



Critical appraisal of the quality of clinical practice guidelines for idiopathic pulmonary fibrosis

Xuanlin Li^{1,2#}, Xueqing Yu^{1,2,3#}, Yang Xie^{1,2,3}, Zhenzhen Feng^{1,2}, Yanfang Ma^{4,5,6}, Yaolong Chen^{4,5,6}, Jiansheng Li^{1,2}

¹Co-construction Collaborative Innovation Center for Chinese Medicine and Respiratory Diseases by Henan & Education Ministry of P.R. China, Zhengzhou, China; ²Henan Key Laboratory of Chinese Medicine for Respiratory Disease, Henan University of Chinese Medicine, Zhengzhou, China; ³Department of Respiratory Diseases, The First Affiliated Hospital of Henan University of Chinese Medicine, Zhengzhou, China; ⁴Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China; ⁵Chinese GRADE Center, Lanzhou, China; ⁶WHO Collaborating Centre for Guideline Implementation and Knowledge Translation, Lanzhou, China

Contributions: (I) Conception and design: J Li, Y Chen; (II) Administrative support: Y Xie; (III) Provision of study materials: X Yu; (IV) Collection and assembly of data: X Li, X Yu, Y Ma; (V) Data analysis and interpretation: X Li, X Yu, Y Xie, Z Feng, Y Ma; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Jiansheng Li. Co-construction Collaborative Innovation Center for Chinese Medicine and Respiratory Diseases by Henan & Education Ministry of P.R. China, Zhengzhou 450046, China; Henan Key Laboratory of Chinese Medicine for Respiratory Disease, Henan University of Chinese Medicine, Zhengzhou 450046, China. Email: li_js8@163.com.

Background: Clinical practice guidelines (CPGs) have long served as an essential tool for clinicians to rationalize their treatment in practice. However, the quality of guidelines varies greatly. The present study aimed to analyze high-quality CPGs of idiopathic pulmonary fibrosis (IPF) and highlight the potential for further improvement.

Methods: Three guideline developers' websites, PubMed, Embase, and Web of Science, as well as a public search engine, Google Scholar, were searched to retrieve CPGs regarding the management of IPF. The methodology and reporting quality of retrieved CPGs were assessed using the validated Appraisal of Guidelines for Research and Evaluation Instrument II (AGREE II) and Reporting Items for Practice Guidelines in Healthcare (RIGHT) checklist.

Results: Twelve IPF CPGs were reviewed, among which 7 (58.3%) were considered as "recommended" and 1 (8.3%) as "recommended with modifications". Among the 6 domains of AGREE II, scope and purpose (70.99%) and clarity of presentation (68.06%) were considered to be the fields in which CPGs performed best, evidenced by the highest mean AGREE II scores. The domains in which the reviewed CPGs received the lowest mean scores were rigor of development (50.87%) and applicability (34.14%). The intraclass correlation coefficient scores were excellent in each domain. The basic information domain received the highest overall reporting rate in the 7 domains of the RIGHT checklist; the other 6 domains had a full reporting rate of <50%. Eight items had a satisfactory level of reporting, whereas 14 items had poor reporting according to the RIGHT checklist. Correlation analysis revealed a highly positive correlation between the methodology and reporting quality of CPGs for IPF ($r=0.872$).

Conclusions: The methodological quality of selected IPF CPGs fluctuated greatly, and the full reporting rate was found to be quite low in some domains. In the future, we should focus not only on improving the methodological quality in the development of guidelines, but also on the reporting quality of guidelines.

Keywords: Idiopathic pulmonary fibrosis (IPF); clinical practice guidelines (CPG); Appraisal of Guidelines for Research and Evaluation Instrument II (AGREE II); Reporting Items for Practice Guidelines in Healthcare checklist (RIGHT checklist); quality

Submitted Apr 08, 2020. Accepted for publication Sep 03, 2020.

doi: 10.21037/atm-20-3200

View this article at: <http://dx.doi.org/10.21037/atm-20-3200>

Introduction

Idiopathic pulmonary fibrosis (IPF), a chronic and progressive fibrotic lung disease, is the most common type of idiopathic interstitial pneumonia (1,2). IPF has a short median survival time, and although it was once considered rare, its incidence and mortality have risen over time (3–5). Its incidence is estimated to range between 3–9 cases/100,000 people per year in Europe and North America, and is lower in East Asia and South America (3). In 2012, the incidence of IPF in the UK was 80% higher than that in 2000 (6). In 2014 alone, approximately 28,000–65,000 people died of IPF clinical syndrome in Europe, and 13,000–17,000 people died of IPF in the USA (4). In 2017, 21/100,000 people died of IPF in the USA; the male mortality rate was approximately 1.26 times than the female mortality rate (7). Acute exacerbations and comorbidities in IPF are major causes of hospitalization, death, and financial burden (8,9). Unfortunately, due to the lack of precise diagnostic techniques and therapeutics, the diagnosis and management of IPF still constitute major challenges with high complexity for clinicians in practice. Therefore, several countries and international organizations have developed and updated the clinical practice guidelines (CPGs) of IPF to improve the efficiency of diagnosis and intervention.

Ideally, CPGs are statements that include recommendations aimed at optimizing patient care through systematic reviews of evidence and assessments of the benefits and harms of alternative care options (10). The practical values of the guidelines are mainly based on rigorous methodology, transparency of development, and quality of reporting (11,12). In general, clinicians and policy-makers now consider CPGs as an essential tool for selecting the most evidenced and cost-effective treatments for their practice (13,14). However, there is a surplus of IPF CPGs with widely variable quality; therefore, it is important to identify the IPF CPGs of high quality. To date, critical appraisal of the quality of existing CPGs for IPF has been limited. Therefore, we thoroughly assessed CPGs on the diagnosis or/and management of IPF using the Appraisal of Guidelines, Research and Evaluation II (AGREE II) and Reporting Items for Practice Guidelines in Healthcare (RIGHT) checklist, with the aim of finding high-quality CPGs for IPF and to highlight their potential for improvement (15,16).

Methods

Study design

We conducted a critical appraisal of the methodology and reporting quality of CPGs for IPF using the AGREE II and RIGHT checklists, respectively (15,16).

Retrieval of guidelines

We performed a systematic database search in PubMed, Embase, and Web of Science to identify CPGs that presented diagnosis or/and treatment of IPF from inception to July 13, 2020. The database search strategy combined the following terms: idiopathic pulmonary fibrosis, pulmonary fibros*, idiopathic, IPF, practice guideline, guidance*, recommendation*, and consensus. We adjusted the search formulae to different databases (a supplementary appendix for this can be found online). For a more comprehensive search, we also searched the websites of The National Institute for Health and Care Excellence (NICE, <https://www.nice.org.uk/>), Scottish Intercollegiate Guidelines Network (<https://www.sign.ac.uk/>), and Guidelines International Network (<https://g-i-n.net>) as supplemental sources. We also performed a search in Google Scholar to obtain CPGs possibly missed by the systematic searches. The searches were independently performed by two reviewers (XLL and XQY). Disagreements were resolved by consulting a third reviewer (YFM).

Selection of guidelines

We imported the acquired records into EndNote X8 software (Clarivate Analytics), and used its command in combination with the manual elimination of duplicates. We then screened the titles and abstracts to eliminate irrelevant records and obtained the full text of prospective content-related CPGs for IPF. All relevant CPGs were examined in accordance with the following inclusion criteria: (I) complete CPG in English version was available; (II) focus on IPF diagnosis or treatment; and (III) included the updated version for the latest CPG. The CPG exclusion criteria were: (I) secondary or multiple publications; and (II) reviews, short summaries, concluding remarks, editorials, interpretations, position papers, and other versions of CPGs still under review. For each filtered

guideline, we thoroughly searched for accompanying appendices and supplementary documents to better inform our assessments. Two reviewers (YX and ZZF) selected the guidelines independently and consulted a methodology expert (YLC) to resolve ambiguities.

Data extraction of guidelines

We designed a standard form for data extraction covering the first author, publication year, country, topic, version, developers, grading system, and number of recommendations. To distinguish between guidelines, we used the first author and publication year as the ID of the guideline. For guidelines in which the first author was not mentioned, we used the abbreviation of the development institution plus publication year as the guideline ID. One reviewer (XLL) extracted data from the CPGs, and a second reviewer (XQY) checked the data.

Methodological quality assessment of guidelines

To evaluate whether the filtered CPGs for IPF met our preset inclusion criteria, the AGREE II instrument was adopted. AGREE II comprises 23 items divided into 6 domains: scope and purpose, stakeholder involvement, rigor of development, clarity of presentation, applicability, and editorial independence. Each domain identified a unique dimension of the methodology quality of the included CPGs, and each item was assessed with a seven-point score, ranging from 1 (strongly disagree) to 7 (strongly agree). Domain scores were calculated according to the following formula: $(\text{obtained score} - \text{minimal possible score}) / (\text{maximal possible score} - \text{minimal possible score}) \times 100\%$. Each included guideline was generally given an overall quality assessment score based on the average score for all domains. Because neither the AGREE tool nor previous studies had defined a uniform criterion for overall quality, we considered domains with scores $<50\%$ to indicate lower quality, $50\text{--}70\%$ as sufficient quality, and $>70\%$ as good quality in our study (17-20). Three reviewers (XQY, XLL, and ZZF) independently performed the methodological quality appraisals based on the AGREE II (15) under the guidance of two methodological experts (YFM and YLC). Before formal assessment, a meeting was held to discuss the appraisal criteria according to the AGREE II (15).

Reporting quality assessment of guidelines

To assess whether the reporting quality of each IPF CPG met our preset inclusion criteria, the RIGHT checklist (16), consisting of 22 items divided into 7 domains, was used. Certain requirements were divided into two or three subitems according to their content. Therefore, the checklist covered 35 items in 7 domains, including basic information (6 items), background (8 items), evidence (5 items), recommendations (7 items), review and quality assurance (2 items), funding and conflicts of interest statements and management (4 items), and other information (3 items). We rated items as “reported” (relevant information was fully presented), “partially reported” (relevant information was partly presented), “unreported” (lack of all relevant information), or “not applicable” (not appropriate for evaluating specific guidelines), based on the protocol of a previous study (21). Two reviewers (XLL and ZZF) performed the reporting quality appraisals independently under the training and guidance of the RIGHT checklist initiator (YLC). We completed two rounds of pre-evaluation before formal assessment, and SPSS version 22.0 software was used to calculate the kappa value of the coefficient of internal consistency.

Statistical analysis

We used Excel 2016 software (Microsoft) to document the rates and percentages of reporting of the RIGHT items, as well as the score for the domains of the AGREE II. The means and ranges of the CPG scores for each AGREE II domain were calculated. Agreement among reviewers was measured by intraclass correlation coefficient (ICC) calculated by SPSS version 22.0, according to previous studies (17-20). The ICC level was classified as poor (<0.40), fair ($0.40\text{--}0.59$), good ($0.60\text{--}0.74$), or excellent ($0.75\text{--}1.00$). The relationship between the methodological quality and the reporting quality of included CPGs were analyzed by Spearman's correlation using SPSS version 22.0 software, according to previous research (22). We made a judgment of “fully reported”, “partially reported”, “unreported”, or “not applicable”, with corresponding scores of 1, 0.5, 0, and 0, respectively. The total scores that could be obtained with the RIGHT checklist and the AGREE II tool were 35 and 161, respectively. The relationship between the methodological quality and the reporting quality is presented in the scatter gram.

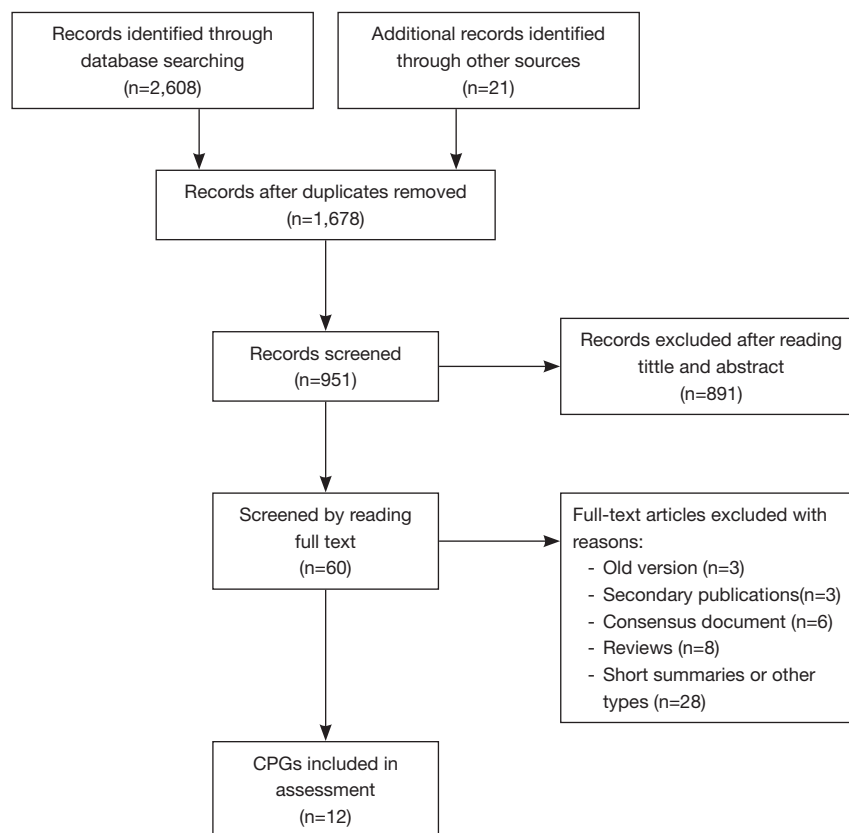


Figure 1 CPGs selection process. CPG, clinical practice guidelines.

Results

Identification of guidelines

The electronic search yielded 2,629 records. Duplicated records were removed, and 1,678 records were excluded according to their titles and abstracts. Eventually, 60 full-text articles were filtered for eligibility, and 12 CPGs (2,23-33) met our preset eligibility criteria and were therefore included in the present study (*Figure 1*).

Characteristics of guidelines

All 12 included CPGs (11 CPGs published in journals and 1 on the NICE website) were published between 2013 and 2020 (2,23-33). All were developed by local or international medical societies. Two CPGs (2,23) were globally developed, and the others came from nine different countries: France, Germany, Spain, Switzerland, Japan, Korea, England, Brazil, and Poland (24-33). Six CPGs (24,26,28,30,31,33) covered the diagnosis and treatment

of IPF, five referred to treatment (23,25,27,29,32), and one referred to diagnosis (2). Six CPGs were updated and developed from original guidelines (2,23-25,27,31), three were adapted from other guidelines (26,28,30), and the remaining three were original versions (2,32,33). The ranking systems of evidence quality also varied among these CPGs. The Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) system (34) was adopted to grade evidence quality in eight CPGs (2,23,26,27,29,31-33). The Oxford Centre for Evidence-based Medicine Levels of Evidence (35) were adopted for one of the selected guidelines (25), and the remaining three CPGs did not report the criteria they used to grade evidence (24,28,30). Recommendations for the included CPGs ranged from 7 to 66 (*Table 1*).

Result of methodology quality

Scope and purpose and clarity of presentation were the two domains where CPGs had the highest mean score, with

Table 1 The characteristics of eligible CPGs

First author	Year	Country	Topic	Version	Developers	Evidence based	Scoring System	No. of recommendations
Raghu G (23)	2015	International	Treatment	Update	ATS, ERS, JRS, and ALAT	RCTs, Observational studies, Systematic review	GRADE	12
Raghu G (2)	2018	International	Diagnosis	Update	ATS, ERS, JRS, and ALAT	RCTs, Observational studies, Systematic review	GRADE	12
Cottin V (24)	2017	French	Diagnosis and management	Update	French IPF experts	Systematic review	-	66
Behr J (25)	2018	German	Pharmacological therapies	Update	AWMF	Prospective or retrospective clinical trials, RCTs, Systematic reviews, Meta-analyses	Oxford levels of evidence	21
Xaubet A (26)	2013	Spain	Diagnosis and treatment	Adaptation	SEPAR	-	GRADE	11
Xaubet A (27)	2017	Spain	Medical treatment	Update	SEPAR	Clinical trials, Meta-analyses	GRADE	13
Funke-Chambour M (28)	2017	Switzerland	Diagnosis and treatment	Adaptation	SRS	Clinical trials, Systematic review	-	19
Homma S (29)	2018	Japan	Treatment	Original	JRS	Clinical trials from Japan and overseas	GRADE	17
Lee SH (30)	2019	Korean	Diagnosis and management	Adaptation	KATRD	Meta-analysis	-	-
NICE (31)	2017	England	Diagnosis and management	Update	NICE	Randomized studies, Observational studies, Diagnostic studies, Prognostic studies, Qualitative studies	GRADE	37
Baddini-Martinez J (32)	2020	Brazil	Medical treatment	Original	SBPT	Systematic review and Meta-analysis	GRADE	7
Piotrowski WJ (33)	2020	Poland	Diagnosis and treatment	Original	PRS	Randomized studies, Systematic review and Meta-analysis	GRADE	28

CPG, clinical practice guidelines. -, Not report; ATS, American Thoracic Society; ERS, European Respiratory Society; JRS, Japanese Respiratory Society; ALAT, Latin American Thoracic Society; RCTs, Randomized controlled trials; GRADE, Grading of Recommendations, Assessment, Development and Evaluation; AWMF, Association of the Scientific Medical Societies in Germany; SEPAR, Spanish Society of Pneumology and Thoracic Surgery; SRS, Swiss Respiratory Society; KATRD, Korean Academy of Tuberculosis and Respiratory Diseases; NICE, National Institute of Health and Care Excellence; SBPT, Brazilian Thoracic Association; PRS, Polish Respiratory Society.

Table 2 AGREE II domain and overall assessment for eligible guidelines

Guideline ID	Domain 1: scope and purpose	Domain 2: stakeholder involvement	Domain 3: rigour of development	Domain 4: clarity of presentation	Domain 5: applicability	Domain 6: editorial independence	Overall assessment
Raghu G (23)	74.07%	66.67%	60.42%	79.63%	34.72%	71.43%	R
Raghu G (2)	83.33%	72.22%	68.06%	83.33%	48.61%	77.14%	R
Cottin V (24)	48.15%	77.78%	29.17%	66.67%	40.28%	60.00%	NR
Behr J (25)	74.07%	75.93%	55.56%	68.52%	51.39%	48.57%	R
Xaubet A (26)	27.78%	5.56%	6.25%	16.67%	4.17%	25.71%	NR
Xaubet A (27)	31.48%	5.56%	7.64%	22.22%	4.17%	8.57%	NR
Funke-Chambour M (28)	85.19%	42.59%	31.94%	79.63%	27.78%	2.86%	NR
Homma S (29)	85.19%	83.33%	67.36%	81.48%	20.83%	68.57%	R
Lee SH (30)	79.63%	55.56%	57.64%	61.11%	5.56%	57.14%	RM
NICE (31)	87.04%	81.48%	75.00%	83.33%	48.61%	60.00%	R
Baddini-Martinez J (32)	88.89%	68.52%	70.14%	85.19%	61.11%	80.00%	R
Piotrowski WJ (33)	87.04%	87.04%	81.25%	88.89%	62.50%	88.57%	R
Mean, % (Range)	70.99 (27.8–88.9)	60.19 (5.6–68.5)	50.87 (6.3–70.1)	68.06 (16.7–85.2)	34.14 (4.2–61.1)	54.05 (25.7–80.0)	-
ICC (95%CI)	0.88 (0.80–0.93)	0.90 (0.84–0.95)	0.91 (0.87–0.93)	0.84 (0.74–0.91)	0.86 (0.79–0.91)	0.92 (0.84–0.96)	-

R, recommended; RM, recommended with modifications; NR, not recommended.

scores of 70.99% (range, 27.8–88.89%) and 68.06% (range, 16.7–85.2%), respectively. The mean score for clarity of presentation had the highest range. CPGs had the lowest mean score for the rigor of development and applicability domains: 50.1% and 34.1% (range, 6.3–70.1% and 4.2–61.1%, respectively). The mean score of the editorial independence domain was 54.1%, with the outlier being 2.9% (28). Overall, ICC scores were >80% in each domain, suggesting considerable agreement between the reviewers (Table 2).

Additionally, two CPGs scored <50% in each domain (26,27), one guideline scored 29.2% for the rigorous development domain (24), and one guideline in the editorial independence domain scored 2.9% (28). Therefore, these CPGs were not recommended. One guideline was recommended with modifications, as it failed to describe facilitators and barriers to its application, and presented monitoring and/or auditing criteria in the applicability domain, where it only scored 13% (30). The remaining CPGs were recommended because they had a score >50%

score in at least 5 domains (2,23,25,29,31–33). In particular, two guidelines (32,33) published in 2020 scored >50% in each domain, indicating high methodological quality (Figure 2).

Result of reporting quality

Our pre-evaluation results ($\kappa=0.831$, $P<0.001$) demonstrated good agreement between the reviewers. Of the seven domains, the basic information domain and other information domain had the highest full reporting rates; however, the full reporting rates of the remaining six domains were <50%. The evidence domain and the review and quality assurance domain had unreported rates of >25%, and the background domain had the lowest unreported rate (8.3%) of all domains. Partial reports existed in all domains, ranging from 12.5% to 34.5% (Figure 3).

For specific items, 8 items had a satisfactory level of reporting quality; $\geq 70\%$ CPGs adhered to the RIGHT checklist: 1a: identify the report as a guideline; 1c:

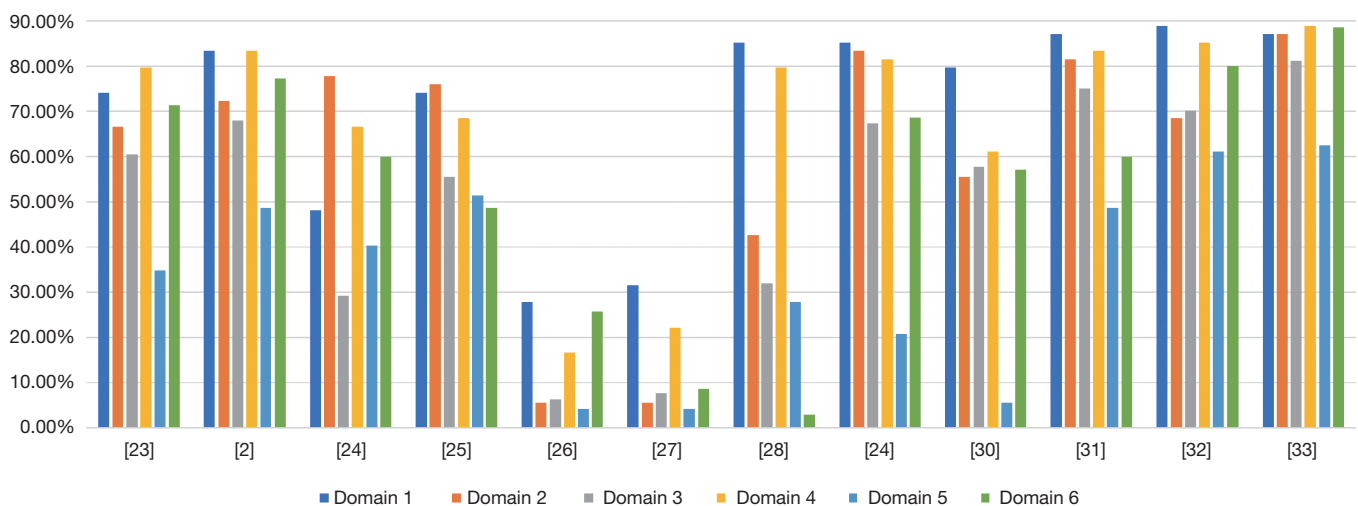


Figure 2 The scores of eligible guidelines in AGREE II domains. AGREE II, Appraisal of Guidelines for Research and Evaluation Instrument II.

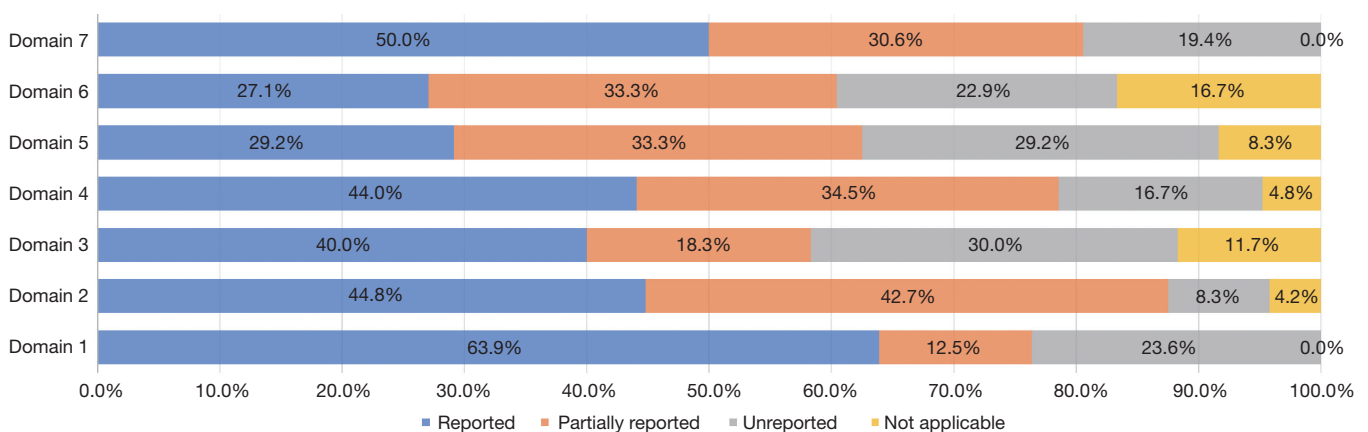


Figure 3 The scores of included guidelines in RIGHT checklist domains. RIGHT, Reporting Items for Practice Guidelines in Healthcare.

describe the focus of the guideline; 4: identify at least one corresponding author; 5: describe the basic epidemiology of the problem; 7a: describe the primary populations; 13a: provide clear recommendations; 15: describe the processes and approaches to formulate the recommendations; and 20: how is the guideline assessed, respectively. However, 14 items had poor reporting; $\leq 25\%$ CPGs adhered to the checklist: 7b: give special consideration for subgroups; 8b: describe the settings for which the guideline is intended; 9b: list all necessary information of the individuals involved in developing the guideline; 10: indicate how to select and sort the outcomes; 11b: describe how to identify and assess the system review; 14a: describe whether values and preferences of the target populations were considered; 14b: describe

whether cost and resource implications were considered; 14c: describe other factors taken into consideration; 17: indicate whether the guideline was subjected to a quality assurance process; 18a: describe the specific sources of funding for all stages of guideline development; 18b: describe the role of funders in the development of guideline; 19b: describe how conflicts of interest were evaluated and managed; and 22: describe any limitations in the guideline development process (Table 3).

Correlation of methodological and reporting quality

The Spearman correlation analysis showed that $R=0.872$, $P=0.000$ (Figure 4).

Table 3 The reporting rate in each RIGHT checklist item of eligible guidelines

No.	Item	Reported, n (%)	Partially reported, n (%)	Unreported, n (%)	Not applicable, n (%)
Basic information					
1a	Identify the report as a guideline, that is, with "guideline(s)" or "recommendation(s)" in the title.	12 (100)	0 (0)	0 (0)	0 (0)
1b	Describe the year of publication of the guideline.	4 (33.3)	0 (0)	8 (66.7)	0 (0)
1c	Describe the focus of the guideline, such as screening, diagnosis, treatment, management, prevention, or others.	12 (100)	0 (0)	0 (0)	0 (0)
2	Provide a summary of the recommendations contained in the guideline.	4 (33.3)	3 (25.0)	5 (41.7)	0 (0)
3	Define new or key terms, and provide a list of abbreviations and acronyms if applicable.	5 (41.7)	6 (50.0)	1 (8.3)	0 (0)
4	Identify at least 1 corresponding developer or author who can be contacted about the guideline.	9 (75.0)	0 (0)	3 (25.0)	0 (0)
Background					
5	Describe the basic epidemiology of the problem, such as the prevalence/incidence, morbidity, mortality, and burden (including financial) resulting from the problem.	10 (83.3)	2 (16.7)	0 (0)	0 (0)
6	Describe the aim(s) of the guideline and specific objectives, such as improvements in health indicators (e.g., mortality and disease prevalence), quality of life, or cost savings.	5 (41.7)	7 (58.3)	0 (0)	0 (0)
7a	Describe the primary population(s) that is affected by the recommendation(s) in the guideline.	9 (75.0)	3 (25.0)	0 (0)	0 (0)
7b	Describe any subgroups that are given special consideration in the guideline.	1 (8.3)	7 (58.3)	1 (8.3)	3(25.0)
8a	Describe the intended primary users of the guideline (such as primary care providers, clinical specialists, public health practitioners, program managers, and policymakers) and other potential users of the guideline.	8 (66.7)	3 (25.0)	1 (8.3)	0 (0)
8b	Describe the setting(s) for which the guideline is intended, such as primary care, low- and -middle income countries, or inpatient facilities.	1 (8.3)	6 (50.0)	4 (33.3)	1(8.3)
9a	Describe how all contributors to the guideline development were selected and their roles and responsibilities (e.g., steering group, guideline panel, external reviewers, systematic review team, and methodologists).	7 (58.3)	4 (33.3)	1 (8.3)	0 (0)
9b	List all individuals involved in developing the guideline, including their title, role(s), and institutional affiliation(s).	2 (16.7)	9 (75.0)	1 (8.3)	0 (0)

Table 3 (continued)

Table 3 (continued)

No.	Item	Reported, n (%)	Partially reported, n (%)	Unreported, n (%)	Not applicable, n (%)
Evidence					
10a	State the key questions that were the basis for the recommendations in PICO (population, intervention, comparator, and outcome) or other format as appropriate.	8 (66.7)	1 (8.3)	3 (25.0)	0 (0)
10b	Indicate how the outcomes were selected and sorted.	3 (25.0)	1 (8.3)	8 (66.7)	0 (0)
11a	Indicate whether the guideline is based on new systematic reviews done specifically for this guideline or whether existing systematic reviews were used.	5 (41.7)	5 (41.7)	1 (8.3)	1 (8.3)
11b	If the guideline developers used existing systematic reviews, reference these and describe how those reviews were identified and assessed (provide the search strategies and the selection criteria, and describe how the risk of bias was evaluated) and whether they were updated.	1 (8.3)	3 (25.0)	3 (25.0)	5 (41.7)
12	Describe the approach used to assess the certainty of the body of evidence.	7 (58.3)	1 (8.3)	3 (25.0)	1 (8.3)
Recommendations					
13a	Provide clear, precise, and actionable recommendations.	9 (75.0)	3 (25.0)	0 (0)	0 (0)
13b	Present separate recommendations for important subgroups if the evidence suggests that there are important differences in factors influencing recommendations, particularly the balance of benefits and harms across subgroups.	3 (25.0)	8 (66.7)	1 (8.3)	0 (0)
13c	Indicate the strength of recommendations and the certainty of the supporting evidence.	8 (66.7)	1 (8.3)	2 (16.7)	1 (8.3)
14a	Describe whether values and preferences of the target population(s) were considered in the formulation of each recommendation. If yes, describe the approaches and methods used to elicit or identify these values and preferences. If values and preferences were not considered, provide an explanation.	3 (25.0)	7 (58.3)	2 (16.7)	0 (0)
14b	Describe whether cost and resource implications were considered in the formulation of recommendations. If yes, describe the specific approaches and methods used (such as cost-effectiveness analysis) and summarize the results. If resource issues were not considered, provide an explanation.	3 (25.0)	3 (25.0)	6 (50.0)	0 (0)
14c	Describe other factors taken into consideration when formulating the recommendations, such as equity, feasibility, and acceptability.	2 (16.7)	7 (58.3)	3 (25.0)	0 (0)
15	Describe the processes and approaches used by the guideline development group to make decisions, particularly the formulation of recommendations (such as how consensus was defined and achieved and whether voting was used).	9 (75.0)	0 (0)	0 (0)	3 (25.0)

Table 3 (continued)

Table 3 (continued)

No.	Item	Reported, n (%)	Partially reported, n (%)	Unreported, n (%)	Not applicable, n (%)
Review and quality assurance					
16	Indicate whether the draft guideline underwent independent review and, if so, how this was executed and the comments considered and addressed.	5 (41.7)	4 (33.3)	3 (25.0)	0 (0)
17	Indicate whether the guideline was subjected to a quality assurance process. If yes, describe the process.	2 (16.7)	4 (33.3)	4 (33.3)	2 (16.7)
Funding and declaration and management of interests					
18a	Describe the specific sources of funding for all stages of guideline development.	2 (16.7)	4 (33.3)	5 (41.7)	1 (8.3)
18b	Describe the role of funder(s) in the different stages of guideline development and in the dissemination and implementation of the recommendations.	2 (16.7)	3 (25.0)	1 (8.3)	6 (50.0)
19a	Describe what types of conflicts (financial and nonfinancial) were relevant to guideline development.	7 (58.3)	4 (33.3)	1 (8.3)	0 (0)
19b	Describe how conflicts of interest were evaluated and managed and how users of the guideline can access the declarations.	2 (16.7)	5 (41.7)	4 (33.3)	1 (8.3)
Other information					
20	Describe where the guideline, its appendices, and other related documents can be accessed.	10 (83.3)	2 (16.7)	0 (0)	0 (0)
21	Describe the gaps in the evidence and/or provide suggestions for future research.	6 (50.0)	6 (50.0)	0 (0)	0 (0)
22	Describe any limitations in the guideline development process (such as the development groups were not multidisciplinary or patients' values and preferences were not sought), and indicate how these limitations might have affected the validity of the recommendations.	2 (16.7)	3 (25.0)	7 (58.3)	0 (0)

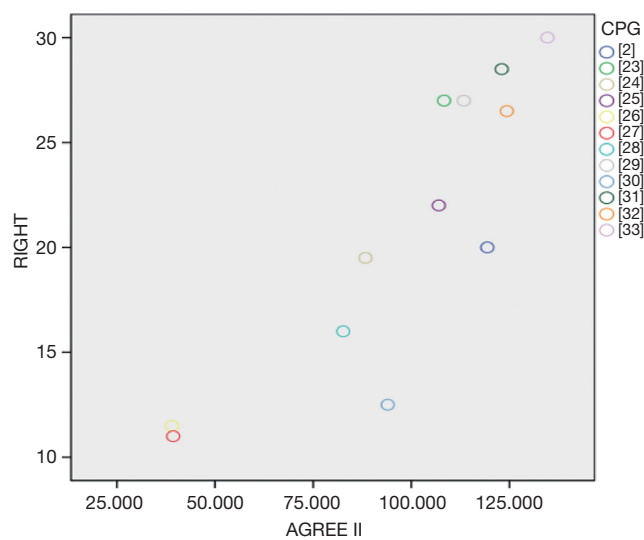


Figure 4 The correlation of methodological and reporting quality.

Discussion

In this critical appraisal of 12 CPGs for IPF using the validated AGREE II tool and RIGHT checklist, 7 CPGs were considered to be high quality and suitable for recommendation to clinical practitioners and policy-makers. The international guidelines and the guidelines issued in 2020, in particular, were of high methodological and reporting quality (2,23,32,33). Although two CPGs used the GRADE system to classify the quality of evidence, their scores in each domain of AGREE II were <50%, indicating that the methodological quality was insufficient (26,27). One guideline scored 29.2% score in the rigorous development domain, as it failed to explicitly describe the criteria for selecting evidence and making recommendations (24), and one guideline in the editorial independence domain scored only 2.9%, as it failed to explicitly describe the conflicts of interest (28); therefore, these CPGs were not recommended. We recommend that these guidelines (24,26-28) use the GRADE system for grading the quality of evidence, and use the AGREE II tool to improve the methodology quality when it is updated or revised.

The performance of the CPGs across the AGREE II domains was poor; the mean CPG scores were low, and ranges differed greatly between the scope and purpose and the clarity of presentation domains. The CPGs had mean scores of >60%, which was consistent with previous studies (17-19). In the rigor of development domain, the mean score of eligible CPGs was 50.9%, with a range of 6.3–70.1%. This domain was used to define the quality of

guidelines because it highlights the difficulties of producing evidence-based recommendations in related research, including whether the methods for guideline development are considered exhaustive enough to produce evidence-based recommendations (36,37). In our study, even recommended CPGs were likely to have lower scores ($\leq 70\%$) in the rigor of development domain; thus, developers of CPGs should pay particular attention to the methods used for developing guidelines. Similar to other studies, the applicability domain had the lowest mean score of 34.14%, with a range of 4.2–61.1% (17-19,36,38,39). Low applicability could lead to limited compliance and failure to achieve the purpose of the guidelines; therefore, formulating an appropriate implementation strategy is important (40).

The reporting quality of CPGs for IPF published in journals and on guideline developers' websites from 2013 to 2020 was evaluated using the RIGHT checklist (16). The RIGHT checklist is generally a reliable and useful tool for the assessment of reporting quality of guidelines (21,41). Among its domains, the basic information domain had the highest overall reporting rate, while the other 6 domains had rates of <50%. Eight items (22.9%) had a satisfactory level of full reporting, and 14 items (40%) had a poor level of reporting, according to the RIGHT checklist. The common deficiencies in the reporting of the included IPF guidelines were: (I) a lack of a summary of the recommendations contained in the guideline; (II) a lack of adequate descriptive methods for the selection and classification of clinical outcomes; (III) the funding for the various stages of guideline development were rarely described; and (IV) a lack of description of the limitations of the guideline development process. These deficiencies called for meticulous evaluation on the reporting quality of existing IPF guidelines and the improvement of quality for reporting in practice.

The AGREE II tool was developed for the methodological rigor and transparency of CPGs (42). Meanwhile, the aim of the RIGHT checklist is to assist guideline developers in reporting CPGs, to support journal editors and peer reviewers when considering guideline reports, and to help clinical practitioners understand and implement a guideline; the checklist differs from the AGREE II in many important aspects (16). Our research showed that IPF guidelines with better methodological quality also had greater reporting quality. Therefore, applying the AGREE II tool and RIGHT checklist together to evaluate the quality of CPGs, we were able to identify the possible gaps in the different aspects as well as areas for further improvement.

To the best of our knowledge, the present study was the first to assess the quality of CPGs for IPF by combining the AGREE II tool with the RIGHT checklist. Taken together, our study highlights the areas for further improvement in the methodology and reporting of CPGs for IPF.

Limitations

The present study has some limitations. First, we included only CPGs published in English. Although we attempted to perform systematic searches and reasonable screening of the literature, the smaller scope might still hinder the diversity of regions and settings. Second, evaluating the methodology and reporting quality of CPGs using the AGREE II and RIGHT checklists is a subjective process. We independently evaluated and used ICC scores to examine agreement; however, bias was still inevitable. Third, due to limitations of journal layout, the reporting of some CPGs might be incomplete, resulting in a possible loss of some important information and reduced reporting quality. Despite these shortcomings, our study was conducted under the guidance of a guideline methodologist, and a comprehensive search for CPGs and excellent agreement enhanced the credibility of our findings.

Conclusions

Our critical appraisal of CPGs for IPF found that more than half of the guidelines are of high quality and could be recommended; however, the methodological quality of IPF CPGs varied greatly, and the full reporting rate is low in some domains. In the future, we should focus not only on improving the methodological quality for the development of guidelines, but also the reporting quality of guidelines.

Acknowledgments

Funding: This study was supported by Key Project of the National Natural Science Foundation of China (81830116); Qihuang Scholars Award of the State TCM Academic Leader Program; Central Plains Thousand People Program (No. 194200510018, ZYQR201810159); Training Project for Top-notch Talent of Traditional Chinese Medicine in Henan Province (2019ZYBJ002).

Footnote

Conflicts of Interest: All authors have completed the ICMJE

uniform disclosure form (available at <http://dx.doi.org/10.21037/atm-20-3200>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Ethics application was not required.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Richeldi L, Collard HR, Jones MG. Idiopathic pulmonary fibrosis. *Lancet* 2017;389:1941-52.
2. Raghu G, Remy-Jardin M, Myers JL, et al. Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. *Am J Respir Crit Care Med* 2018;198:e44-68.
3. Hutchinson J, Fogarty A, Hubbard R, et al. Global incidence and mortality of idiopathic pulmonary fibrosis: a systematic review. *Eur Respir J* 2015;46:795-806.
4. Hutchinson JP, McKeever TM, Fogarty AW, et al. Increasing global mortality from idiopathic pulmonary fibrosis in the twenty-first century. *Ann Am Thorac Soc* 2014;11:1176-85.
5. Raghu G, Chen SY, Yeh WS, et al. Idiopathic pulmonary fibrosis in US Medicare beneficiaries aged 65 years and older: incidence, prevalence, and survival, 2001-11. *Lancet Respir Med* 2014;2:566-72.
6. Strongman H, Kausar I, Maher TM. Incidence, Prevalence, and Survival of Patients with Idiopathic Pulmonary Fibrosis in the UK. *Adv Ther* 2018;35:724-36.
7. Dove EP, Olson AL, Glassberg MK. Trends in Idiopathic Pulmonary Fibrosis-related Mortality in the United States: 2000-2017. *Am J Respir Crit Care Med* 2019;200:929-31.
8. Diamantopoulos A, Wright E, Vlahopoulou K, et al. The Burden of Illness of Idiopathic Pulmonary Fibrosis: A Comprehensive Evidence Review. *Pharmacoeconomics*

- 2018;36:779-807.
9. Yu YF, Wu N, Chuang CC, et al. Patterns and Economic Burden of Hospitalizations and Exacerbations Among Patients Diagnosed with Idiopathic Pulmonary Fibrosis. *J Manag Care Spec Pharm* 2016;22:414-23.
 10. Institute of Medicine (US) Committee on Standards for Developing Trustworthy. *Clinical Practice Guidelines. Clinical practice guidelines we can trust.* Washington, DC: The National Academies Press, 2011.
 11. Murad MH. *Clinical Practice Guidelines: A Primer on Development and Dissemination.* *Mayo Clin Proc* 2017;92:423-33.
 12. Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014;383:267-76.
 13. Shekelle PG. *Clinical Practice Guidelines: What's Next?* *JAMA* 2018;320:757-58.
 14. Steel N, Abdelhamid A, Stokes T, et al. A review of clinical practice guidelines found that they were often based on evidence of uncertain relevance to primary care patients. *J Clin Epidemiol* 2014;67:1251-57.
 15. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *J Clin Epidemiol* 2010;63:1308-11.
 16. Chen Y, Yang K, Marušić A, et al. A Reporting Tool for Practice Guidelines in Health Care: The RIGHT Statement. *Ann Intern Med* 2017;166:128-32.
 17. Andrade R, Pereira R, van Cingel R, et al. How should clinicians rehabilitate patients after ACL reconstruction? A systematic review of clinical practice guidelines (CPGs) with a focus on quality appraisal (AGREE II). *Br J Sports Med* 2020;54:512-9.
 18. Jiang M, Guan WJ, Fang ZF, et al. A Critical Review of the Quality of Cough Clinical Practice Guidelines. *Chest* 2016;150:777-88.
 19. Shen WQ, Yao L, Wang XQ, et al. Quality assessment of cancer cachexia clinical practice guidelines. *Cancer Treat Rev* 2018;70:9-15.
 20. Sigfrid L, Perfect C, Rojek A, et al. A systematic review of clinical guidelines on the management of acute, community-acquired CNS infections. *BMC Med* 2019;17:170.
 21. Wang Q, Duan Y, Liang J, et al. Reporting quality of 2014-2018 clinical practice guidelines on diabetes according to the RIGHT checklist. *Endocrine* 2019;65:531-41.
 22. Yun X, Yaolong C, Zhao Z, et al. Using the RIGHT statement to evaluate the reporting quality of clinical practice guidelines in traditional Chinese medicine. *PLoS One* 2018;13:e0207580.
 23. Raghu G, Rochwerg B, Zhang Y, et al. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline: Treatment of Idiopathic Pulmonary Fibrosis. An Update of the 2011 Clinical Practice Guideline. *Am J Respir Crit Care Med* 2015;192:e3-19.
 24. Cottin V, Crestani B, Cadranel J, et al. French practical guidelines for the diagnosis and management of idiopathic pulmonary fibrosis - 2017 update. Full-length version. *Rev Mal Respir* 2017;34:900-68.
 25. Behr J, Günther A, Bonella F, et al. German Guideline for Idiopathic Pulmonary Fibrosis - Update on Pharmacological Therapies 2017. *Pneumologie* 2018;72:155-68.
 26. Xaubet A, Ancochea J, Bollo E, et al. Guidelines for the diagnosis and treatment of idiopathic pulmonary fibrosis. Sociedad Española de Neumología y Cirugía Torácica (SEPAR) Research Group on Diffuse Pulmonary Diseases. *Arch Bronconeumol* 2013;49:343-53.
 27. Xaubet A, Molina-Molina M, Acosta O, et al. Guidelines for the medical treatment of idiopathic pulmonary fibrosis. *Arch Bronconeumol* 2017;53:263-69.
 28. Funke-Chambour M, Azzola A, Adler D, et al. Idiopathic Pulmonary Fibrosis in Switzerland: Diagnosis and Treatment. *Respiration* 2017;93:363-78.
 29. Homma S, Bando M, Azuma A, et al. Japanese guideline for the treatment of idiopathic pulmonary fibrosis. *Respir Investig* 2018;56:268-91.
 30. Lee SH, Yeo Y, Kim TH, et al. Korean Guidelines for Diagnosis and Management of Interstitial Lung Diseases: Part 2. Idiopathic Pulmonary Fibrosis. *Tuberc Respir Dis (Seoul)* 2019;82:102-17.
 31. The National Institute for Health and Care Excellence. Idiopathic pulmonary fibrosis in adults: diagnosis and management (Last updated May 2017). Accessed March 1, 2020. Available online: <https://www.nice.org.uk/>
 32. Baddini-Martinez J, Ferreira J, Tanni S, et al. Brazilian guidelines for the pharmacological treatment of idiopathic pulmonary fibrosis. Official document of the Brazilian Thoracic Association based on the GRADE methodology. *J Bras Pneumol* 2020;46:e20190423.
 33. Piotrowski WJ, Bestry I, Białas AJ, et al. Guidelines of the Polish Respiratory Society for diagnosis and treatment of idiopathic pulmonary fibrosis. *Adv Respir Med* 2020;88:41-93.
 34. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-26.

35. Oxford Centre for Evidence-based Medicine. Levels of Evidence (March 2009). Accessed March 1, 2020. Available online: <https://www.cebm.net/2009/06/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/>
36. Molino CGRC, Leite-Santos NC, Gabriel FC, et al. Factors Associated With High-Quality Guidelines for the Pharmacologic Management of Chronic Diseases in Primary Care: A Systematic Review. *JAMA Intern Med* 2019;179:553-60.
37. Sox HC. Conflict of Interest in Practice Guidelines Panels. *JAMA* 2017;317:1739-40.
38. Acuña-Izcaray A, Sánchez-Angarita E, Plaza V, et al. Quality assessment of asthma clinical practice guidelines: a systematic appraisal. *Chest* 2013;144:390-7.
39. 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.
40. Chan WV, Pearson TA, Bennett GC, et al. ACC/AHA Special Report: Clinical Practice Guideline Implementation Strategies: A Summary of Systematic Reviews by the NHLBI Implementation Science Work Group: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation* 2017;135:e122-37.
41. Tokalić R, Viđak M, Buljan I, et al. Reporting quality of European and Croatian health practice guidelines according to the RIGHT reporting checklist. *Implement Sci* 2018;13:135.
42. Burls A. AGREE II-improving the quality of clinical care. *Lancet* 2010;376:1128-29.

Cite this article as: Li X, Yu X, Xie Y, Feng Z, Ma Y, Chen Y, Li J. Critical appraisal of the quality of clinical practice guidelines for idiopathic pulmonary fibrosis. *Ann Transl Med* 2020;8(21):1405. doi: 10.21037/atm-20-3200