Original Article

The Chinese Thoracic Oncology Group (CTONG) therapeutic option scoring system: a multiple-parameter framework to assess the value of lung cancer treatment options

Jiu-Wei Cui[1#], Qing Zhou[2#], Shun Lu[3], Ying Cheng[4], Jie Wang[5], Ri-Lan Bai[1], Wen-Qian Li[1], Lei Qian[1], Xiao-Yuan Chen[6], Yun Fan[7], Cheng Huang[8], Xiao-Qing Liu[9], Hai-Yan Tu[2], Jin-Ji Yang[2], Li Zhang[10], Jian-Ying Zhou[11], Wen-Zhao Zhong[2], Yi-Long Wu[2]; on behalf of the Chinese Thoracic Oncology Group (CTONG)

[1]Cancer Center, the First Hospital of Jilin University, Changchun, China; [2]Guangdong Lung Cancer Institute, Guangdong Provincial People's Hospital & Guangdong Academy of Medical Sciences, Guangzhou, China; [3]Department of Shanghai Lung Cancer Center, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai, China; [4]Department of Medical Oncology, Jilin Province Tumour Hospital, Changchun, China; [5]Department of Medical Oncology, National Cancer Center/Cancer Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China; [6]School of Medicine, Tsinghua University, Beijing, China; [7]Department of Medical Oncology, Zhejiang Cancer Hospital, Hangzhou, China; [8]Department of Respiratory Medicine, Fujian Provincial Tumor Hospital, Fuzhou, China; [9]Department of Pulmonary Oncology, 307 Hospital of the Academy of Military Medical Sciences, Beijing, China; [10]Department of Medical Oncology, Sun Yat-Sen University Cancer Center, Guangzhou, China; [11]Department of Respiratory Diseases, The First Affiliated Hospital of College of Medicine, Zhejiang University, Hangzhou, China

Contributions: (I) Conception and design: YL Wu, JW Cui, Q Zhou; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Yi-Long Wu. Guangdong Lung Cancer Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, China. Email: syylwu@live.cn.

Background: Currently, there is no standard context that conforms to the Chinese national framework for evaluating medical decisions regarding the treatment of lung cancer.

Methods: This draft was formulated after a systematic review and a focus group discussion among 20 experts, who were senior physicians with extensive clinical experience from the Chinese Thoracic Oncology Group (CTONG) task force. Subsequently, a draft and a five-point Likert scale were sent to 300 CTONG working group members. These were modified according to feedback from a four-round modified Delphi approach. Hence, the first version of the 'Therapeutic option of lung cancer: CTONG scoring system' was formulated. Afterward, a corresponding questionnaire was designed to collect opinions on the weight allocation of various indicators. This was issued through the WeChat platform, "Oncology News" application and e-mails from October 23, 2020, to November 25, 2020. Participants from numerous occupations in cancer-related fields from various regions of China were included in the study. Overall and subgroup analyses regarding weight allocations were performed. The differences between participant-allocated and reference weights were considered to adjust the framework.

Results: The framework contained four aspects and six indicators, including efficacy [progression-free survival (PFS)/overall survival (OS) and subsequent treatment], safety [treatment-related severe adverse event (SAE), dose adjustment], quality of life (Qol), and compensation. The reference weights were 50%, 5%, 10%, 5%, 10%, and 20% for each indicator. By November 25, 2020, 1,043 valid questionnaires had been obtained. The majority of the questionnaires were completed by physicians (86.5%). Subgroup analysis among the various groups showed an overall consistent trend. Besides, significant differences between the participant-allocated and reference weights were found among PFS/OS (difference: –11.5%), compensation (difference: –10.1%), and subsequent treatment (difference: 9.7%) indicators. After discussion, the final

weight allocations were set at 45%, 10%, 15%, 5%, 10%, and 15% for PFS/OS, subsequent treatment, treatment-related SAE, dose adjustment, Qol, and compensation, respectively.

Conclusions: The CTONG scoring system, as an objective evaluation model that involves multiple parameters, is a breakthrough method for evaluating the therapeutic value of lung cancer treatment options in China, which is worthy of further verification in future clinical practice.

Keywords: Lung cancer; therapeutic option; Chinese Thoracic Oncology Group score (CTONG score); questionnaire

Submitted May 08, 2021. Accepted for publication Aug 17, 2021.
doi: 10.21037/tlcr-21-388
View this article at: https://dx.doi.org/10.21037/tlcr-21-388

Introduction

Lung cancer is a major threat to human health. It is associated with the highest incidence of malignant tumors and mortality in China (1). In recent years, the treatment of lung cancer has been found to be associated with two challenges. Firstly, the advancement of science and technology and an increase in the number of treatment/drug options has made the choice of an appropriate treatment challenging (2,3). Secondly, the increase in psychological, social, and economic needs and considerations can affect patient treatment decisions. This includes a variety of factors, such as efficacy, subsequent regimens, drug toxicity, quality of life (Qol), and treatment costs. However, currently, the lung cancer treatment guidelines regarding clinical decision-making are primarily based on the efficacy of drug therapy. The aforementioned factors that affect treatment are not considered quantitatively or in detail. This makes the development and selection of treatment options potentially unsuitable for specific patients. Therefore, it is necessary to consider and evaluate multidimensional indicators and select the most appropriate and effective treatment plan for patients after weighing the advantages and disadvantages.

At present, for patients with lung cancer in China, there is no standard method for evaluating the suitability of a medical decision that is in line with the national framework. Based on these observations, the Chinese Thoracic Oncology Group (CTONG) recently established an expert working group to develop the first version (draft version) of the CTONG consensus scoring system. This group solicited multidisciplinary opinions from professionals in the field of lung cancer. The group conducted comprehensive deliberations, repeated discussions, as well as revisions of the first draft. A questionnaire was used to collect the opinions of relevant clinicians throughout China on the CTONG score. Finally, the CTONG working group repeatedly revised and improved the first version of the CTONG scoring system according to the feedback and the survey data. They worked to ensure that the system fully reflected the requirements and opinions of the majority of the participants. The working group determined that the CTONG score was in line with China's national framework. In the process of forming the scoring system, we synthesized data from clinical practice, consulted the literature, and referred to the opinions of clinicians, patients, and other relevant individuals. We also reviewed other scoring systems such as the American Society of Clinical Oncology (ASCO) scoring system (4), the European Society for Medical Oncology (ESMO) scoring system (5), and the National Comprehensive Cancer Network (NCCN) guidelines (6). Considering that the socioeconomic development level and medical insurance system in various countries impacts the evaluation, we combined the social and economic characteristics of China to construct a unique scoring system suitable for the Chinese population.

Methods

It should be pointed out that our study did not involve such contents as human experiments, animal experiments and case report, but only conducted expert consensus and online questionnaire survey and therefore waived from an ethical approval.

Formulation of the first version of the scoring system

In April 2020, a CTONG task force was established online to develop the "Therapeutic option of lung cancer: CTONG

scoring system". This system includes scoring criteria and assignment of weights as dimensions of the system. The task force included 20 representative experts with the following criteria: (I) the professional titles of experts were senior doctors with extensive clinical experience; (II) experts had a specialty in lung cancer; (III) experts were chosen from various disciplines, including oncology, thoracic surgery, and radiotherapy; (IV) experts were chosen from various regions and cities in China to minimize geographical and economic heterogeneity. The task force was headed by Dr. Wu, Chairman of CTONG. The process was based on the modified Delphi approach, a consensus-based methodology.

First, to synthesize the current criteria for tumor evaluation, a systematic literature review was conducted by experts from the CTONG task force. Websites including PubMed, Web of science, EMBASE, Cochrane Library, and Chinese National Knowledge Infrastructure (CNKI) were searched with keywords such as "lung cancer", "therapeutic", "treatment", "evaluation", and "score". Subsequently, a CTONG task force focus group discussion was initiated to draft the framework based on literature reviews, evidence from clinical practice, and comprehensive deliberation. Indicators, scoring criteria, assignment of weights, and cutoff values for weight differences were proposed by the experts. The task force voted on these factors to reach a consensus.

After the draft was prepared, it was used as the basis to design a five-point Likert scale (Table S1). The scale was emailed to 300 members of the CTONG working group from various fields of lung cancer along with the draft and relevant materials (7,8). The CTONG members were able to score the degree of their agreement and complete corresponding comments. Two CTONG staff members were responsible for collecting anonymous Likert score sheets and calculating the consent rate for each question.

Subsequently, the CTONG task force held a focus group discussion. According to the consent rates and aggregated opinions, experts first gave their opinions independently and then participated in online discussions. The results were decided upon by a vote. Decisions were made when the agreement threshold was reached. Otherwise, the above process was repeated. Finally, the first version of the framework was formulated as "Therapeutic option of lung cancer: CTONG scoring system".

Questionnaire design and distribution

An Internet-based questionnaire was designed to widely solicit opinions. This was especially evident for the weight allocations to further verify the practicability and acceptability of the first-version (draft version) context for various groups. The participants included in this study were doctors with various professional titles, health care workers, including nurses and pharmacists, patients and their families, as well as other corporate staff in various regions of China. Participants were associated with cancer-related fields and were over 18 years of age. However, medical students without clinical experience, health care workers without oncologic experience, patients without cancer histories, and corporate staff aged under 18 years working in non-cancer fields were not included. We also excluded participants with potential conflicts of interest, including staff from medical companies, insurance companies, and those working for the government health insurance administration.

The questionnaire was designed to include two sections: records of basic information and surveys of weight distribution. Basic information such as occupation, city, hospital, professional title, and department was included. The survey of weight allocation was based on the indicators in the first-generation version. Reference weights were provided in the questionnaire. Participants could adjust weight allocations according to their intention. In addition, a free text space was available to independently collect supplementary subjective opinions on the dimensions and scoring criteria.

The questionnaire was distributed through the WeChat (a very popular communication application in China) platform, "Oncology News" application, and E-mails from October 23, 2020, to November 25, 2020. The participants completed the questionnaire voluntarily and independently. The results could not be queried to avoid bias.

Questionnaire analysis and framework application

The results of the questionnaires were collected, eliminated, and statistically analyzed by an independent third party, "Oncology News", a company focusing on tumor-associated information. Questionnaires that failed to meet the inclusion criteria or those with incomplete information were excluded. In addition, the CTONG task force held a third meeting to discuss the framework adjustments. They also discussed the difference between participant-allocated and reference weight allocations. This was analyzed based on the consensual cutoff value of weight difference during the first version of the framework formulation. After individual

expression and face-to-face discussions, the decision was finalized by a vote.

In addition, the CTONG scale was applied to lung cancer settings to test its' feasibility and utility. We tested the CTONG scale in the first-line application of the second- and third-generation epidermal growth factor receptor-tyrosine kinase inhibitors (EGFR-TKIs), dacomitinib, and osimertinib for the treatment of patients with advanced non-small cell lung cancer (NSCLC) with common EGFR mutations (ex19del and L858R).

Statistical analysis

The analytical methods in the current study mainly included the Delphi method and descriptive statistical analysis. Firstly, in order to reach a consensus for various versions of the CTONG scale framework in the focus group discussion of CTONG task force, the agreement threshold of the modified Delphi approach was defined as a consent rate of ≥75% (9,10). Besides, in terms of questionnaire analyses, a descriptive analysis, including frequency and mean values of weight allocation, was performed among various subgroups to evaluate the feasibility and popularization of the framework. Statistical analyses were performed by SPSS 26.0 and GraphPad 18.0.

Results

First version framework overview

After the four-round modified Delphi approach, the CTONG task force and members of CTONG working group reached a consensus on the first version of "Therapeutic option of lung cancer: CTONG scoring system". It contained four dimensions (efficacy, safety, Qol, and compensation) and six indicators [progression-free survival (PFS)/overall survival (OS), subsequent treatment, treatment-related severe adverse event (SAE), dose adjustment, Qol, and compensation]. While evaluating the therapeutic options, each indicator was assigned a score ranging from 0 to 2, with a score of two points representing the best option. The sum of the weight allocation was 100.0%. Therefore, the evaluation of the PFS (time from tumor diagnosis to tumor progression or death)/OS (time from tumor diagnosis to death of any cause) indicator (reference weight: 50.0%) was dominated by OS results. Differences in OS analyses should be more than 2 months with statistical significance. When OS was unavailable, differences in PFS analyses should be significantly greater than 3 months. Subsequent treatment (reference weight: 5.0%) reflected the efficacy. This was evaluated by the classification of treatment options, including therapies targeting specific biomarkers, general therapy such as chemotherapy, or no standard therapy. With regard to safety parameters, both treatment-related SAE (reference weight: 10.0%) and dose adjustment (reference weight: 5.0%) were assessed with cutoff values of 10.0% and 30.0%. In addition, both the improvement in Qol and the presence of a treatment holiday were considered for measuring the Qol indicator (reference weight: 10.0%). Qol is the metric for both life quantity and quality. An improved Qol should be analyzed using validated scales and based on corresponding reports from clinical studies. Treatment holidays are time off from all anti-tumor therapies without disease progression. The best example of Qol was improvement in Qol with a treatment holiday. This was followed by improvement in Qol without a treatment holiday and absence of improvement in Qol or absence of a treatment holiday. Finally, the compensation indicator (reference weight: 20.0%) reflects the current Chinese health insurance policies and funding policies from pharmaceutical companies. The details of the CTONG scores are presented in *Table 1*.

In addition, a consensus was reached on the cutoff value for weight differences. A difference of less than 3% was considered consistent, while a difference of over 5% indicated a significant difference and signaled that adjustments should be made. Indicators with a difference of 3–5% could also be adjusted according to the overall scores and clinical considerations. In addition, adjustments should be made in units of 5% to facilitate clinical applications.

Results of the questionnaires

Basic characteristics

On November 25, 2020, a total of 1,075 questionnaires had been collected, 32 questionnaires were excluded, and 1,043 valid questionnaires were included. Of the 1,043 participants, 136 (13.0%), 219 (21.0%), and 688 (66.0%) participants were from first-tier cities (Beijing, Shanghai, Shenzhen, and Guangzhou; FTC group), provincial capital cities (excluding the above four cities, PCC group), and other cities (OC group), respectively. The majority of the participants were physicians (902 participants, 86.5%), followed by patients or family members of patients (69 participants, 6.6%), pharmacists (27 participants, 2.6%), and nurses (10 participants, 1.0%).

Table 1 Therapeutic option for lung cancer: CTONG scoring system

Score	Efficacy		Safety		Qol	Compensation	Total score
	PFS/OS	Subsequent treatment	Treatment-related SAE	Dose adjustment			
2 score	High efficacy	Targeted therapy	≤10.0%	≤10.0%	Improvement in Qol with treatment holiday	Within medicare reimbursement	
1 score	General efficacy	General treatment	>10.0%, ≤30.0%	>10.0%, ≤30.0%	Improvement in Qol without treatment holiday	Within other compensation	
0 score	Low efficacy	No standard treatment	>30.0%	>30.0%	absence of improvement in Qol or absence of treatment holiday	Without compensation	

Efficacy: includes both PFS/OS and subsequent treatment. PFS/OS: OS predominates, and differences in OS and PFS analyses should be more than 2 and 3 months, respectively, with statistical significance; Subsequent treatment included targeted therapy with a specific biomarker, general therapy such as chemotherapy, and no standard therapy. Qol: both Qol and treatment holidays were considered. Improved Qol should be analyzed using validated scales and reference to corresponding clinical studies; treatment holidays were defined as time-off anti-tumor therapies without disease progression. CTONG, Chinese Thoracic Oncology Group; PFS, progression-free survival; OS, overall survival; SAE, severe adverse events; Qol, quality of life.

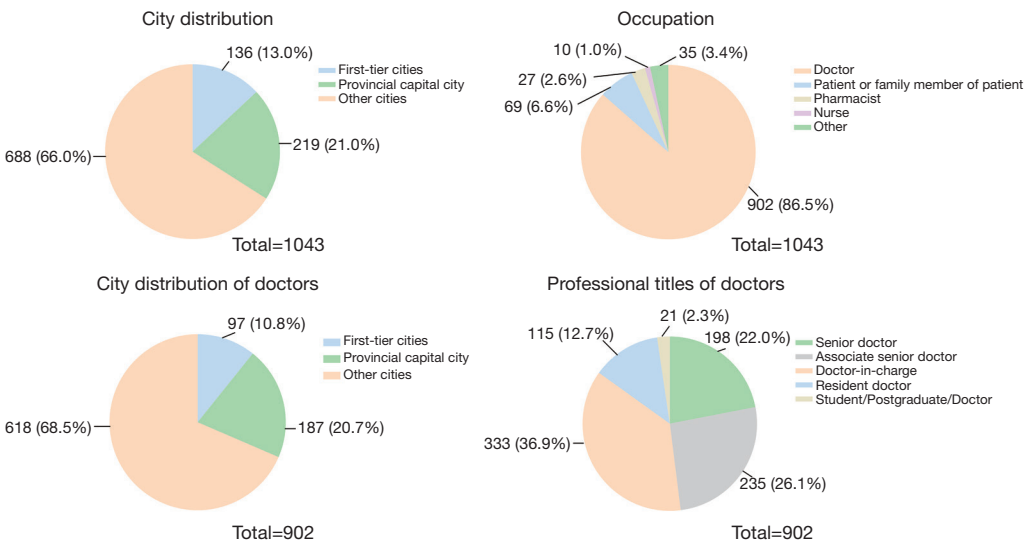

Figure 1 Basic characteristics of questionnaire participants.

Professional titles of the 902 doctors included senior doctors (198 participants, 22.0%), associate senior doctors (235 participants, 26.1%), doctor-in-charge (333 participants, 36.9%), resident doctors (115 participants, 12.7%), and students (including students with postgraduate and doctoral degrees, 21 participants, 2.3%). Basic information is shown in *Figure 1*. The content of the questionnaire is described in Table S2.

Overall and subgroup analysis of the weight allocations

The average weight assignments indicated by all participants were calculated from questionnaires (PFS/OS, 38.5%; subsequent treatment, 14.7%; treatment-related SAE, 14.7%; dose adjustment, 9.4%; Qol, 12.7%; compensation, 9.9%). Differences between participant-assigned and reference weight allocations were found among the various occupation-based subgroups (*Figure 2A*). Pharmacists paid more attention to treatment-related SAEs (16.4%); patients and their family members focused more on Qol (15.8%), and nurses were more concerned with subsequent treatment (19.9%), and dose adjustment (13.5%) than were participants of other professions. With regard to doctors,

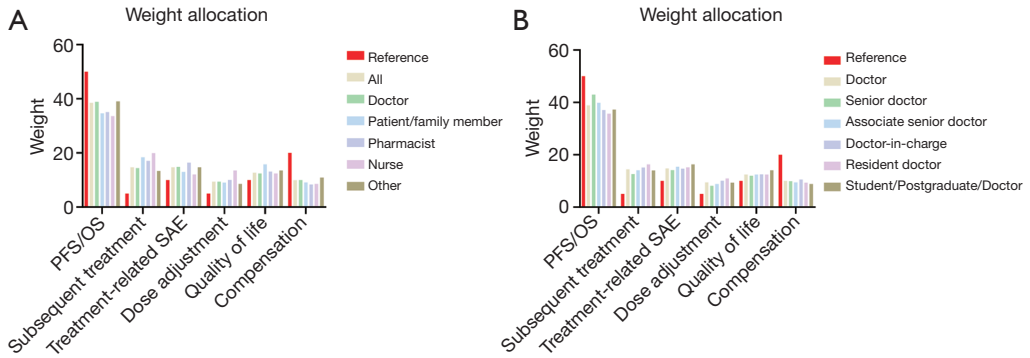

Figure 2 Weight allocation among participants with various occupations (A) and doctors with various professional titles (B). PFS, progression-free survival; OS, overall survival; SAE, severe adverse events.

the weight allocation of PFS/OS (38.9%), treatment-related SAE (14.8%), and compensation (10.0%) indicators were higher than the overall average value, while the weight allocation values of subsequent treatment (14.4%) and Qol (12.4%) were lower.

Furthermore, we performed professional title-based subgroup analyses. Senior doctors allocated higher weight (43.0%) to PFS/OS and lower weight to subsequent treatment (12.7%), treatment-related SAE (14.2%), dose adjustment (8.2%), and Qol (12%). Doctors-in-charge paid greater attention to terms of compensation (10.6%). Resident doctors provided higher allocations to subsequent treatment (16.4%) and dose adjustment (10.9%), and lower allocation to PFS/OS (35.7%). Students allocated higher weight to treatment-related SAE (16.4%) and Qol (14.1%) and lower weight to compensation (8.8%) than doctors of other professional titles. The details are presented in *Figure 2B*.

Moreover, geographical subgroup analyses revealed that doctors from the FTC group conferred higher weight on PFS/OS (42.0%) and Qol (12.8%) and a lower weight on compensation (9.0%). Physicians in the OC group conferred higher weight on subsequent treatment (14.6%), safety (treatment-related SAE: 15.0%, dose adjustment: 9.8%), and compensation (10.1%), and a lower weight on PFS/OS (38.1%). Doctors in the PCC group conferred higher weight on treatment-related SAEs (15.0%) than did those from other groups. In addition, we further explored the differences in weight allocation among doctors with various professional titles in the geographical subgroups. In terms of PFS/OS, significant differences between the FTC and OC groups were found for both resident doctors (8.7%) and associate senior doctors (5.0%). The results are presented in *Table 2*.

Framework adjustments

To synthesize the results from questionnaires and further adjust the CTONG scale, differences between average participant-assigned weights and reference weights from the first-version framework were further analyzed. The most significant differences (participant-assigned *vs.* reference weight) were found in PFS/OS (38.5% *vs.* 50.0%, difference: –11.5%), compensation (9.9% *vs.* 20.0%, difference: –10.1%), and subsequent treatment (14.7% *vs.* 5.0%, difference: 9.7%), followed by treatment-related SAE (14.7% *vs.* 10.0%, difference: 4.7%), and dose adjustment (9.4% *vs.* 5.0%, difference: 4.4%). After the third meeting of the CTONG task force, the weights of PFS/OS and compensation were both adjusted with a reduction of 5%, while that of subsequent treatment was adjusted with an increase of 5%. The participant-assigned weight of Qol from the questionnaires was consistent with the reference weight (12.7% *vs.* 10.0%, difference: 2.7%), which did not require adjustment. In addition, to achieve a total weight score of 100%, treatment-related SAE showed both a higher difference (participant-assigned *vs.* reference weight: 4.7% *vs.* 4.4%) and reference weight value (10% *vs.* 5%) than did dose adjustment. Treatment-related SAE was also a more commonly reported indicator to evaluate the safety and efficacy of therapeutic drugs in clinical trials. It was also adjusted by a 5% increase. In general, with unanimous approval by the CTONG task force, the final weight allocation of the CTONG score was finalized as follows: PFS/OS, 45%; treatment-related SAE and compensation, 15% each; subsequent treatment and Qol, 10% each, and dose adjustment, 5%. The selection of indicators, as well as

Table 2 Weights assigned by doctors from various cities to the various indicators

Professional title of doctors	City distribution	PFS/OS		Subsequent treatment		Treatment-related SAE		Dose adjustment		QoL		Compensation	
		Value	D	Value	D	Value	D	Value	D	Value	D	Value	D
All doctors	FTC	42.0	3.9	13.7	0.9	13.7	1.3	8.8	1.2	12.8	0.7	9.0	1.1
	PCC	40.3		14.2		15.0		8.6		12.1		9.8	
	OC	38.1		14.6		15.0		9.8		12.4		10.1	
Senior doctors	FTC	44.4	4.1	12.1	1.2	13.2	1.2	7.6	2.1	14.5	3.4	8.1	2.6
	PCC	45.8		11.9		13.8		6.8		11.1		10.7	
	OC	41.7		13.1		14.4		8.9		11.9		9.9	
Associate senior doctors	FTC	43.6	5.0*	14.0	0.9	12.4	3.3	8.1	1.9	12.1	0.9	9.8	0.5
	PCC	42.3		13.4		15.5		7.5		11.8		9.5	
	OC	38.6		14.3		15.7		9.4		12.7		9.3	
Doctors-in-charge	FTC	42.0	3.9	13.7	0.9	13.7	1.3	8.8	1.2	12.8	0.7	9.0	1.1
	PCC	40.3		14.2		15.0		8.6		12.1		9.8	
	OC	38.1		14.6		15.0		9.8		12.4		10.1	
Resident doctors	FTC	42.8	8.7*	13.1	3.9	12.9	2.9	9.1	2.5	13.9	1.9	8.4	1.8
	PCC	36.6		17.0		14.7		9.5		12.0		10.2	
	OC	34.1		17.0		15.8		11.6		12.2		9.4	

*, significant difference. FTC, first-tier cities; PCC, provincial capital cities; OC, other cities; PFS, progression-free survival; OS, overall survival; SAE, severe adverse events; QoL, quality of life; D, difference.

Table 3 Weight adjustments

Parameters	Efficacy		Safety		QoL	Compensation
	PFS/OS	Subsequent treatment	Treatment-related SAE	Dose adjustment		
Weight reference	50.0%	5.0%	10.0%	5.0%	10.0%	20.0%
Weight in questionnaires	38.5%	14.7%	14.7%	9.4%	12.7%	9.9%
Difference	–11.5%*	+9.7%*	+4.7%	+4.4%	+2.7%	–10.1%*
Weight adjustment	–5%	+5%	+5%	–	–	–5%
Final weight	45%	10%	15%	5%	10%	15%

*, significant difference. PFS, progression-free survival; OS, overall survival; SAE, severe adverse events; QoL, quality of life.

the scoring criteria, did not change according to the first-generation framework. The results are presented in *Table 3*.

Framework test

Based on the final version of the CTONG scale, we used the scale to score the first-line dacomitinib and osimertinib treatment in patients with advanced NSCLC showing ex19del and L858R mutations according to the phase 3 ARCHER1050 trial and phase 3 FLAURA trial (11-14). For patients with the ex19del mutation, the total CTONG score was 1.5 for osimertinib and 0.5 for dacomitinib. For patients with the L858R mutation, dacomitinib and osimertinib achieved total CTONG scores of 1.4 and 0.6, respectively (details shown in *Figure 3*). Thus, according to the CTONG score, osimertinib and dacomitinib were

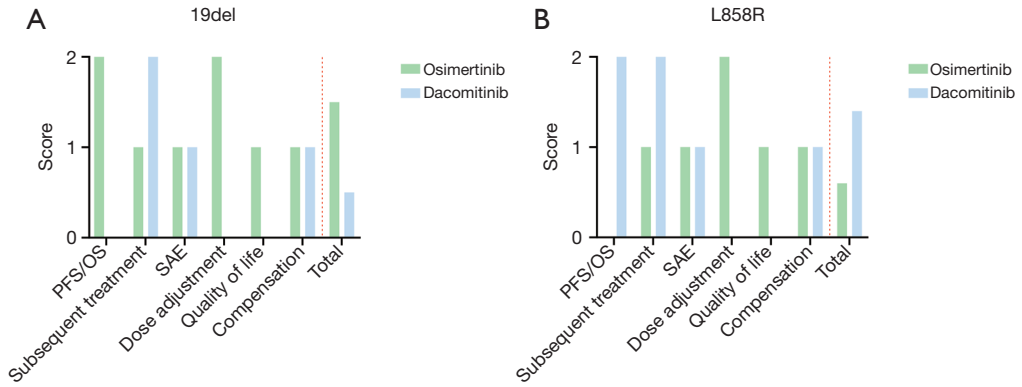

Figure 3 Results of the framework test for first-line dacomitinib and osimertinib treatment in patients with advanced NSCLC showing ex19del (A) and L858R (B) mutations. NSCLC, non-small cell lung cancer; PFS, progression-free survival; OS, overall survival; SAE, severe adverse events.

recommended for patients with NSCLC showing an ex19del mutation and L858R mutation, respectively.

Discussion

With the rapid development of targeted therapy and immunotherapy for lung cancer, the number of anti-tumor drugs used in clinical practice has increased. This has increased the difficulty of decision-making regarding treatment/drug options. In the past, the clinical efficacy of the drug was the most important factor. However, with the increasing consideration of social, psychological, and economic factors, it has gradually been recognized that clinical decision-making is affected by numerous factors. Driven by the above factors, the CTONG expert group developed the CTONG scoring system for Chinese patients with lung cancer. They conducted a reasonable comprehensive revision by soliciting the opinions of various types of participants. This process was completed in order to increase the system's suitability for China's national circumstances and to fully reflect the treatment needs of various levels of the population. Overall, the development of the CTONG scale is scientific and reasonable and has been recognized by the majority of the population.

During the development process, the CTONG working group solicited the opinions of expert groups from multidisciplinary areas in the field of lung cancer. This was based on the four dimensions that are currently widely recognized: drug efficacy, safety, Qol, and compensation. This process comprehensively considered and preliminarily developed a scoring framework and the various indicators to be assessed. In terms of efficacy, in addition to focusing on the most traditional indicators, PFS and OS, the CTONG group believed that considering whether patients have the support of other subsequent treatments while judging the efficacy of the primary treatment could be innovative and important. In addition to efficacy, the use of reasonable and safe medications should be emphasized to ensure the safety of drugs. For example, the adverse events (AEs) of patients should be minimized while achieving the best clinical efficacy of drugs and achieving the goal of "high efficiency and low toxicity." In terms of safety, SAEs that have a great impact on efficacy, treatment process, and various drug dose modifications are mainly considered. The CTONG group believes that the reduction of a drug dose or discontinuation of a drug can more accurately visualize the overall safety of the drug. This is different from the meaning expressed by SAEs and thus it is included in the safety category. Furthermore, they noted that tumor-bearing survival, efforts to improve the Qol of patients, and a reduction in their hospital stay cannot be ignored. A reduction in hospital stays as well as making the patient more comfortable outside the hospital should also be regarded as one of the factors to judge the Qol. In addition, in recent years, the value of medicine (15) has been gradually drawn attention. It is believed that "the relative relationship between the benefits obtained by patients and the cost—that is, efficiency—can be recognized as the core of the value of cancer treatment". Therefore, the inclusion of the economic situation of patients and drug compensation in the process of cancer treatment is recommended. This component is greatly affected by national and regional policies. Thus,

emphasizing the importance of formulating unique scoring criteria that are in line with the national context of developing countries such as China is critical. Overall, the framework development and indicator selection processes of the CTONG score were reasonable and comprehensive. They are widely recognized as the main factors affecting the selection of treatment options/drugs.

After selecting the evaluation indicators, to measure the importance of various indicators more reasonably and objectively, the CTONG expert group preliminarily distributed the weights and grade scores of various indicators. This took place after repeated discussions and voting while evaluating the existing clinical trial results and relevant evidence combined with clinical experience. To reasonably revise the weight of the CTONG score and to improve its' applicability to clinical practice, the CTONG working group used a questionnaire to solicit the opinions of thousands of people from diverse backgrounds throughout China on the CTONG scale. Overall, the results of the survey were roughly in line with the framework of the first-generation scale: efficacy, safety, and Qol were the focus of clinicians' considerations. Drug efficacy (OS/PFS) accounted for the largest proportion, followed by treatment-related SAEs, subsequent treatment, Qol, and dose adjustment. This signified that the indicator selection and weight allocation of the first-generation CTONG scale were positively recognized by most groups. This further validated the scientific accuracy and rationality of the strategy developed by the CTONG expert group.

Most of the differences in the average weights allocated by individuals with various professional titles between various cities were within 3%. The differences were within the allowable range. However, there are significant differences in weight considerations among certain groups. This reflects the differences in their needs. Regarding the weight distribution of OS/PFS, the difference between various cities was greater than 3% for all physicians with various professional titles. This was especially true for resident doctors and associate senior doctors. Considering that the basic weight of OS/PFS is relatively large, it is greatly affected by the economy and preferential policies among various cities. Analysis of the weight bias of doctors with various professional titles in various cities showed that physicians from FTCs and other PCCs, especially those with senior professional titles, assigned high weight to PFS/OS and Qol. They assigned relatively little weight to compensation. This may be related to the fact that these PCCs have better economic conditions and greater support for medical and commercial insurance than OCs. This results in an increase in the number of drugs in the compensation category. Thus, the doctors that assigned less weight to compensation will be less likely to understand the overall treatment of patients as well as humanistic care. Physicians from OCs conferred a relatively higher weight allocation on subsequent treatment, safety, and compensation than on efficacy. The results of the survey also showed that participants with various occupations and doctors with various professional titles showed subtle differences in the allocation of weights of the value indicators due to differences in their respective professional duties and degree of contact with patients. For example, senior doctors and associate senior doctors primarily considered the overall efficacy of treatment and treatment-related SAEs. Residents mostly considered subsequent treatment and dose adjustment. Students, patients, and their families considered Qol and the safety issues of patients. Nurses paid more attention to dose adjustment, subsequent treatment, and other factors affecting patient management because of their close contact with patients. Pharmacists paid the highest attention to drug treatment-related SAEs. The analysis of these data is helpful for the subsequent consideration of various regions and groups with various professional titles for dynamic adjustment in the future.

Although the questionnaire results were consistent with the overall trend of the first version, some significant differences remained. First, compensation was associated with the greatest difference. The first-generation framework defined its' weight as 20%, while the investigational results defined it as 9.9%. Due to the influence of policies, compensation mechanisms, etc., there are differences in the weight allocation results of the various groups. In recent years, with the popularization of national preferential policies such as medical insurance and other compensation the accessibility of various drugs has significantly improved. Thus, this situation reduced the weight distribution of compensation. This also reflects dynamic changes in each factor in clinical practice. Therefore, the evaluation of weight should also be dynamically monitored and adjusted according to the changes in national priorities and policies to better meet the clinical needs of patients. Second, in terms of efficacy as measured by PFS/OS, the weight in the first version of the CTONG score was set at 50%, while the mean value of the survey results was 38.5%. Considering that there has been an increase in the interest of the Chinese

population regarding psychological, social, and humanistic factors in recent years, the weight of efficacy may be appropriately affected by these aspects. This may result in a moderate decrease in the weight of PFS/OS indicators. Third, in terms of safety, the weight of SAEs after the actual survey increased by approximately 5%. This indicated that the interest in the occurrence of treatment-related SAEs in clinical practice has increased. This is directly related to the patient's tolerance and whether they could continue to receive the treatment. However, for dose adjustment, the weight in the first-generation CTONG score accounted for only 5%, while the average weight determined by the survey was 9.4%. Drug dose reduction or discontinuation is an accurate indication of the overall safety of a drug. It includes modifications due to low-grade AEs or other reasons. This reflects the importance physicians and patients placed on the overall safety and dose management of a drug.

Considering the difference between the questionnaire results and the initial set point, the CTONG expert group developed the scale revision by consulting the existing literature, evaluating the relevant evidence, combining statistical analysis, and repeatedly deliberating and voting. Based on these principles and findings, the authors finalized the CTONG scoring system after repeated discussions. The weight distribution of the final CTONG score was adjusted relative to the initial version as follows: the weights for PFS/OS and compensation were decreased by 5% and those of subsequent treatment and SAEs were increased by 5%. The weight of Qol and dose modifications did not require any changes. Among these, SAE and dose adjustments were controversial. First, the revised value of SAEs was slightly higher than the original values. Second, the initial weight of SAEs was higher. This indicated that SAEs are critically important in the opinion of experts. Considering that the clinical impact of drug SAEs on efficacy, treatment course, and the condition and mortality of the patient is higher than that of dose adjustment, the CTONG working group choose to add additional weight to SAEs. Thus, in the process of developing scoring systems, experts in relevant fields should not only scientifically, objectively, and rationally analyze the value of drugs, but also solicit the opinions of various types of closely related groups for comprehensive consideration. This method serves to correct the preliminary assumption of the CTONG score, to improve its' scientific accuracy and reasonability, and to increase its' practicability for future clinical applications. This is the scientific method that was necessary for strategy development.

Several organizations have assessed the value of various treatment/drug options for patients with tumors. They have proposed comprehensive methods for judging the suitability of treatment choices using various indicators. However, adjustment and revision of these methods using questionnaires were not performed. The ASCO scoring system is based on a comparison of net health benefits (NHB) (including clinical benefits, toxicity, and bonus points) and costs (including drug acquisition cost and patient co-pay) (4). The NCCN guidelines provide comprehensive recommendations by considering the efficacy, safety, quality, consistency of evidence, and economic affordability of the regimen/drug (6). The ESMO established the ESMO Magnitude of Clinical Benefit Scale (ESMO-MCBS) to comprehensively consider the therapeutic option for tumors, based on six aspects: curative and non-curative setting, prognosis, PFS/OS, HR, Qol, and toxicity (5). The development of the CTONG system may compensate for the shortcomings of existing scoring systems. It is not only highly scientific and involves the selection of simple indicators but also quantifies and reasonably revises the weight of indicators. This increases the applicability in clinical settings and is expected to effectively assess the value of treatment options. The overall trend of the CTONG scale is consistent with that of the previous ASCO scoring framework. However, there remain subtle differences. These differences may be associated with variations in drug availability, health insurance, economic status, and the humanistic environment in various geographical countries. This may result in different assessment principles and weights. Overall, the design framework and indicator weights of the CTONG scale have been generally recognized by doctors, nurses, pharmacists, patients, and their families. This further affirms the rationality and standardization of this scale. Moreover, this scoring system is in line with China's national priorities and is suitable for the comprehensive assessment of the value of various regimens/drug treatments for the treatment of lung cancer.

Nevertheless, there may be many challenges associated with the actual popularization of the CTONG score in a clinical setting. A complete understanding of the development, revision process, and consideration basis of the CTONG score is helpful for dynamic assessment, adjustment, continuous updates, and optimization in the future. The following are some suggestions that we enumerated after synthesizing the opinions of multiple parties: (I) Although the development of this scoring

system is standardized, its' overall design requires further optimization and improvement. For example, each indicator in the ASCO scoring system is divided into five scoring levels, while that in the CTONG scoring system is divided into only three. Which of these two methods is superior and more applicable during clinical application must be continuously examined and adjusted. (II) The severity and management of various treatment-related AEs have dissimilar effects on the efficacy of treatment. For example, although some treatment-related AEs (such as skin and endocrine events) reached grade 3 or higher and were not easily or completely reversed, most were easy to manage and had little effect on the patients' treatment process. Moreover, some patients had only grade 2 treatment-related AEs (such as pneumonia, pancreatitis, etc.). However, the effects of these AEs on the health of patients are significant and are difficult to manage and reverse. Therefore, when assessing the therapeutic value of a drug, it is recommended that the degree of impact on a patient's health and the corresponding management score be considered. This will enable patients to fully understand the severity of AEs and their impact on treatment efficacy. (III) In recent years, with the advancement of China's national policies, the scope of medical and commercial insurance support has expanded. The number of drugs that have been placed under the compensation category has increased and the product value has been gradually adjusted. Furthermore, various generic drugs are now available and the living standards of residents are continuously improving. Therefore, under the umbrella of economic security, most residents consider the efficacy and safety of drugs. Qol is more important to them than compensation. In practice, the developed scoring systems should be dynamically adjusted according to changes in economic policies, etc. (IV) Although the weight assigned to compensation is lower in the current version of the scoring system than in the previous version, compensatory mechanisms may vary across economic levels. This affects the participants' consideration of their scoring weights and the selection of the treatment regimen during the survey. For county-level or municipal hospitals in remote areas with relatively poor economic conditions, the impact of treatment costs, drug accessibility, and medical and commercial insurance is of great significance. It is the basic indicator of efficacy, Qol, and safety. Therefore, regional differences may influence the framework of the scoring system in the future. (V) The assessment of the value of any treatment should be dynamically tailored to accommodate new medical information. This will ensure that the weights of various items are sufficiently reflected in the treatment setting. In practice, the scoring system should be dynamically adjusted according to national priorities, the development, and changes in various indicators, and the opinions of other residents. These actions will gradually promote the upgrading and improvement of the system. Finally, regarding the statistical research method, the Delphi method we applied is only one of the methods available for experts to form a preliminary consensus. We formed the first draft of the scoring system using this method. This method may not be the most reliable research methodology for answering a research question. Therefore, our expert panel conducted a focus group discussion, an expert face-to-face discussion, and a questionnaire survey to gather additional evidence. However, there remain many shortcomings that require further verification in clinical practice.

The CTONG score is proposed as an objective evaluation model with multiple factors, multiple parameters, and far-reaching clinical and social significance. First, the CTONG score provides reasonable guidance and an objective basis for the selection of cancer treatment/drug options in clinical practice. It is beneficial when a single disease can be targeted with several drugs. Second, the proposed CTONG score can indicate the demand for new drugs. This can determine the direction of research focusing on the development of novel therapeutics and corresponding strategies. It can also ascertain the direction of market operations. The CTONG score suggests that the value of drugs should be comprehensively assessed from many aspects and perspectives. For example, not only should the efficacy of drugs but also the safety data and improvement of Qol should be assessed. When data obtained by these assessments are available, drugs are more likely to be approved for marketing based on a comprehensive consensus. An efficacy evaluation should be performed after the completion of this process. Therefore, the CTONG score is helpful in guiding pharmaceutical companies to investigate future drug indications and the research and development of new regimens/drugs. Thus, it can direct the development of new drugs with "high efficiency and low toxicity". Furthermore, the utilization of the CTONG score can also direct the design of future clinical trials. It can promote the rationality of the study design, allowing for a breakthrough over previous designs that relied only on PFS/OS. By incorporating relevant data such as the data obtained from large clinical trials and expert consensus of the CTONG group, the CTONG

scoring system provides corresponding evaluation criteria for multiple factors. These include the judgment of PFS/OS, dose adjustment, improvement of Qol as well as a reference for the development of evaluation indicators and endpoints in clinical trials. For example, the evaluation of the improvement of Qol in the CTONG score is primarily based on the results reported in previous large phase III clinical trials. This suggests that Qol should be used as an important assessment indicator in the design of future clinical trials. Finally, the CTONG score may have reference significance for recommendations regarding cancer treatment. As shown in the results section, osimertinib is preferentially recommended and dacomitinib is generally recommended for patients with advanced NSCLC showing the ex19del mutation based on the CTONG composite score. However, for patients with NSCLC showing the L858R mutation, dacomitinib is preferentially recommended, whereas osimertinib is generally recommended. It is noted that for the EGFR-targeted drugs recommended in parallel, the guidelines in the CTONG score are subdivided into priority, general, and other recommendations by comprehensively considering multiple factors. Thus, it provides a basis for refining clinical guidelines in the future.

In particular, it should be observed that our study supports "patient-centered" treatment. When determining the evaluation indicators, we solicited the opinions of front-line clinicians, carried out face-to-face discussions by experts in focus group discussions, and widely solicited the questionnaire survey results of various groups. We finally determined six indicators that affected cancer treatment decisions. First, experts in the development of the scoring framework are front-line clinicians who are knowledgeable regarding patient needs and factors that influence treatment options. These specialists understand the real needs of patients through direct contact in clinical practice. Second, we did not only adopt the opinions of the expert group. In order to accurately reflect the needs of the public, we distributed a questionnaire to include patients and considered this issue from the perspective of patients. This also preliminarily represented the opinions of patients. Finally, the developed CTONG score is a quantitative decision framework that considers various factors that may affect patients' treatment decisions. However, the implementation process for clinical practice may be individualized. That is, patients may consider their own conditions (such as drug accessibility and economic factors) for individualized treatment selection. Adherence during treatment is also one of the factors that affect treatment decisions. In the process of identifying indicators that influence cancer treatment decisions, we considered treatment-related SAE and dose adjustments. In terms of Qol, we also considered a treatment holiday, which is the embodiment of patient treatment adherence. Finally, it should be noted that various countries may have various national priorities and characteristics owing to differences in their cultures, policies, as well as other factors. Therefore, the evaluation criteria for the therapeutic options for cancer may be different than in China. The CTONG score was primarily developed based on the investigation carried out in the Chinese population. Therefore, its' suitability for other countries should be further verified. This scoring system must be adjusted accordingly in the future.

Conclusions

Overall, the CTONG score considers various specific factors corresponding to the Chinese population with lung cancer. It includes efficacy, safety, Qol, and compensation. It was officially proposed after conducting a large questionnaire survey, expert group discussions, and repeated revisions. This is in line with China's national priorities and has been widely recognized. The current CTONG score is an effective and breakthrough method established for evaluating the therapeutic value of treatment options for lung cancer in China. It provides a reference for the formulation of guidelines regarding the selection of treatment options and the development of novel therapeutic medicines. However, it should be noted that our study only provides a reference or decision option for clinical practice and future guideline development, rather than a guideline. It is worthy of continuous verification of its future effectiveness in clinical practice studies. The ultimate purpose of this scoring system is to provide comprehensive information regarding the treatment options to doctors and patients. This will enable them to make fully informed decisions regarding the best available therapy with the least number of side effects and the lowest economic cost.

Acknowledgments

We would like to thank medical writing service, Editage, for the help in polishing our paper.
Funding: The study was supported by National Key R&D

Program of China (Grant No. 2016YFC1303800).

Footnote

Data Sharing Statement: Available at https://dx.doi.org/10.21037/tlcr-21-388

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/tlcr-21-388). YLW reports advisory services for AstraZeneca, Boehringer Ingelheim, Novartis, Takeda; personal fees from AstraZeneca, Beigene, Boehringer Ingelheim, BMS, Eli Lilly, MSD, Pfizer, Roche, Sanofi; grants from AstraZeneca, Boehringer Ingelheim, BMS, Hengrui, and Roche, outside the submitted work. WZZ reports honoraria from AstraZeneca, Eli Lilly, Pfizer, Roche, and Sanofi, outside the submitted work. QZ reports honoraria from AstraZeneca, Boehringer Ingelheim, BMS, Eli Lilly, MSD, Pfizer, Roche, and Sanofi, outside the submitted work. The other 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. Our study did not involve such contents as human experiments, animal experiments and case report, but only conducted expert consensus and online questionnaire survey and therefore waived from an ethical approval.

References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
2. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. Nature 2018;553:446-54.
3. de Mello RA, Neves NM, Tadokoro H, et al. New Target Therapies in Advanced Non-Small Cell Lung Cancer: A Review of the Literature and Future Perspectives. J Clin Med 2020;9:3543.
4. Schnipper LE, Davidson NE, Wollins DS, et al. American Society of Clinical Oncology Statement: A Conceptual Framework to Assess the Value of Cancer Treatment Options. J Clin Oncol 2015;33:2563-77.
5. Cherny NI, Sullivan R, Dafni U, et al. A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). Ann Oncol 2017;28:2901-5.
6. Carlson RW, Jonasch E. NCCN Evidence Blocks. J Natl Compr Canc Netw 2016;14:616-9.
7. Xue Y, Hauskrecht M. Active Learning of Classification Models with Likert-Scale Feedback. Proc SIAM Int Conf Data Min 2017;2017:28-35.
8. Likert R. A technique for the measurement of attitudes. Archives of Psychology 1932:140:141-55.
9. Loblaw DA, Prestrud AA, Somerfield MR, et al. American Society of Clinical Oncology Clinical Practice Guidelines: formal systematic review-based consensus methodology. J Clin Oncol 2012;30:3136-40.
10. Qaseem A, Forland F, Macbeth F, et al. Guidelines International Network: toward international standards for clinical practice guidelines. Ann Intern Med 2012;156:525-31.
11. Wu YL, Cheng Y, Zhou X, et al. Dacomitinib versus gefitinib as first-line treatment for patients with EGFR-mutation-positive non-small-cell lung cancer (ARCHER 1050): a randomised, open-label, phase 3 trial. Lancet Oncol 2017;18:1454-66.
12. Mok TS, Cheng Y, Zhou X, et al. Improvement in Overall Survival in a Randomized Study That Compared Dacomitinib With Gefitinib in Patients With Advanced Non-Small-Cell Lung Cancer and EGFR-Activating Mutations. J Clin Oncol 2018;36:2244-50.
13. Soria JC, Ohe Y, Vansteenkiste J, et al. Osimertinib in Untreated EGFR-Mutated Advanced Non-Small-Cell Lung Cancer. N Engl J Med 2018;378:113-25.
14. Ramalingam SS, Vansteenkiste J, Planchard D, et

al. Overall Survival with Osimertinib in Untreated, EGFR-Mutated Advanced NSCLC. N Engl J Med 2020;382:41-50.

15. Zhou Q, Wu YL. Developing CSCO Lung Cancer Practice Guidelines Stratified by Resource Availability and Treatment Value. J Glob Oncol 2017;3:285-8.

Cite this article as: Cui JW, Zhou Q, Lu S, Cheng Y, Wang J, Bai RL, Li WQ, Qian L, Chen XY, Fan Y, Huang C, Liu XQ, Tu HY, Yang JJ, Zhang L, Zhou JY, Zhong WZ, Wu YL; on behalf of the Chinese Thoracic Oncology Group (CTONG). The Chinese Thoracic Oncology Group (CTONG) therapeutic option scoring system: a multiple-parameter framework to assess the value of lung cancer treatment options. Transl Lung Cancer Res 2021;10(8):3594-3607. doi: 10.21037/tlcr-21-388

