

# Quality assessment of traditional and conventional medicine clinical practice guidelines for osteoporosis

Jung-Hyun Kim, PhD<sup>a</sup>, Byung-Kwan Seo, PhD<sup>b</sup>, Yong-Hyeon Baek, PhD<sup>b,\*</sup>

## Abstract

Numerous studies have reported the variable quality of clinical practice guidelines (CPGs) across various domains. The aim of this study was to systematically assess the quality, methodology, and consistency of recently developed traditional and conventional medicine CPGs that focus on the management of osteoporosis and provide helpful recommendations for patients with osteoporosis.

From June 2020 to July 2020, CPGs with osteoporosis targeting any age were systematically retrieved. All CPGs of traditional and conventional medicine related to the assessment and diagnosis, management, and clinical therapeutic and pharmacological recommendations with osteoporosis were eligible for inclusion in this study. The excluded documents included guidelines without recommendations, secondary publications derived from CPGs, consensus statements, or consensus conferences based on the opinion of panelists, systematic reviews, editorials, clinical trials, and single-author documents. The quality of CPGs was independently examined by three assessors using the Appraisal of Guidelines for Research and Evaluation II (AGREE II) instrument. AGREE II consists of 6 domains; scope and purpose, stakeholder involvement, rigor of development, clarity of presentation, applicability, and editorial independence. Consequently, selected CPGs were graded as recommended (A), recommended with modifications (B), or not recommended (C), and the specific treatments and preventive recommendations in the CPGs were summarized.

The quality of the 15 CPGs assessed varied across the AGREE II domains. The overall quality ranged from 3.0 to 6.0 out of 7. The domain that had the highest scores were “clarity of presentation,” with a mean value of 69.0% (range 46%–83%); “editorial independence” had the lowest score of 30.2% (range 0%–75%). The conventional CPGs focused on pharmacological treatments, calcium and vitamin D intake, and prevention, while the traditional CPGs consistently emphasized on herbal medicine and non-pharmaceutical treatment and management.

Further development of CPGs will require improvement in domains where low item scores have been obtained in the quality assessment in this present study. Further research is needed on alternative modalities for osteoporosis, especially complementary approaches, and higher quality CPGs are needed to facilitate evidence-based clinical practice.

**Abbreviations:** AGREE = Appraisal of Guidelines for Research and Evaluation, CPG = clinical practice guideline.

**Keywords:** clinical practice guidelines, conventional medicine, osteoporosis, quality assessment, systematic review, traditional medicine

## 1. Introduction

A clinical practice guideline (CPG) was defined as “a statement developed systematically to help practitioners and patients make decisions about appropriate medical services” in 1990 by the Institute of Medicine (IOM),<sup>[1]</sup> and “a statement containing a systematic literature review of clinical evidence and recommen-

dations made to assess the benefits and risks of alternative modalities.”<sup>[2]</sup> CPGs provide clinicians with specific recommendations based on the best available evidence. Several recent studies have found that CPGs have fairly variable methodological quality; there are several poor quality CPGs that are not based on the best available evidence. Therefore, the quality assessment of

Editor: Jorddy Neves Cruz.

J-HK and B-KS contributed equally to this work as primary authors.

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HF20C0013).

The authors have no conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are publicly available.

<sup>a</sup> Department of Acupuncture and Moxibustion, Kyung Hee University Hospital, Gangdong 892, Dongnam-ro, Gangdong-gu, <sup>b</sup> Department of Acupuncture and Moxibustion Medicine, College of Korean Medicine, Kyung Hee University, 26, Kyungheedae-ro, Dongdaemun-gu, Seoul, Republic of Korea.

\* Correspondence: Yong-Hyeon Baek, 892, Dongnam-ro, Gangdong-gu, Seoul, Republic of Korea (e-mail: byhacu@khu.ac.kr).

Copyright © 2021 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Kim JH, Seo BK, Baek YH. Quality assessment of traditional and conventional medicine clinical practice guidelines for osteoporosis. *Medicine* 2021;100:5(e24559).

Received: 28 July 2020 / Received in final form: 14 December 2020 / Accepted: 11 January 2021

<http://dx.doi.org/10.1097/MD.00000000000024559>

CPGs using systematic methods is crucial.<sup>[3]</sup> The Appraisal of Guidelines for Research and Evaluation (AGREE) is an appraisal tool that has been validated and endorsed by leading raters or compilers of international CPGs.

Osteoporosis is a systemic skeletal disease characterized by low bone mass and microarchitectural deterioration in bone tissue, which leads to enhanced bone fragility and increased fracture risk. The operational definition of osteoporosis is based on bone mineral density measurement.<sup>[5]</sup> Fragility fractures are the most debilitating consequences of osteoporosis, and they are associated with increased morbidity and mortality.<sup>[6]</sup>

The economic burden of osteoporosis in the US is approximately \$17.9 billion per annum. In the EU, it was reported that 6.6% of men and 22.1% of women older than 50 years had osteoporosis in 2010, with 3.5 million fragility fractures,<sup>[7]</sup> and this is set to rise following the increasing skew towards older adults.<sup>[8]</sup> In spite of the indication of less benefit, even in patients who sustain a fragility fracture, fewer than 20% receive treatment to cut down the risk of fracture in the year following the fracture.<sup>[9]</sup>

While standardized treatments such as bisphosphonates, raloxifene, denosumab, teriparatide, abaloparatide, and romosozumab have recently been used to manage osteoporosis,<sup>[10]</sup> traditional medical treatments are not recommended as alternatives.<sup>[11]</sup> Traditional herbal decoctions and other complementary modalities can be used for managing osteoporosis to reduce or delay the use of conventional medications.<sup>[12,13]</sup>

In a limited way, traditional medicine refers to Traditional Chinese Medicine in the present study. Other than conventional medicine, Traditional Chinese medicine has discrete advantages that include the following:

1. it is based on individual symptoms;
2. it has low resistance;
3. it is administered in various forms;
4. its use is characterized by convenience, portability, and acceptability;
5. it is safe, reliable, and effective, and it has low toxicity and adverse effects;
6. it is affordable.<sup>[14]</sup>

Moreover, traditional Chinese medical doctors have rich experience in treating osteoporosis.

Nevertheless, guidelines that have been developed that reflect the therapeutic point of traditional medicine are scarce. It would be clinically significant if researchers could find deficiencies in existing guidelines and utilize them as clues in the process of developing CPGs in the future. The objective of this systematic review was

1. to examine the explicit instructions for the conventional or traditional medicine modalities in the CPGs found and
2. assess the methodological quality of CPGs for patients with osteoporosis.

## 2. Methods

### 2.1. Retrieval of clinical guidelines

Authors searched electronic databases and websites for osteoporosis CPGs; there were no restrictions on the dates of publication of CPGs. Since CPGs are updated as necessary, only the latest and oldest versions were screened.

Medical subject headings and text related to osteoporosis and guidelines were used to search MEDLINE, EMBASE, and the Cochrane Library using the OVID interface. Websites and Google Scholar were also searched to complement the electronic database search. The sites that were searched include Pubmed (<https://pubmed.ncbi.nlm.nih.gov/>), National Institute for Health and Clinical Excellence (NICE; [www.nice.org.uk](http://www.nice.org.uk)), Guidelines International Network (G-I-N; [www.g-i-n.net](http://www.g-i-n.net)), National Guidelines Clearinghouse (NGC; [www.guideline.gov](http://www.guideline.gov)), and Scottish Intercollegiate Guidelines Network (SIGN; [www.sign.co.uk](http://www.sign.co.uk)). CPGs published in the Japanese database were retrieved via Citation Information by NII (CiNii; <https://ci.nii.ac.jp/>). Chinese language retrieval was processed via the China National Knowledge Infrastructure (CNKI). Additional retrieval involved a computerized search of three major academic databases in Korea, which are Korean Medical Guideline Information Center (KoMGI; <https://www.guideline.or.kr/>), Korea Education and Research Information Service (RISS; <http://www.riss.kr/index.do#thesis>), and National Clearing House for Korean Medicine (NCKM; <http://www.nckm.or.kr/main/index.do>). Dissertations, letters, gray literature, government documents, study reports, papers from conferences, and abstracts were manually reviewed to avoid publication bias. Personal contacts were made with former authors of the retrieved literature to resolve missing data from the publications.

The search keywords for CPGs were (CPGs and osteoporosis) for each of the databases mentioned above. These search terms were combined as follows: CPGs OR guideline OR specification AND osteoporosis. This search strategy was adjusted for each database. The references retrieved were for guidelines that met the definition proposed by the Institute of Medicine. Furthermore, only CPGs written in English, Chinese, Japanese, and Korean were included.

### 2.2. Selection of guidelines

All CPGs of traditional and conventional medicine supported by an official government and global organizations and associations related to the assessment and diagnosis, management, and clinical therapeutic and pharmacological recommendations were eligible for inclusion in this study. The excluded documents included guidelines without recommendations, secondary publications derived from CPGs, consensus statements, or consensus conferences based on the opinion of panelists, systematic reviews, editorials, clinical trials, and single-author documents. Referring to the above-mentioned definition of evidence-based CPG, CPGs that were based on expert opinions, and not literature, were excluded. If the full guideline was not available in the public domain, we bought a copy. Two reviewers independently screened and assessed paper titles and abstracts for potential eligibility using a predefined relevance criteria form. Full-text articles were obtained for potentially relevant CPGs, and these were subsequently screened by 2 independent reviewers. Disagreements at any stage were resolved by discussion or the involvement of a third reviewer.

### 2.3. Data extraction and quality assessment

A draft data extraction form was piloted and modified. Two reviewers independently extracted all of the data using the standardized data extraction form. Disagreements were resolved by discussion or the involvement of a third reviewer. All the

relevant documents and websites of the selected CPGs were examined. The extracted data included CPG characteristics (eg, country, organization, year of publication, number of authors, number of references, target population, and treatments), recommendations related to the pharmacological, non-pharmacological, and preventive modalities.

Three researchers (co-authors) who were experienced in the quality assessment of CPGs independently scored each guideline using the AGREE II instrument. All of them were Korean medicine clinicians and professors of colleges. Before evaluating the CPGs, the assessors received the same training using the AGREE II online tutorial and reviewed the CPGs on osteoporosis using the AGREE II to familiarize them with the instrument. Each appraiser independently scored the CPGs using the AGREE II items and assigned an overall quality rating using the same 7-point scale. The kappa score was calculated for consistency verification.

The AGREE II Instrument, which was utilized for the assessment of guidelines, is an updated version of the original AGREE Instrument developed in 2003 by the AGREE Collaboration.<sup>[15]</sup> AGREE II consists of 23 items grouped into 6 domains:

1. scope and purpose – overall aim of the guideline, specific health problems, and target group;
2. stakeholder involvement – extent to which appropriate stakeholders were involved in developing the guideline, which represents the views of its intended users;
3. rigor of development – the process of gathering and summarizing the evidence, methods used to develop and update recommendations;
4. clarity of presentation – language, structure, and format of guideline;
5. applicability – potential barriers and facilitators to implementation, strategies to improve uptake, and resources needed to implement the guideline;
6. editorial independence – biases due to competing interests.

The overall assessment included rating the overall quality of the guideline and determining whether the guideline would be recommendable to practitioners. The assessors compared their scores for each item and came to a consensus on discrepant scores (defined as scores varying by 3 points or more on the seven-point AGREE II scale). This approach accounted for frank errors on the part of an assessor when they had missed the relevant part of the guideline in their original assessment. Inter-rater reliability was examined by comparing the individual item scores of each assessor and ensuring that there was only a low discrepancy (less than 1.5 standard deviations from the mean domain score) with the concordance calculator.

Standardized domain scores (expressed on a scale of 0–100) were calculated using the approach of AGREE II ([obtained score – minimum possible score] divided by [maximum possible score – minimum possible score]). After the AGREE II evaluation, the final decision of “recommend,” “recommend with modification,” or “do not recommend” was independently made by each assessor, and a consensus was reached for each guideline. The decisions for the guidelines were divided into 3 levels according to the score of each field and the final judgment of the evaluators: Class A (recommended): the scores of all 6 fields of the guideline were  $\geq 60\%$ , and the guideline could be recommended without modification; Class B (recommended after modification of different degrees): the number of domains with scores of  $\geq 30\%$  was  $\geq 3$ , but there were  $< 60\%$  fields that needed

modification of varying degrees; Class C (not recommended): the number of fields with scores of less than 30% was  $\geq 3$ , and the guideline could not be recommended due to the inferior methodology of development or the poor quality of evidence.

## 2.4. Summary of CPGs recommendations

The recommendations provided by each CPG, stratified by the subjects and treatment modalities, were presented; treatment modalities recommended by the CPGs were summarized in comparative tables to highlight possible gaps. The contents of CPGs were summarized for reference and to facilitate the comparison of conventional and traditional medical CPGs.

## 2.5. Statistical analysis

The authors used the total score provided by each reviewer and the score per domain to obtain descriptive statistics. An average score of 60% was chosen (standardization of total points as a percentage over maximum points), and the proportion of CPGs that had scores above it in each domain was established. The data were obtained by analyzing the guidelines after applying the AGREE II instrument using Excel version 2010 (Microsoft Corporation, Redmond, WA). The institutional review board approval was not necessary because there was no direct involvement with patients or body samples.

## 3. Results

### 3.1. Study selection

Of the 221 papers retrieved from the search, 79 were selected for full-text screening and 15 were included in the final assessment. The flow diagram (Fig. 1) illustrates the main reasons for the exclusion of certain documents.

### 3.2. Clinical practice guideline components

The general characteristics of the selected guidelines are presented in Table 1. A total of 15 CPGs published between 2005 and 2019 were included; only 2 were published in 2005 and 2008 and the remaining 13 were published in 2010 or later. Of these, 46.7% ( $n=7$ ) were developed by Asian organizations, followed by the US ( $n=3$ ), Italy ( $n=2$ ), Republic of Korea ( $n=2$ ), and other countries (Canada, Spain, and France). Six (40.0%) guidelines were developed by medical societies, followed by government agencies ( $n=5$ , 33.3%) and colleges ( $n=4$ , 26.7%). Their content subjects included diagnosis, treatment, prevention, management, conventional modalities, and traditional Chinese medicine. Only one of the CPGs did not report the number of authors, and 13 (86.7%) of them had more than 10 authors. Additionally, 14 (93.3%) of them cited references (range, 51–216; mean, 116.4).

Most of the target populations covered by the selected CPGs were menopausal women and older adult women with primary osteoporosis; some CPGs focused on other specific gender populations such as men with osteoporosis. The subject matter, which included diagnosis, prevention, treatment, and management, were mostly similar, and only a few CPGs covered risk factors and adverse effects.

Treatment was classified into pharmaceutical and nonpharmaceutical. Only one CPG covered the treatment of traditional medicine and Western medicine concomitantly. Some of the

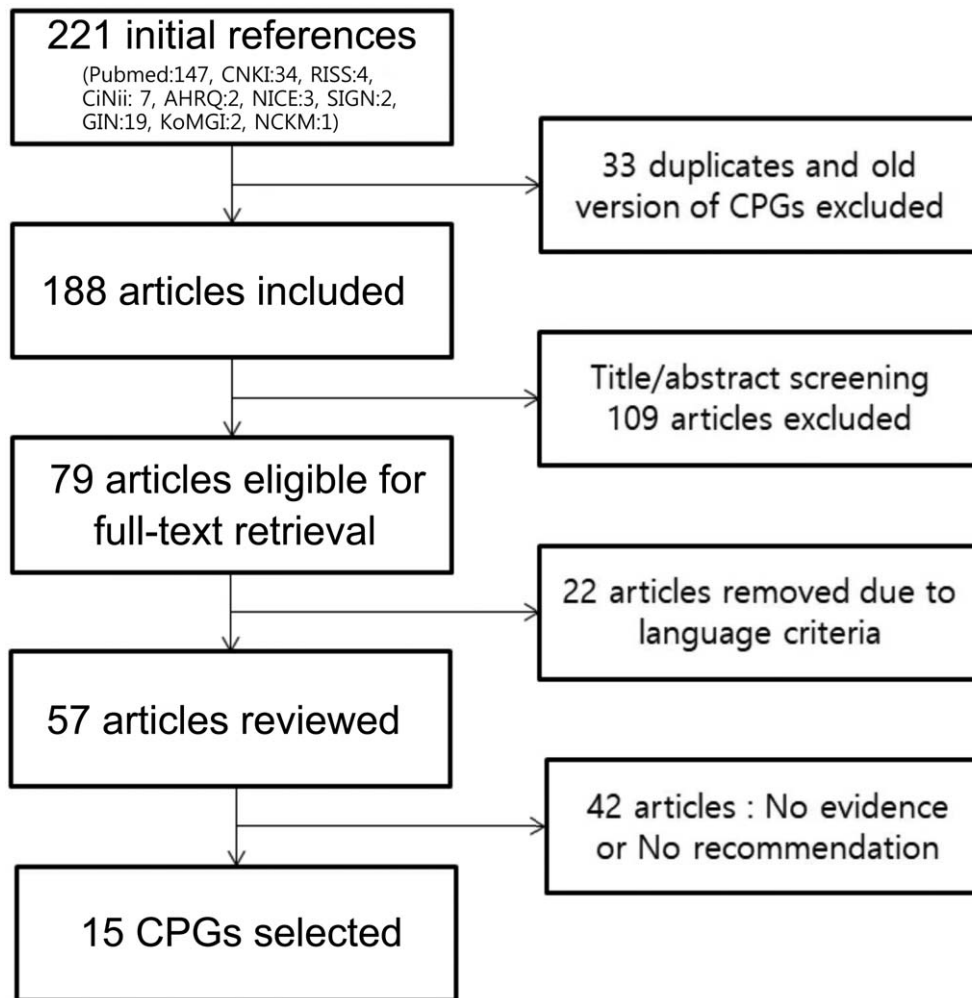


Figure 1. Procedure for CPG selection.

CPGs highlighted only exercise and rehabilitation as treatment modalities.

### 3.3. AGREE II appraisal results

A total of 15 CPGs of osteoporosis were examined using the AGREE II; three reviewers were assigned to each guideline. Table 2 presents the results of the assessments for each of the guidelines using the AGREE II tool. Kappa score among researchers was 0.61.

In general, the guidelines received the lowest scores for editorial independence among all 6 domains (mean 30.2, range 0%–75%), whereas, they scored highest on the clarity of presentation (mean 69.0%, range 46%–83%). One CPG could have low scores in certain domains and high scores in others, and this influenced reviewers' recommendations.

The frequencies of the final reviewer recommendations after the overall evaluation were as follows: class A=1; class B=13, class C=1. Based on the analysis, AACE 2016<sup>[24]</sup> was the CPG with the highest grading. This guideline scored the highest in the following 4 domains; scope and purpose, stakeholders involvement, clarity of presentation, and applicability.

The latter guideline was not recommended because of a lack of rigor in development and editorial independence although it scored highly for “clarity of presentation.” All the other guidelines were recommended even though some needed revisions in one or more domains. Most of the guidelines provided easily relatable recommendations and various options for managing osteoporosis.

The assessment results stratified by the AGREE II domains are shown in Table 2.

**3.3.1. Purview and purpose.** The score of this domain demonstrates the general aim of the clinical guidelines, the clinical issues covered, and the applicable population. The scores for this domain were moderate, with a mean of 50.1% and range of 20 to 94. KSBMR 2018<sup>[29]</sup> scored the highest, at 94%. Two scores were above 80%, 9 were below 60%, and the lowest score by PRMA 2005<sup>[16]</sup> was 20%. PRMA 2005 scored low because it provided a generalized coverage of the tackled health issues, and there was no clear indication of the population to which the guidelines would apply. The common problems with the other guidelines were that they described specific clinical or health issues, and they did not report the expected benefits.

**Table 1**

**General information on the included guidelines.**

Study ID	Country	Organization	Year of publication	Number of authors	Number of references	Target population	Subject	Treatment
PRMA 2005 <sup>[16]</sup>	Italy	Physical and Rehabilitation Medicine Association	2005	8	122	1. Postmenopausal 2. Senile	1. Prevention 2. Rehabilitations 3. Management 4. Evaluation	1. Exercises 2. Physical therapy*(TENS, Bracing treatment)
ACP 2008 <sup>[17]</sup>	USA	American College of Physicians	2008	11	49	Men	1. Diagnosis 2. Risk factors 3. Prevention	I/A
SACOC 2010 <sup>[18]</sup>	Canada	Scientific Advisory Council of Osteoporosis Canada	2010	13	71	Women and men over age 50	1. Diagnosis 2. Management 3. Adverse effects	1. Exercises 2. Calcium and vitamin D 3. Pharmacologic therapy 1. Herbal medicine
CATCM 2011 <sup>[19]</sup>	China	Chinese Academy of Traditional Chinese Medicine	2011	19	108	1. Postmenopausal 2. Senile	1. Diagnosis 2. Treatment 3. Management 4. Prevention	1. Herbal medicine 2. Acupuncture 3. Moxibustion 4. Conventional treatment
CACMS 2012 <sup>[20]</sup>	China	China Academy of Chinese Medical Services	2012	12	Not reported	Adults	1. Diagnosis	1. Herbal medicine
SMS 2013 <sup>[21]</sup>	Spain	Spanish Menopause Society	2013	16	51	Postmenopausal	2. Management 1. Pharmacological treatment 2. Prevention	1. Pharmacological treatment 1. Calcium and vitamin D 2. Pharmacological treatment
FSR 2014 <sup>[22]</sup>	France	French Society of Rheumatology	2014	15	68	Glucocorticoid-induced	1. Pathophysiology 2. Management (follow-up) 3. Treatment	1. Calcium and vitamin D 2. Pharmacological treatment 1. Calcium and vitamin D
SOS 2015 <sup>[23]</sup>	Saudi Arabia	Saudi Osteoporosis Society	2015	14	87	Women over age 60	1. Conventional treatment 2. Prevention	1. Calcium and vitamin D 2. Pharmacological treatment 1. Calcium and vitamin D
AACE 2016 <sup>[24]</sup>	US	American Association of Clinical Endocrinologists	2016	11	209	Postmenopausal	1. Diagnosis 2. Risk factors 3. Conventional treatment	1. Calcium and vitamin D 2. Pharmacological treatment 1. Calcium and vitamin D
ACR 2017 <sup>[25]</sup>	US	American College of Rheumatology	2017	21	57	Children and adults	1. Conventional treatment 2. Prevention	1. Calcium and vitamin D 2. Pharmacological treatment 1. Exercises
ISOT 2017 <sup>[26]</sup>	Italy	Italian Society for Orthopaedics and Traumatology	2017	20	185	Children and adults	1. Conventional treatment 2. Prevention	1. Exercises 2. Nutrition 3. Pharmacological treatment 1. Calcium and vitamin D
CATCM 2017 <sup>[27]</sup>	China	Chinese Academy of Traditional Chinese Medicine	2017	Not reported	154	Adults and adolescents	1. Diagnosis 2. Pathology 3. Treatment	1. Calcium and vitamin D 2. Pharmacological treatment 3. Herbal medicine 1. Calcium
CSGGM 2018 <sup>[28]</sup>	China	Chinese Society of Gerontology and Geriatric Medicine Branch	2018	14	178	Women and men over age 60	1. Diagnosis 2. Treatment	2. Pharmacological treatment 1. Calcium and vitamin D
KSBMR 2018 <sup>[29]</sup>	Republic of Korea	Korean Society of Bone and Mineral Research	2018	11	75	Women and men over age 19	1. Risk of fracture assessment 2. Conventional treatment 1. Diagnosis	2. Pharmacological treatment 1. Calcium and vitamin D
KSBMR 2019 <sup>[30]</sup>	Republic of Korea	Korean Society of Bone and Mineral Research	2019	41	216	Adults	2. Prevention 3. Pharmacological treatment	2. Nutrition and diet 3. Rehabilitation 4. Pharmacological treatment

\*TENS = Transcutaneous Electrical Nerve Stimulation.

**Table 2**  
**AGREE II domain-standardized scores for selected CPGs and average overall quality score and overall assessment of the reviewers.**

Study ID	Scope and purpose (%)	Stakeholders involvement (%)	Rigor of development (%)	Clarity of presentation (%)	Applicability (%)	Editorial independence (%)	Overall assessment
PRMA 2005 <sup>[16]</sup>	20	46	51	70	47	31	B
ACP 2008 <sup>[17]</sup>	70	46	53	61	43	42	B
SACOC 2010 <sup>[18]</sup>	54	41	57	76	63	47	B
CATCM 2011 <sup>[19]</sup>	56	48	63	48	21	0	B
CACMS 2012 <sup>[20]</sup>	33	31	45	46	32	42	B
SMS 2013 <sup>[21]</sup>	46	30	39	56	31	47	B
FSR 2014 <sup>[22]</sup>	59	30	58	76	39	39	B
SOS 2015 <sup>[23]</sup>	59	74	53	89	67	11	B
AACE 2016 <sup>[24]</sup>	85	81	60	83	69	72	A
ACR 2017 <sup>[25]</sup>	65	56	42	74	39	0	B
ISOT 2017 <sup>[26]</sup>	46	43	51	83	61	75	B
CATCM 2017 <sup>[27]</sup>	28	35	44	65	39	8	B
CSGGM 2018 <sup>[28]</sup>	54	24	44	69	54	39	B
KSBMR 2018 <sup>[29]</sup>	94	37	66	61	50	0	B
KSBMR 2019 <sup>[30]</sup>	28	39	12	78	58	0	C

**3.3.2. Stakeholder involvement.** This domain assessed whether the CPGs were suitable for professional participants. It examined the degree to which guidelines represented the perspectives of their designated users and whether the target groups were clearly defined and their opinions and selections were collected. The guideline development team should include all relevant professionals, patient groups, and the delegates from the general public. The guidelines published by AACE 2016<sup>[24]</sup> and SOS 2015<sup>[23]</sup> scored well; they provided basic demographic data and the roles of the development team members. The guidelines by ACR 2017<sup>[25]</sup> and CATCM 2011<sup>[19]</sup> did not state the geographical locations and professions. CACMS 2012,<sup>[20]</sup> SMS 2013,<sup>[21]</sup> FSR 2014,<sup>[22]</sup> and CSGGM 2018<sup>[28]</sup> scored poorly; the information they provided was not detailed enough. SACOC 2010<sup>[18]</sup> and ISOT 2017<sup>[26]</sup> indicated that the development team included epidemiologists and methodologists, while PRMA 2005<sup>[16]</sup> and ACP 2008<sup>[17]</sup> included only epidemiologists. The views of the target population should be considered during the development of the guidelines. Most of the guidelines scored poorly in this area, and only the target populations acknowledged in SOS 2015,<sup>[23]</sup> AACE 2016,<sup>[24]</sup> and ACR 2017<sup>[25]</sup> were preferred. With the exception of CATCM 2017,<sup>[27]</sup> CSGGM 2018,<sup>[28]</sup> and KSBMR 2019,<sup>[30]</sup> most (n=12, 80%) guidelines defined the target population and this subsection scored better.

**3.3.3. Rigor of development.** This domain, which was regarded as the most crucial, assessed the rigor of the process of development of the guidelines, which is mainly related to the method and standard of evidence retrieval, the tradeoffs of evidence, the process of forming recommendations, and the correlation with existing evidence, external review, and the update of guidelines.

The average score for this domain was 49.2%, with a range of 12 to 63. The recommendations and evidence for the evidence-based guidelines were verifiable and explicit. Six of the 15 evidence-based guidelines used the GRADE evaluation, making the guidelines more rigorous. Three (50%) of them scored above 60% for this domain.

Since one of the ultimate objectives of this paper is to cover the more methodologically complete items to be included in the CPG

in the future, the comparison of scores of the subitems in this section is considered meaningful (Refer to Table 3). AACE 2016<sup>[24]</sup> details the superiority and drawbacks of the evidence in the annex; other guidelines did not provide such elaboration. On the application of systematic methodology to the search for evidence, CATCM 2011,<sup>[19]</sup> FSR 2014,<sup>[22]</sup> AACE 2016,<sup>[24]</sup> and KSBMR 2018<sup>[29]</sup> had better scores; they detailed their strategies for searching evidence, such as retrieval from databases and the use of search terms and periods. Not only did they deliver recommendations from expert panelists, but CATCM 2011<sup>[19]</sup> and FSR 2014<sup>[22]</sup> went through external review before publication, while other guidelines were not cited. KSBMR 2018<sup>[29]</sup> detailed the health benefits, side effects, and risks of recommendations, and it had the highest score. It also detailed the procedures for updates although it did not indicate the schedule. All the remaining guidelines did not indicate the renewal procedure or schedule, and the average score for this item was the lowest among the subsections for this domain.

**3.3.4. Clarity of presentation.** This domain assesses the clarity of the CPGs. Reviewers determined if

1. the recommendations were comprehensive and concrete,
2. different modalities or health issues were listed,
3. significant recommendations were easy to pinpoint,
4. and the recommendations could help users elucidate problems better.

This domain had the highest score with an average of 69.0% and a range of 46 to 83. All guidelines were clearly presented, and the recommendations were easy to comprehend. CATCM 2011<sup>[19]</sup> and CACMS 2012,<sup>[20]</sup> which mainly covered traditional medicine, scored lower for the descriptions of the modalities of traditional medicine than the CPGs that covered conventional medicine due to the ambiguity of expression.

**3.3.5. Applicability.** This domain primarily covered the developers and the obstacles they encountered during the development of the guideline; the areas assessed included the provision of the application methods or recommendations, considerations of possible resources required during implementation, and the provisions for monitoring standards.

**Table 3****Rigor of development: standardized average scores in AGREE II domain for selected CPGs.**

Title ID	Systematic methods used	Selecting criteria presented	Formulating the recommendations described	Benefits, side effects, and risks considered	Explicit link between the recommendations and evidence	Reviewed by external experts	Updating procedure provided
PRMA 2005 <sup>[16]</sup>	67	78	100	100	50	17	0
ACP 2008 <sup>[17]</sup>	83	100	83	67	61	0	33
SACOC 2010 <sup>[18]</sup>	78	100	83	67	72	33	22
CATCM 2011 <sup>[19]</sup>	72	78	100	78	50	39	89
CACMS 2012 <sup>[20]</sup>	50	33	83	83	28	67	17
SMS 2013 <sup>[21]</sup>	83	72	94	0	17	33	11
FSR 2014 <sup>[22]</sup>	61	61	83	100	50	39	67
SOS 2015 <sup>[23]</sup>	67	56	83	67	50	50	50
AACE 2016 <sup>[24]</sup>	89	67	100	67	83	44	33
ACR 2017 <sup>[25]</sup>	67	44	50	44	78	39	17
ISOT 2017 <sup>[26]</sup>	61	78	83	67	83	39	0
CATCM 2017 <sup>[27]</sup>	50	67	44	50	50	83	6
CSGGM 2018 <sup>[28]</sup>	61	50	83	67	67	22	0
KSBMR 2018 <sup>[29]</sup>	83	67	100	67	83	28	100
KSBMR 2019 <sup>[30]</sup>	0	17	0	0	61	17	0

The average score for this domain was 47.5%, and the highest score by AACE 2016<sup>[24]</sup> was 69%. The annexes listed the promoting and hindering factors, the recommendations for priority recommendation, and related fees and resources. Most guidelines, including ISOT 2017,<sup>[26]</sup> provided tools for practice and flow charts to guide decision making. SACOC 2010,<sup>[18]</sup> AACE 2016,<sup>[24]</sup> and ISOT 2017<sup>[26]</sup> enhanced their applicability by clarifying the cost issues that may have arose when applying the recommendations. Only SOS 2015<sup>[23]</sup> presented key criteria for monitoring and evaluating the extent to which the recommendations could be implemented.

**3.3.6. Editorial independence.** For this domain, guidelines were expected to indicate factors that influenced the development of content, such as the perspectives of sponsors, and any conflicts of interest of the development team members.

The scores for this domain were fickle. ISOT 2017<sup>[26]</sup> had the highest score of 75%, followed by AACE 2016<sup>[24]</sup> with a score of 72%. CATCM 2011,<sup>[19]</sup> ACR 2017,<sup>[25]</sup> KSBMR 2018,<sup>[29]</sup> and KSBMR 2019<sup>[30]</sup> had no information for this section. SOS 2015<sup>[23]</sup> and CATCM 2017<sup>[27]</sup> had statements on editorial independence and conflicts of interest, but they were not detailed enough.

**3.3.7. Overall evaluation.** The 15 CPGs were of variable quality: if 6 areas scored  $\geq 60\%$ , the guideline was classified as A (can be recommended directly without change); if a guideline scored between 30% and 60% for 3 or more areas, it was classified as B (recommended after different degrees of modification and improvement), and modifications of varying degrees were requested; if the guideline had scores of  $\leq 30\%$  in 3 or more fields, it was classified as C (not recommended) because the development method of the guidelines or the quality of the evidence was defective. Of the 15 CPGs, 1 was A, 13 were B, and 1 was C. The domain of AGREE II with the highest score was clear-cut (69.0%), followed by applicability (55.1%), the rigor of development (54.1%), scope and purpose (53.1%), stakeholder involvement (44.1%), and editorial independence (30.2%), which was lowest.

## 4. Discussion

In this section, researchers firstly have glanced the overall methodological quality of retrieved guidelines. To obtain clues in developing guidelines for traditional medicine for the treatment of osteoporosis in the future, researchers analyzed the distinctions of both kinds of medicine and the recommendations for osteoporosis. Summed up these insights, researches summarized the future recommendations for the following development of CPGs.

### 4.1. Evaluation of the quality of osteoporosis guidelines

CPG encourages physicians to assess treatment outcomes more often and revise treatments accordingly in managing osteoporosis. Additionally, guideline application has potential to improve patients' treatment adherence, as patient buy-in and shared decision-making have been identified as essential components of an osteoporosis treatment approach.<sup>[31]</sup> This study is the first attempt at using the AGREE II system to assess the methodological quality of CPGs on osteoporosis, its pharmacological and nonpharmacological management approaches, prevention, and traditional Chinese medicine interventions. Only 15 guidelines met the inclusion criteria and were included. Considering the high prevalence of osteoporosis, the number of guidelines that were eventually included in this study was less than expected. Specifically, all the guidelines covered the pharmacological or nonpharmacological treatment of osteoporosis, while 8 covered the prevention. Various age and gender distributions of the target group were covered by the guidelines. Three guidelines targeted only women: SMS 2013,<sup>[21]</sup> SOS 2015,<sup>[23]</sup> and AACE 2016.<sup>[24]</sup> ACP 2008<sup>[17]</sup> targeted men. Among the 15 selected CPGs, 3 included recommendations for children and adolescents: ACR 2017,<sup>[25]</sup> ISOT 2017,<sup>[26]</sup> and CATCM 2017.<sup>[27]</sup>

The overall quality varied significantly across the 15 guidelines, and no guideline had a high score for all the domains. Most of the 15 guidelines (13 of 15) needed modifications, and only 1 was recommended for use in clinical practice. However, the function of the guideline leans on its specific quality. Only strictly designed

**Table 4**  
Average scores of selected CPGs for AGREE II domains and comparison of the average scores of traditional and western CPGs.

	Scope and purpose (%)	Stakeholders involvement (%)	Rigor of development (%)	Clarity of presentation (%)	Applicability (%)	Editorial independence (%)
Average scores of CPGs	53	44	49	69	48	30
Average scores of traditional CPGs	45	40	54	47	27	21
Average scores of Western CPGs	57	46	49	73	52	34

guidelines can guide decision-makers in healthcare domains, and the quality of guidelines for osteoporosis need to be improved.

Armstrong<sup>[32]</sup> assessed osteoporosis CPGs using AGREE II tools. Only 19 published CPGs were included in his study; however, CPGs that were not based on evidence were also included in the final retrieval results, and it was considered that stricter selection criteria should have been applied. Ten of the assessed CPGs were to be recommended with revisions, and only 4 were recommended for use in clinical practice. Since the existing literature only focused on CPGs that focused on physical activity and safe movement, our study is the first to involve CPGs covering the overall management of osteoporosis. This study covered various aspects, including the latest evidence, the target population, and the quality of coverage of traditional and conventional medicine approaches to osteoporosis. In addition, the modalities utilized in traditional medicine was summarized so that it could provide a methodological standards for future guideline development.

#### 4.2. Appraisal of traditional and conventional CPGs in the methodological perspective

Table 4 shows the average scores of the CPGs for the AGREE II domains and the comparison between average scores of traditional medicine and conventional medicine guidelines. The 5 domain scores of the conventional medicine guidelines were higher than the

traditional guidelines. The lowest score was for editorial independence, and this needs to be considered for improvements in quality for the future development of traditional medicine CPGs on osteoporosis. Items that were lacking for the traditional CPGs, compared to conventional medicine CPGs, are clarity of presentation and applicability, which need to be supplemented. The domain of rigor of development, which recorded the highest score, also seems to have received relatively high scores due to the recent development of traditional medicine CPGs, but the score was only 54%, and improvements for this domain are also required.

#### 4.3. Analysis of recommendations for osteoporosis

It was observed that the pharmacological approaches recommended by conventional guidelines were consistent in general, but the complementary or nonpharmacological managements by experts were not. For example, only 2 CPGs made clinical recommendations for rehabilitation; most CPGs made recommendations for preventing and managing fragility fractures, but they did not emphasize the need for rehabilitation. In addition, only 1 CPG delivered recommendations on physical therapy, such as transcutaneous electrical nerve stimulation and bracing treatments, which are thought to be derived considering the distribution of older adults in the patient groups and the importance of activity limitations.

Table 5 shows the contents included in the recommendations of the traditional medicine CPGs. From the details of the

**Table 5**  
Recommendations for osteoporosis in traditional CPGs.

Title ID	Diagnosis	Herbal medicine treatment	Nonherbal medicine treatment	Management
CATCM 2011 <sup>[19]</sup>	1. Diagnosis with conventional medicine	1. Prescribed herbal medicine according to pattern identification; (1) Kidney Yang Deficiency, (2) Liver-Kidney Yin Deficiency, (3) Spleen-Kidney Yang Deficiency, (4) Qi obstruction due to bloodstream malfunction	1. Acupuncture	1. Pain management
	2. Pattern identification in traditional Chinese medicine	2. Manufactured herbal medicine	2. Moxibustion 3. Tui-na manual therapy	2. Follow-up examination 3. Prevention; (1) Diet, (2) Exercises, (3) Lifetime habits, (4) Prevention for fragility fracture
CACMS 2012 <sup>[20]</sup>	Pattern identification in traditional Chinese medicine	1. Prescribed herbal medicine according to pattern identification; (1) Kidney Yang Deficiency, (2) Liver-Kidney Yin Deficiency, (3) Spleen-Kidney Yang Deficiency, (4) Qi obstruction due to bloodstream malfunction 2. Manufactured herbal medicine	1. Acupuncture 2. Moxibustion	1. Diet 2. Prevention
CATCM 2017 <sup>[27]</sup>	Not reported	Manufactured herbal medicine	Physical therapy	1. Diet 2. Prevention



recommendations, the CPGs presented on the types of diagnosis, herbal medicine treatment, nonherbal medicine treatment, and the approaches to management. The traditional medicine approach to the diagnosis of relevant symptoms differed from that of conventional medicine. These were common to the 2 selected CPGs:

1. Kidney Yang Deficiency,
2. Liver-Kidney Yin Deficiency,
3. Spleen-Kidney Yang Deficiency,
4. Qi construction due to Bloodstream were presented.

Correspondingly, the right prescription of herbal decoction was administered. In addition to herbal decoction treatment, modalities such as acupuncture, moxibustion, Tui-na manual treatment, and physical therapy were recommended. Recommendations on management covered pain management, follow-up examination, diet at a preventive level, exercise, and lifestyle management.

From a methodological perspective, recommendations on diagnosis and management presented to the CPGs need to be made to supplement the areas of Editorial Dependence, Clarity of Presentation, and Applicability. During further development of CPGs, recommendations that complement the domains of Editorial Independence, Clarity of Presentation, and Applicability should be drawn to reinforce the recommendations for diagnosis and management presented in existing CPGs. In addition, more systematic, evidence-based recommendations for the herbal and nonherbal treatments presented should be made, and they should be derived through systematic considerations of other modalities of traditional medicine not previously presented.

#### **4.4. Future recommendations for the development and assessment of clinical guidelines**

First, research, cooperation, and advocacy should be strengthened. Through analyses and summarization of strictly designed guidelines as well as increased cooperation with foreign organizations and professional bodies that develop guidelines, the quality of traditional medicine guidelines can improve, and this will provide a far-reaching and organized foundation for guideline development.

Second, the registration of the guide on the International Guide Registration Platform not only improves the transparency of the development process and averts offset and redundancy, but also it aids the distribution and application of the guidelines and strengthens collaboration among the various guideline development agencies. The writing of the plan can also ensure high-quality guidelines and the determination to, according to the purview of the plan, build the required human and financial resources, predict possible difficulty promptly, and advocate for the integrity of the final version of the guide to lessen bias.

Lastly, the quality assessment and promotion of traditional medicine guidelines need to be modified. While the common international assessment tool, AGREE II, improves the quality of guideline development, it is challenging to evaluate the development methods, the reliability and veracity of contents, the guide registration, plan writing, and the systematic evaluation. Evaluators often give variable scores due to their different professional backgrounds and level of training. It takes a relatively long time to assess each guide. In addition, AGREE II cannot be used to judge the applicability of external evidence and the specificity of recommendations related to traditional Chinese

medicine or Korean medicine. After the development of large CPGs for osteoporosis, suitable guideline evaluation tools that are suitable for traditional medicine should be considered; this will facilitate the development of high-quality guidelines.

#### **4.5. Limitations**

First, due to language capabilities and database utilization rights issues, this study only covered English, Chinese, Korean, and Japanese. The authors judged that if they included other languages and used them for the present study, they would cause biases due to language misgivings. This could result in search bias, and CPGs that are not published in English, Chinese, Korean, and Japanese, which have a strict methodological quality could be missed.

Second, non-evidence-based guidelines were excluded. Because evidence-based CPGs are developed based on stricter criteria than their non-evidence-based counterparts,<sup>[33]</sup> attention needs to be paid to interpreting the results of “rigor of development.”

Lastly, statistical disparities may also be biased when adopting the AGREE II tool for scoring. Future studies may consider more strict evaluation tools for assessing the methodological quality of CPGs.

#### **5. Conclusions**

There were significant quality disparities between the methodologies of the 15 CPGs found across the AGREE II domains; the overall quality ranged from 3.0 to 6.0 out of 7 using the AGREE II tool. One CPG was classified as A, 13 as B, and 1 as C. The contents of the guidelines were generally consistent, in comparison with the recommendations of CPGs with conventional medicine. The CPGs with traditional medicine on osteoporosis presented the use of various methods such as herbal decoction, proprietary traditional medicine, acupuncture, and moxibustion application. Comparing the quality of CPGs, the latter CPGs still need to be improved in the areas of stakeholder involvement, clarity of presentation, applicability, and editorial independence. In addition to the need for a larger number of CPGs for managing osteoporosis with traditional medicine, building a guideline assessment tool fitted for traditional medicine is required to help establish and apply rigorously developed guidelines.

#### **Author contributions**

**Conceptualization:** Jung-Hyun Kim.

**Data curation:** Jung-Hyun Kim, Byung-Kwan Seo, Yong-Hyeon Baek.

**Formal analysis:** Jung-Hyun Kim, Byung-Kwan Seo, Yong-Hyeon Baek.

**Funding acquisition:** Yong-Hyeon Baek.

**Methodology:** Byung-Kwan Seo.

**Supervision:** Yong-Hyeon Baek.

**Visualization:** Jung-Hyun Kim.

**Writing – original draft:** Jung-Hyun Kim.

**Writing – review & editing:** Jung-Hyun Kim.

#### **References**

- [1] Field MJ, Lohr KN. Committee to advise the public health service on clinical practice guidelines, Institute of Medicine. Clinical practice guidelines: directions for a new program. National Academies Press; 1990.
- [2] Institute of Medicine Committee on Standards for Developing Trustworthy Clinical Practice Guidelines, Graham R, Mancher M,

- Miller Wolman D, et al. *Clinical Practice Guidelines We Can Trust*. Washington (DC): National Academies Press (US) Copyright 2011 by the National Academy of Sciences. All rights reserved; 2011.
- [3] Armstrong JJ, Goldfarb AM, Instrum RS, et al. Improvement evident but still necessary in clinical practice guideline quality: a systematic review. *J Clin Epidemiol* 2017;81:13–21.
- [4] Yao L, Chen Y, Wang X, et al. Appraising the quality of clinical practice guidelines in traditional Chinese medicine using AGREE II instrument: a systematic review. *Int J Clin Pract* 2017;71:
- [5] Ensrud KE, Crandall CJ. Osteoporosis. *Ann Intern Med* 2017;167:17–32.
- [6] Bledsoe L, Alessi K, Toro JB, et al. Fragility fractures: diagnosis and treatment. *Am J Orthop (Belle Mead NJ)* 2018;47:
- [7] Hernlund E, Svedbom A, Ivergård M, et al. Osteoporosis in the European Union: medical management, epidemiology and economic burden. A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). *Arch Osteoporos* 2013;8:136.
- [8] Kanis JA, Odén A, McCloskey EV, et al. A systematic review of hip fracture incidence and probability of fracture worldwide. *Osteoporos Int* 2012;23:2239–56.
- [9] Giangregorio L, Papaioannou A, Cranney A, et al. Fragility fractures and the osteoporosis care gap: an international phenomenon. *Semin Arthritis Rheum* 2006;35:293–305.
- [10] Liu J, Curtis EM, Cooper C, et al. State of the art in osteoporosis risk assessment and treatment. *J Endocrinol Investig* 2019;42:1149–64.
- [11] Wang T, Liu Q, Tjhiow W, et al. Therapeutic potential and outlook of alternative medicine for osteoporosis. *Curr Drug Targets* 2017;18:1051–68.
- [12] Jin YX, Wu P, Mao YF, et al. Chinese herbal medicine for osteoporosis: a meta-analysis of randomized controlled trials. *J Clin Densitom* 2017;20:516–25.
- [13] Wang YC, Chiang JH, Hsu HC, et al. Decreased fracture incidence with traditional Chinese medicine therapy in patients with osteoporosis: a nationwide population-based cohort study. *BMC Complement Altern Med* 2019;19:42.
- [14] Wang WJ, Zhang T. Integration of traditional Chinese medicine and Western medicine in the era of precision medicine. *J Integ Med* 2017;15:1–7.
- [15] Hoffmann-Eißer W, Siering U, Neugebauer EAM, et al. Guideline appraisal with AGREE II: online survey of the potential influence of AGREE II items on overall assessment of guideline quality and recommendation for use. *BMC Health Serv Res* 2018;18:143.
- [16] Bonaiuti D, Arioli G, Diana G, et al. SIMFER Rehabilitation treatment guidelines in postmenopausal and senile osteoporosis. *Eura Medicophys* 2005;41:315–37.
- [17] Qaseem A, Snow V, Shekelle P, et al. Screening for osteoporosis in men: a clinical practice guideline from the American College of Physicians. *Ann Intern Med* 2008;148:680–4.
- [18] Papaioannou A, Morin S, Cheung AM, et al. 2010 clinical practice guidelines for the diagnosis and management of osteoporosis in Canada: summary. *CMAJ* 2010;182:1864–73.
- [19] Chinese Academy of Traditional Chinese Medicine. *Clinical practice guide for traditional Chinese medicine in Chinese medicine internal medicine*. Beijing Chin J Tradit Chin Med 2011;42:3233–7.
- [20] Xie Y-M, Yu W-Y, Dong F-H, et al. Clinical practice guideline of traditional Chinese medicine for primary osteoporosis (extract). *Chin J Integr Med*. 2012;17:52–63.
- [21] Mendoza N, Sánchez-Borrego R, Villero J, et al. 2013 Up-date of the consensus statement of the Spanish Menopause Society on postmenopausal osteoporosis. *Maturitas* 2013;76:99–107.
- [22] Briot K, Cortet B, Roux C, et al. 2014 update of recommendations on the prevention and treatment of glucocorticoid-induced osteoporosis. *Joint Bone Spine* 2014;81:493–501.
- [23] Al-Saleh Y, Sulimani R, Sabico S, et al. 2015 guidelines for osteoporosis in Saudi Arabia: recommendations from the Saudi Osteoporosis Society. *Ann Saudi Med* 2015;35:1–2.
- [24] Camacho PM, Petak SM, Binkley N, et al. American Association of Clinical Endocrinologists and American College of Endocrinology Clinical Practice Guidelines for the diagnosis and treatment of postmenopausal osteoporosis - 2016. *Endocr Pract* 2016;22(Suppl 4):1–42.
- [25] Buckley L, Guyatt G, Fink HA, et al. 2017 American College of Rheumatology Guideline for the Prevention and Treatment of Glucocorticoid-Induced Osteoporosis. *Arthritis Care Res (Hoboken)* 2017;69:1095–110.
- [26] Tarantino U, Iolascon G, Cianferotti L, et al. Clinical guidelines for the prevention and treatment of osteoporosis: summary statements and recommendations from the Italian Society for Orthopaedics and Traumatology. *J Orthop Traumatol* 2017;18(Suppl 1):3–6.
- [27] Xia Weibo, Zhang Zhen Lin, Lin Hua, et al. Guidelines to primary osteoporosis treatment. *Chin J Osteoporos* 2019;1:281–309.
- [28] 2018;Ma Yuan Zheng, Wang Yi-peng, Liu Qiang, et al. Guideline to the diagnosis and treatment of osteoporosis in China. 24:1541–65.
- [29] Park SY, Gong HS, Kim KM, et al. Korean guideline for the prevention and treatment of glucocorticoid-induced osteoporosis. *J Bone Metab* 2018;25:195–211.
- [30] Research) KKSfBaM, Committee GBC. *Physician’s Guide for Osteoporosis*. KSBMR (Korean Society for Bone and Mineral Research) 2019.
- [31] Thomas T, Casado E, Geusens P, et al. Is a treat-to-target strategy in osteoporosis applicable in clinical practice? Consensus among a panel of European experts. *Osteoporos Int* 2020;1:1–9.
- [32] Armstrong JJ, Rodrigues IB, Wasiuta T, et al. Quality assessment of osteoporosis clinical practice guidelines for physical activity and safe movement: an AGREE II appraisal. *Arch Osteoporos* 2016;11:6.
- [33] Korppi M. What are evidence-based guidelines and what are they not? *Acta Paediatr* 2016;105:11–2.