not stated until the Questions for Deliberation (footnote 1): "High risk for fracture defined as the presence of a prior fragility fracture and a bone mineral density T-score of -2.5 or lower." From a practical standpoint, it is impossible to answer the 6 questions posed as none of the clinical trials used this definition as inclusion criteria for enrollment. As outlined in the list below, all included studies offered alternate definitions for identifying patients at high risk for fracture.	We agree that we were not clear. The intent was to describe a group of women at exceptionally high risk for fracture. Clearly women with a prior fragility fracture or a T-score less than -2.5 have osteoporosis and warrant treatment. Most guidelines recommend oral bisphosphonates as first line drug therapy (in addition to calcium, vitamin D and weight bearing exercise). We were attempting to define an exceptionally high-risk group of women with osteoporosis, based on feedback received, who would warrant first line anabolic therapy rather than oral therapy. We have changed the language in the voting questions to describe the same population described by the FDA indication.
Patients in the pivotal teriparatide trial (Neer et al NEJM 2001) were enrolled if they had at least 1 prior vertebral body compression fracture; BMD was not an entry criteria. The mean lumbar spine T-score at baseline was -2.6, so a significant number of women had T-scores of better than -2.5.	We agree.
Patients enrolled in the ACTIVE trial (abaloparatide; Miller et al NEJM 2016) were included if the lumbar spine or femoral neck BMD was between -2.5 and -5.0. Fractures were not an entry	We agree.

criteria and in fact 37% of patients at baseline had no prior fractures.	
The FRAME trial (romosozumab; Cosman et al NEJM 2016) also enrolled based on BMD (here a T-score of between -2.5 and -3.5 in the total hip or femoral neck). Approximately 18% had 1 or more vertebral fractures and 21.8% had a non-vertebral fracture at baseline; so at least 60% did not have a baseline fragility fracture.	We agree.
The HORIZON trial (zoledronate; Black et al NEJM 2007) enrolled patients if they had femoral neck T-score of less than or equal to -2.5 with or without vertebral fracture OR if T-score was less than or equal to -1.5 then at least 2 mild vertebral fractures or at least one moderate vertebral fracture had to be present. 72 % of patients had BMD of < -2.5 and 63% of patients enrolled had evidence of a prior vertebral fracture.	We agree.
No other organization has defined high risk for fracture as defined in ICER's Questions for deliberation. The FDA defines high risk for fracture as a "history of osteoporotic fracture, or multiple risk factors for fracture; or patients who have failed or are intolerant to other available osteoporosis therapy," and currently applies this label to teriparatide, abaloparatide and densoumab. The American Association of Clinical Endocrinologists (AACE) defines patients at high risk for fracture as: 1) those with a prior fragility fracture; or 2) with low bone density and additional risk factor(s) including advanced age, frailty, glucocorticoids, very low T-scores, or increased fall risk.	We have adopted the FDA language for the voting question.
In addition to its lack of precedence, UR is concerned that the ICER definition, which requires not only a prior fragility fracture but also a BMD T-score of < -2.5, would unnecessarily restrict coverage of these drugs and exclude patients with fragility fracture(s) who have low bone mass/osteopenia on DXA study.	We have removed that definition.
United Rheumatology disagrees with ICER's approach to compare anabolic therapies to the anti-resorptive drug zoledronate for the treatment of patients at high risk for osteoporotic fracture. Herein, we acknowledge a change in thinking from our earlier public comments that advocated studying all osteoporosis drug therapies used in women at high risk for fracture including denosumab (Prolia) and zoledronate as well as teriparatide (Forteo), abaloparatide (Tymlos) and Romosozumab (Evenity). Having also been involved with the ICER Evidence Report for Targeted Immune Modulators (TIMs) in the treatment of Rheumatoid Arthritis, which reviewed nine biologics and targeted synthetic DMARDs encompassing five distinct mechanisms of action in patients with moderate to severe rheumatoid arthritis, we initially thought that a similar broad approach should be employed in osteoporosis. The critical difference in osteoporosis is that the structural effects on the target organ (bone) are profoundly different for anabolic	Thank you for the input. It is common practice to compare drugs with different mechanisms of action when the share a common indication. We received substantial input recommending the use of zoledronic acid as the comparator in this review.

as compared to anti-resorptive therapies whereas – in the joint – different TIMs could lead to similar effects on the inflamed synovium.	
We would strongly encourage ICER and those viewing the public comments from stakeholders to once again review the elegant and masterfully written letter from Drs. Felicia Cosman and David Dempster in their roles as Co-Editor and Associate Editor respectively of Osteoporosis International dated December 23, 2016 and posted on the ICER website. The letter provides a cogent overview of: the head to head trials between anabolic and anti-resorptive therapies that favor anabolic drugs; the significant fracture benefit that occurs within 12 to 18 months of anabolic therapies; and the bone biopsy data that shows that anabolic drugs can restore microstructural integrity rather than simply preventing further structural deterioration. In addition, the importance of appropriate sequential therapy is underscored; treating with an anabolic followed by an anti-resorptive will lead to far greater improvements in BMD, especially in the hip, than simply treating patients with an anti-resorptive or using an anabolic drug after a course of anti-resorptive therapy.	We have re-read their comments. Note: the effects on fracture reduction are also rapid with Zoledronic acid (please review the Kaplan Meier curves in the HORIZON trial). Indeed the effect on vertebral fractures was highly statistically significant at 1 year ($p<0.001$, Black et al, 2007). The pivotal trials of the bisphosphonates were longer because they were powered to demonstrate a significant effect on hip fractures. The evidence base for the anabolic agents suffers from a lack of evidence on hip fractures within 12 to 18 months. We agree that the intermediate outcome data (BMD, bone biopsy) support the hypothesis that the anabolic agents may be more effective, but fracture trials like the unpublished ARCH and VERO trials are needed to confirm this hypothesis.
As demonstrated, ICER should acknowledge the roles of anabolic and anti-resorptive drugs as complementary in sequential therapy and not as comparators or appropriate substitutes for one another, which could seriously limit the positive patient outcomes achieved through their treatment as such.	We modeled this approach in our cost-effectiveness model in order to maximize our estimation of the value of anabolic therapy, but do not feel that the optimal sequencing of therapy has been convincingly demonstrated. Please see Dr. Cosman's provocative recent article promoting anabolic therapy as first line therapy and the published responses. *Grey et al 2017 PMID 28294409*
United Rheumatology disagrees with ICER's refusal to acknowledge that radiographic vertebral fractures are an important clinical outcome. ICER instead assigns them to the "non-clinical outcomes" along with BMD and bone turnover markers. In earlier written communications with ICER regarding this topic, ICER incorrectly claims that "while radiographic fractures may be a risk factor for clinical fractures, they are asymptomatic events that are not treated". In fact, radiographic vertebral fractures (or what has also been called morphometric vertebral fractures) are actively sought out by practicing clinicians and finding them can significantly alter treatment. Many current DXA machines include software that allows either single or dual energy imaging of the lateral spine; this procedure called VFA (Vertebral Fracture Assessment) is an essential element in the risk assessment, diagnosis and treatment. Depending on the age of the population studied, 16-45% of patients with low bone mass/osteopenia on DXA study have evidence of morphometric vertebral fractures. This finding changes the clinical diagnosis to osteoporosis and would lead to	Many clinical findings that are asymptomatic affect treatment decisions, often in dramatic ways (for instance, an asymptomatic blood pressure reading of 200/110). Surrogate outcomes can be extremely important for clinical decision making.

treatment where drug therapy otherwise may not have been indicated. In the spine, vertebral fractures can be graded mild/moderate/severe (Grade 1/2/3) and a spinal deformity index (SDI) can be calculated in which risk for future fragility fracture at any site increases with increasing grade and number of vertebral fractures.vi The finding of several vertebral fractures and increasing grade of fracture will occasionally lead to use of an anabolic drug as compared to an anti-resorptive. Other studies have shown that morphometric vertebral fractures of the thoracic spine are clinically significant and impact pulmonary function studies.	
ICER acknowledges some of the limitations inherent in its modelling assumptions. UR agrees that the model has a number of flaws and omissions, including but not limited to the following: it does not address the common scenario of postmenopausal women who are already being treated for osteoporosis with an anti-resorptive.	We agree that the model does not address women already treated with an anti-resorptive agent. The pivotal trials for the anabolic agents excluded women who were recently treated with anti-resorptive agents, so there are no data on which to base estimates of fracture efficacy in this population. As you are aware, BMD and bone turnover data suggest that the anabolic agents are likely to be less effective in patients pre-treated with anti-resorptive therapy.
Adherence/persistence rates are inappropriately assumed to be 100% for all drugs studied. ICER states that there is "a lack of real world adherence data for newer anabolic agents (abaloparatide and romosozumab) and the impact of lower adherence on efficacy for all three anabolics". Yet, adherence will be far better for drugs that are administered in the clinic (zoledronate, romosozumab and denosumab) compared to those self-administered at home, especially drugs requiring a daily injection. Adherence/persistence data is available for teriparatide in the United States (74% at 6 months and 57% at 12 months)viii and there is no reason to assume it would be any different for abaloparatide. Moreover, the importance of the complexity of the treatment regimen on adherence and persistence is acknowledged by ICER in Section 5 "Other Benefits or Disadvantages".	We explored multiple adherence scenarios including one where we "turn off" zoledronic acid (and accompanying efficacy influence) after the first year, effectively mimicking a situation in which a patient stops using ZA the first injection. However, this scenario, as well as other (lower) adherence and treatment effect decline scenarios did not produce a cost-effective result for the anabolics.
In all the major trials of teriparatide, abaloparatide, romosozumab, zoledronate and denosumab, radiographic vertebral fractures are considered the primary outcome; yet, in the ICER cost-effectiveness model, cost/disutilities are only applied to clinical fractures.	We have added a scenario analysis that explores the addition of morphometric vertebral fracture disutility. This scenario showed little difference compared to the base case results, as most of the differences in QALYs were canceled out among the comparators.
In the ICER model, quality of life never improves after a fracture.	The model applies a disutility associated with a fracture for year 1 and for the years 2+ separately. The year 1 disutility is the largest and the utility improves in the years 2+. For

	hip fractures the utility jumps back up to 0.8 of the general age-specific population utility in year 2+, for vertebral it jumps back to 0.931 of baseline and for other fractures the utility is fully restored to baseline.
The model assumes that all patients are subsequently treated with yearly zoledronate for 6 years and that clinical fracture benefit appears immediately. No accommodation is made to include patients with renal insufficiency with creatinine clearance below 30-35.	You are correct. The subset of patients not eligible for bisphosphonate therapy was not the focus of the model. That group would require anti-resorptive therapy with denosumab, which is beyond the scope of this assessment.
The baseline population had a "fracture risk…similar to that observed in the clinical trials of the anabolic agents", but what that fracture risk is does not appear to be stated in the report. It appears that this fracture risk varies from what ICER defines as the high risk fracture patient with a prior fragility fracture and BMD T-score of lower than or equal to -2.5 as previously addressed within this response.	The baseline age-specific fracture risks used in the model are reported in Table 10 and the previously reported ICER definition of high-risk has been removed.
Correct the inexplicable dismissal of denosumab (Prolia), which is inaccurately grouped with calcitonin, raloxifene and estrogen as alternate anti-resorptives that "are not considered first-line therapies because of side effects or less evidence of efficacy."	Thank you. We have changed the language.
Include some discussion related to the relevance of grading of vertebral fractures since Table 2 introduces the grading criteria.	We have included these details.
Update its reference to abaloparatide (Tymlos) on page 8, within the first paragraph, to reflect that it is now approved by the FDA.	We have updated this text.
Correct the error within Table 8; the 5th row states that romosozumab is modeled based on trial data of two years; in actuality, the trial with active drug was for 12 months followed by 12 months of denosumab. This is correctly stated on page 34 under Treatment Strategies paragraph 2 and should be corrected here.	This has been corrected in the revised report.
Correct the error within Table 13, where zoledronic acid is listed as 5mg but strength in column 2 states 4mg/5ml. Zometa, zoledronate for oncology use, is available as a 4 mg dose.	This has been revised in the report in table 15. For Zoledronic acid, we use a dose of 5mg/100ml and have included only generic zoledronic acid in this report.
In the section titled "Other Benefits and Disadvantages" there is no discussion of how a patient's health insurance coverage will affect access to these drugs. A significant number of Medicare patients do not have either a low-income subsidy or supplemental health benefit that would allow them to afford the cost of Part D drugs such as teriparatide and abaloparatide purchased at an outpatient pharmacy. In contrast, since zoledronate, romosozumab and denosumab are administered in-office, they would be covered under Part B (medical benefit) and thus would be far more affordable.	These issues are discussed in the policy roundtable but are not part of what is intended in "Other Benefits and Disadvantages". The results of the discussion will be included in the written report following the public meeting.
Finally, an FDA decision on romosozumab/Evenity was initially expected on 7/19/2017 but Amgen just announced on	Romosozumab has been removed from the analysis. The other anabolic regimens as well

5/20/2017 that they do not expect FDA approval this year. The delay is related to a new cardiovascular safety signal in the ARCH study which compared 12 months of romosozumab followed by 12 months of Fosamax with Fosamax alone in post-menopausal women with osteoporosis at high risk for fracture. The incidence of positively adjudicated cardiovascular serious adverse events (SAEs) at 12 months was 1.9% in the Fosamax arm and 2.5% in the romosozumab arm (a 32% increase).ix In contrast, the FRAME study which had been submitted to the FDA did not report an imbalance in cardiovascular SAEs when romosozumab for 12 months was compared with placebo for 12 mos. with both followed by denosumab for 12 mos. The ICER model currently does not assume any serious adverse events for the anabolic therapies. It is too soon to know whether this will need to be modified.	as zoledronic acid exhibited similar serious adverse event rates compared to placebo and each other in their respective trials. These small event rate differences are unlikely to impact cost-effectiveness results. With newer therapies, whenever a report is published new evidence will become available in the future. Clinicians and patients still need to make treatment decisions now.