



Systematic Review and Quality Appraisal of Literature on Economic Evaluation of Drugs for Type 2 Diabetes

Raela·Abduhilil, Lu Siyu, Sun Lihua*

(School of Business Administration, Shenyang Pharmaceutical University, Shenyang 110016, China)

Abstract

Objective To systematically evaluate the overall characteristics and quality of literature on economic evaluation of drugs for type 2 diabetes, and to provide recommendations for future research and related decision-making.

Methods The economic evaluation literature on drugs for type 2 diabetes based on the Chinese population were searched from CNKI, Wanfang, VIP and PubMed databases. The literature that met the inclusion criteria was selected, and the key research elements were extracted. Meanwhile, the quality of health economics studies (QHES) was used to evaluate the quality of the literature. **Results and Conclusion** A total of 325 articles (296 in Chinese and 29 in English) were included. Most of the studies were conducted by medical institutions (247 articles, 76.0%) and they used cost-effectiveness analysis (295 articles, 90.8%). However, most of the included literature did not report study perspective (267 articles, 90.2%). The average QHES score of the included literature was 57.29, and the quality of the literature was concentrated in “low quality” (123 articles, 37.8%) and “average quality” (138 articles, 42.5%). Literature published in English journals, or using modeling methods, or from universities had relatively higher quality. The quality of economic evaluation literature on drugs for type 2 diabetes based on the Chinese population is generally low at present. There are many problems, such as single research institutions and methods, lack of research perspectives, and no sensitivity analysis.

Keywords: type 2 diabetes; economic evaluation; quality evaluating; QHES

With the development of economy, the improvement of living standards, the change of lifestyle and the acceleration of population aging, the prevalence and mortality of diabetes mellitus (DM) are on the rise all over the world, and it has become one of the seriously endangering human health chronic diseases ^[1]. According to the International Diabetes Federation (IDF) ^[2], the prevalence of diabetes in the world is 10.5% in 2021, and there

are totally 536.6 million adults (aged between 20 and 79) with diabetes, among which the number of Chinese diabetics is 140.9 million. The number of Chinese diabetics is the highest in the world and it is much higher than that in India, which ranks the second. Type 2 diabetes mellitus (T2DM) is the most common diabetes, accounting for more than 90% of all diabetes mellitus ^[3]. The management of T2DM is comprehensive, including lifestyle management, drug therapy, education management, and psychological counseling ^[3, 4]. Among them, drug therapy is the most common and direct way for patients with T2DM to

* Corresponding author: Sun Lihua, professor. Major research area: Pharmacoeconomics, medical investment performance and management, etc. Tel: 024-23986553, E-mail: slh-3632@163.com.



control blood sugar. If the blood sugar of patients fails to be controlled by lifestyle intervention in a certain period, hypoglycemic drugs must be used^[5], and a lifelong medication intervention is required.

The economic evaluation of drugs can select the optimal treatment strategy on the basis of ensuring safety and effectiveness, providing an economic reference for patients with T2DM^[6, 7]. The quality of economic evaluation of drugs for T2DM is particularly important, because decisions made based on high-quality economic evaluation results can maximize the reduction of decision risks and errors^[8]. Therefore, this study systematically evaluated the overall characteristics and quality of economic evaluation literature on drugs for T2DM, identified their problems, and proposed some suggestions for subsequent research and related decisions.

1 Data and methods

1.1 Data source

The CNKI, WanFang, VIP and PubMed databases were searched respectively and systematically from the establishment of each database to December 2022. The Chinese and English search terms were “cost-effectiveness analysis” “cost-benefit analysis” “cost-utility analysis” “minimum cost analysis” “pharmacoeconomics” “pharmaceutical economics” and “type 2 diabetes mellitus”, etc. The combination of keywords and other free words was used for advanced retrieval according to the characteristics of each database.

1.2 Inclusion and exclusion criteria

Inclusion criteria: (1) The disease type was T2DM or its complications; (2) Both the experimental group and the control group must be medication intervention; (3) The research type was a complete economic evaluation study, such as cost-effectiveness analysis (CEA), cost-benefit analysis (CBA), cost-utility analysis (CUA), and cost-minimization analysis (CMA); (4) Full text literature; (5) Papers published in Chinese or English; (6) The population studied must

be Chinese.

Exclusion criteria: (1) The disease was non-T2DM; (2) Economic study of non-drug intervention; (3) Inconsistent types of research, such as reviews, case studies, reviews, etc.; (4) Repeated or unavailability of full text publication; (5) The population studied was foreign population.

1.3 Data extraction and quality evaluation

An independent data extraction table was designed, and the extraction contents were: (1) Literature publication information such as literature title, journal name, journal level (core or non-core journals), publication year, and publication language; (2) Author information such as name of the first author, type of the first author institution, location of the first author institution, and number of authors; (3) Study information such as intervention type, design type, study perspective, economic evaluation methods, incremental analysis and threshold, sensitivity analysis and methods.

The quality of health economics studies (QHES) evaluation scale was used for the economic evaluation literature on drugs for T2DM that met the inclusion criteria. Each term of QHES scale had a weight, ranging from 1 to 9. All the terms answering “yes” would be scored accordingly, while those answering “no” would be scored as zero. The total QHES score of the literature would be obtained by summing up the scores of each term. QHES scale did not provide a score range to specifically distinguish high- and low-quality research, but to more clearly and intuitively describe the quality of included literature, this study used the common grading criteria adopted by several researchers^[9-11]. Quality assessment results were divided into four grades according to QHES scores: high quality (75 to 100), average quality (50 to 74), low quality (25 to 49), and very low quality (0 to 24).

Literature screening, data extraction and quality assessment were independently conducted and cross-checked by two researchers. If there were any disagreements, they were resolved through discussion or consensus with a third party.

1.4 Statistical analysis

Excel 2019 was used for data extraction and SPSS 27.0 software was used for statistical analysis. The results of systematic review were showed by frequency and percentage. The comparison of composition ratio between groups was performed by chi-square test. The quality evaluation results were represented by mean and standard deviation as well as the number and percentage of literature in each quality grade. The Mann-Whitney U test was used to compare the quality scores between groups. The test level was $\alpha = 0.05$.

2 Results

2.1 Literature inclusion results

A total of 6 201 papers were retrieved, including 5 665 in Chinese and 638 in English. 2 935 duplicate papers were excluded. After reading the titles and abstracts of the remaining papers, 2 866 irrelevant papers were excluded. With further reading of full text, another 61 papers were excluded, and 325 papers were finally included, including 296 Chinese ones and 29 English ones. The literature screening process is shown in Fig. 1.

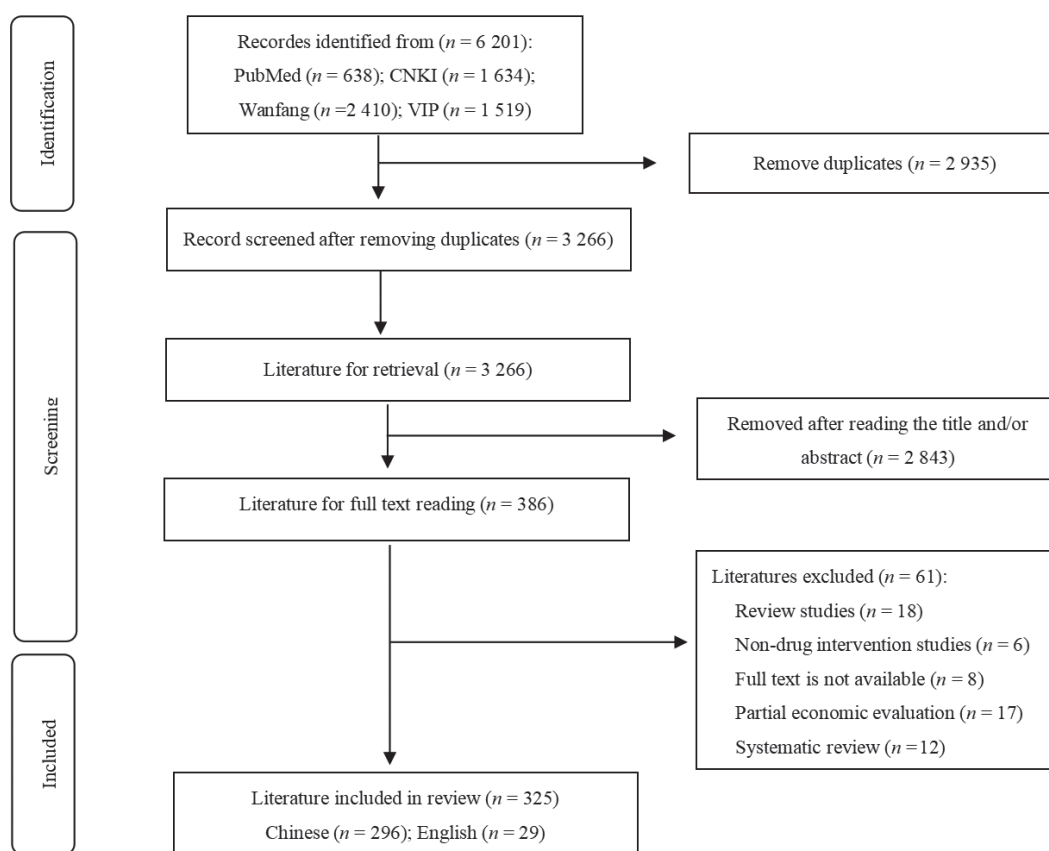


Fig. 1 Literature screening process for economic evaluation of drugs for T2DM

2.2 Publication information of literature

2.2.1 Publication time

Fig. 2 shows the number and dynamic trend of published economic evaluation studies on drugs for

T2DM based on Chinese population. The number of published papers on economic evaluation of drugs for T2DM had slight overall fluctuation in different periods. Before 2007, the papers increased slowly, and then started to increase significantly. The period from 2010 to 2016 witnessed the peak of related literature



publication, with 29 papers published in 2016. Then, the numbers began to decline after 2016. In 2018, the number of published papers dropped to the lowest in

nearly 10 years, with only 8 articles published. After 2018, the number of published papers began to show a fluctuating growth trend again.

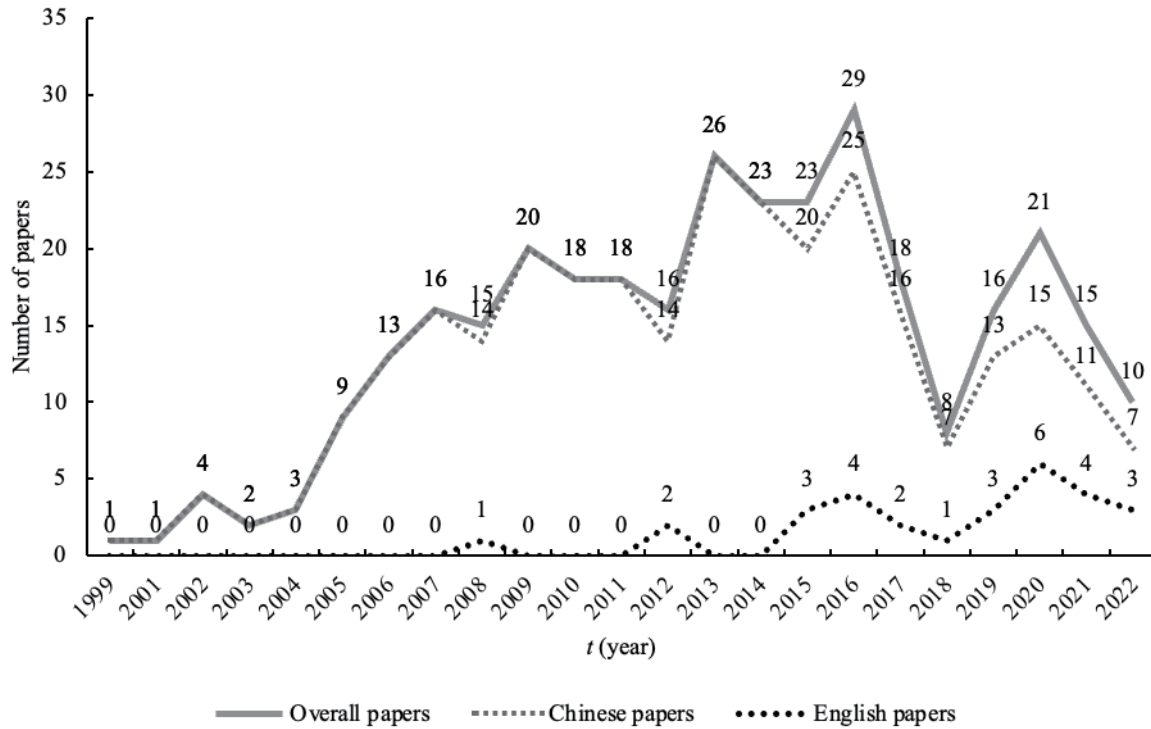


Fig. 2 The number of publications of literature by year

2.2.2 Author information

The 325 included articles involved 304 first authors. Among them, one author published a total of 6 papers, indicating that the relevant researchers on the economic evaluation of drugs for T2DM were relatively scattered and the research team was not concentrated.

Among the Chinese papers, there were 111 papers with only one author, accounting for 37.5%.

There were 232 papers with less than three authors, accounting for 77.7%. Besides, there were only 10 papers with more than seven authors, accounting for 3.4%. Among the English papers, the least number of authors was three, 22 papers with 4–6 authors, accounting for 75.9%, and 15 papers with more than 7 authors, accounting for 17.2%. The number of authors on economic evaluation of drugs for T2DM based on Chinese population are shown in Table 1.

Table 1 The number of authors of included literature

| Number of authors | Overall literature (%) | Chinese literature (%) | English literature (%) |
|-------------------|------------------------|------------------------|------------------------|
| 1 | 111 (34.2) | 111 (37.5) | 0 |
| 2–3 | 121 (37.2) | 119 (40.2) | 2 (6.9) |
| 4–6 | 78 (24.0) | 56 (18.9) | 22 (75.9) |
| > 7 | 15 (4.6) | 10 (3.4) | 5 (17.2) |



2.2.3 Source of research units

The unit of the first author listed in the literature was counted as the research unit, as shown in Table 2. The results displayed that the main research units for economic evaluation of drugs for T2DM were medical institutes. Among the included papers, 247 (76.0%) were carried out in hospitals, 58 (17.8%) were published by academic institutions, 13 (4.0%) were published by community-based public health organizations, and a small part of the

papers were written by pharmaceutical companies or pharmaceutical consulting companies (5, 1.5%) and independent scientific research institutions (2, 0.6%). In addition, the main research units for Chinese literature were hospitals (237 articles, 80.1%), while only 41 articles (13.9%) were published by universities and colleges. The majority of research units in English literature were universities and colleges (17, 58.6%), and there was a statistical difference in the composition ratio of research units in Chinese and English literature ($P < 0.001$).

Table 2 The type of research units of included literature

| Type of research unit | Overall literature (%) | Chinese literature (%) | English literature (%) |
|--|------------------------|------------------------|------------------------|
| Hospital institutions | 267 (82.2) | 267 (90.2) | 0 |
| Academic institutions | 58 (17.8) | 29 (9.8) | 29 (100) |
| Others | 18 (5.5) | 10 (3.4) | 8 (27.6) |
| Community-based health units | 19 (5.9) | 7 (2.4) | 12 (41.4) |
| Pharmaceutical/consulting companies | 12 (3.7) | 9 (3.0) | 3 (10.3) |
| Independent scientific research institutions | 8 (2.5) | 2 (0.7) | 6 (20.7) |
| Total | 1 (0.3) | 1 (0.3) | 0 |

2.2.4 Source of published journals

The names of published journals of all the included papers were statistically summarized, and it was found that the published journals of economic evaluation of drugs for T2DM were widely distributed. The 325 included articles covered 168 journals, including 149 Chinese journals and 19 English journals. In terms of the level of published journals, 133 articles (40.9%) were published in core journals, and the remaining 192 articles (59.1%) were published in non-core journals.

articles (34.8%) compared the economic efficiency of combined drug regimens with different drugs, 47 articles (14.5%) compared the economic efficiency of single drug regimens and combined drug regimens, 27 articles (8.3%) compared the economic efficiency of domestically produced and imported drugs with the same generic name, 22 articles (6.8%) evaluated the economic efficiency of different formulations of the same variety, and 7 articles (2.1%) compared the economic efficiency of insulin with different injection methods.

2.3 Information of literature

2.3.1 Intervention type

The literature was categorized according to the type of intervention studied in the literature. A total of 109 articles (33.5%) compared the economic efficiency of different single drug regimens, 113

2.3.2 Design and time horizon

The included literature mainly consisted of retrospective studies (256 articles) and literature-based studies (69 articles). Among the 69 articles using literature-based studies, 13 articles collected relevant literature or data on the drugs for T2DM, which formed a more general evaluation result by integrating the collected information. Modeling method was



used for the other 56 articles, including 29 Chinese articles (9.8%) and 27 English articles (93.1%). The most commonly used model was the CORE model (26 articles, 46.4%), followed by the Cardiff diabetes model (9 articles, 16.1%), Markov model (9 articles, 16.1%), UKPDS model (3 articles, 5.4%), and COMT model (3 articles, 5.4%).

In terms of time horizon, the 256 retrospective articles included a total of 48 articles (18.75%) that had no specific time of treatment intervention. Many of the papers that described intervention time only mentioned blood glucose control time, hospitalization time as the study time horizon. In some articles, the treatment time was not clear and only made statements such as “no less than a certain number of days”. Among all the literature reporting time horizon, the shortest intervention time was 7 days and the longest was 2 years (2 articles). The three most frequently used time horizon were 12 weeks (49 articles), 8 weeks (33 articles), and 3 months (21 articles). Among the 56 articles using modeling method, 6 articles conducted short-term drug economic evaluations

(less than 2 years), which were all Chinese papers. 47 articles were long-term economic evaluations (5 years or more), including 20 Chinese papers and 27 English papers. In addition, only one study conducted both short-term and long-term economic evaluations.

2.3.3 Study perspective

Of the 325 included articles, only 58 (17.8%) specified the study perspective, while the remaining 267 (90.2%) articles did not mention it. All the included English literature clearly indicated the study perspective, while only 29 Chinese literature (11.3%) reported it and the remaining 267 (90.2%) articles did not report study perspective, which caused inconvenience or misleading the literature users and decision-makers, and eventually affected the application and reference of evaluation results. There was a statistically significant difference in the composition ratio of the research perspective between Chinese and English literature ($P < 0.001$), as shown in Table 3.

Table 3 The study perspective of the included literature

| Type of study perspective | Overall literature (%) | Chinese literature (%) | English literature (%) | <i>P</i> |
|---|------------------------|------------------------|------------------------|----------|
| Literature did not report study perspective | 267 (82.2) | 267 (90.2) | 0 | < 0.001 |
| Literature report study perspective | 58 (17.8) | 29 (9.8) | 29 (100) | |
| Health care system perspective | 18 (5.5) | 10 (3.4) | 8 (27.6) | |
| Payer’s perspective | 19 (5.9) | 7 (2.4) | 12 (41.4) | |
| Societal perspective | 12 (3.7) | 9 (3.0) | 3 (10.3) | |
| Health care providers perspective | 8 (2.5) | 2 (0.7) | 6 (20.7) | |
| Patient’s perspective | 1 (0.3) | 1 (0.3) | 0 | |

2.3.4 Evaluation methods

Of the 325 included articles, CEA was the main method used in 295 articles (90.8%), followed by CUA in 16 articles (4.9%). 2 studies simultaneously used CEA and CUA. A total of 10 studies (3.1%) used the CMA, and 2 studies (0.6%) used CBA. It

can be seen that CEA was the main method used in economic evaluations of drugs for T2DM based on Chinese population. There was no statistically significant difference in the composition ratio of pharmacoeconomic evaluation methods between Chinese and English literature ($P > 0.05$), as shown in Table 4.



Table 4 Type of the evaluation methods of included literature

| Type of evaluation methods | Overall literature (%) | Chinese literature (%) | English literature (%) | <i>P</i> |
|----------------------------|------------------------|------------------------|------------------------|----------|
| CEA | 295 (90.8) | 272 (91.9) | 23 (79.3) | |
| CUA | 16 (4.9) | 11 (3.7) | 5 (17.2) | |
| CEA, CUA | 2 (0.6) | 2 (0.7) | 0 | > 0.05 |
| CMA | 10 (3.1) | 9 (3.0) | 1 (3.4) | |
| CBA | 2 (0.6) | 2 (0.7) | 0 | |
| Total | 325 (100) | 296 (100) | 29 (100) | |

2.3.5 Incremental analysis

Of the 325 included articles, over half of the studies conducted incremental analysis (178 articles, 54.8%), while the remaining literature did not. Among the studies that conducted incremental analysis, 161 articles (90.4%) used CEA, and 17 articles (9.6%) used CUA. Among the articles that used CEA, 23 articles (19.7%) reported the threshold and the reasons for selecting the threshold. Among the articles that used CUA, 5 articles (29.4%) did not report the threshold. In the literature that reported thresholds, most of them used 1–3 the per capita GDP as the threshold standard.

2.3.6 Sensitivity analysis

Among the 325 included articles, only 194 articles (59.7%) conducted uncertainty analysis, including 165 Chinese articles (55.7%) and all English articles (29 articles, 100%). There was a statistically significant difference in the composition ratio of uncertainty analysis between Chinese and English-language articles ($P < 0.001$), as shown in Table 5. Among the literature conducted sensitivity analysis, 162 articles (83.5%) chose the simplest sensitivity analysis method, with the majority of Chinese articles (156 articles, 94.6%) conducting only one-way sensitivity analysis, and the vast majority of English-language articles (20 articles) conducting both one-way and probability sensitivity analysis.

Table 5 Type of the sensitivity analysis of included literature

| Sensitivity analysis | Overall literature (%) | Chinese literature (%) | English literature (%) | <i>P</i> |
|---|------------------------|------------------------|------------------------|----------|
| Literature did not conduct sensitivity analysis | 131 (40.3) | 131 (44.3) | 0 | < 0.001 |
| Literature conduct sensitivity analysis | 194 (59.7) | 165 (55.7) | 29 (100) | |
| One-way, sensitivity analysis | 162 (83.5) | 156 (94.6) | 6 (20.7) | |
| One-way, probability sensitivity analysis | 25 (12.9) | 7 (4.2) | 18 (62.1) | |
| One-way, probability, case analysis | 6 (3.1) | 2 (1.2) | 4 (13.8) | |
| One-way, multi-way sensitivity analysis | 1 (0.5) | 0 | 1 (3.4) | |

2.4 Results of literature quality assessment

2.4.1 Quality of included literature

The quality assessment results showed that the average QHES score for the 325 included literature was 57.29, with a standard deviation of 18.39. The lowest-scoring literature received a score of 15,

which did not meet the other requirements except for “research objectives” and “research conclusions”. The highest-scoring literature received a score of 94. The quality of the included literature was concentrated in “low quality” (123 articles, 37.8%) and “average quality” (138 articles, 42.5%), indicating that the quality of economic evaluation literature on drugs for T2DM based on the Chinese population was relatively



low at present. Among them, almost all English literature was of high quality (27 articles, 93.1%). Only 30 articles (10.1%) of Chinese literature were considered “high quality”, while literature classified as

“very low quality” was all written in Chinese. There was a significant statistical difference in the quality grades between Chinese and English literature ($P < 0.001$), as shown in Table 6.

Table 6 Quality grade of included literature

| QHEs score | Quality grade | Overall literature (%) | Chinese literature (%) | English literature (%) | <i>P</i> |
|------------|---------------|------------------------|------------------------|------------------------|----------|
| 0–24 | Very low | 7 (2.2) | 7 (2.4) | 0 | < 0.001 |
| 25–49 | Low | 123 (37.8) | 123 (41.6) | 0 | |
| 50–74 | Average | 138 (42.5) | 136 (45.9) | 2 (6.9) | |
| 75–100 | High | 57 (17.5) | 30 (10.1) | 27 (93.1) | |

Table 7 displays the number of papers that meets the requirements of each item of QHEs scale. Most literature received scores in items such as “research objectives” “methodology stated” and “appropriate measurement of costs”, and over 80% of literature clearly explained their content. The description status of items such as “uncertainty handled” “incremental analysis” “time horizon” and “primary outcome stated” were moderate. The item “accurate conclusions and recommendations of the study” was

the only one that all 325 included papers explicitly stated. The number of papers that met the items “research perspective” and “best available source for variable estimates” was relatively small. There were also few papers that satisfy the items like “transparent descriptions of analysis” and “statement and justification of models and assumptions”, probably because there was not many literature that used model research methods. In addition, only 40 (12.3%) studies declared their funding sources.

Table 7 The compliance rate of QHEs terms of included literature

| QHEs term | Overall literature (%) | Chinese literature (%) | English literature (%) |
|--|------------------------|------------------------|------------------------|
| QHEs-1: Objective | 324 (99.7) | 294 (99.3) | 29 (100) |
| QHEs-2: Perspective | 55 (16.9) | 26 (8.8) | 29 (100) |
| QHEs-3: Best available source for variable estimates | 63 (19.4) | 59 (19.9) | 4 (13.8) |
| QHEs-4: Subgroup analysis described | 91 (28) | 64 (21.6) | 27 (93.1) |
| QHEs-5: Uncertainty handled | 194 (59.7) | 165 (55.7) | 29 (100) |
| QHEs-6: Incremental analysis | 177 (54.5) | 148 (50) | 29 (100) |
| QHEs-7: Methodology stated | 266 (81.8) | 238 (80.4) | 28 (96.6) |
| QHEs-8: Time horizon and discount | 186 (57.2) | 160 (54.1) | 26 (89.7) |
| QHEs-9: Appropriate measurement of costs | 270 (83.1) | 251 (84.8) | 19 (65.5) |
| QHEs-10: Primary outcome stated | 176 (54.2) | 151 (51) | 25 (65.5) |
| QHEs-11: Valid and reliable health outcomes measures | 174 (53.5) | 153 (51.7) | 21 (72.4) |
| QHEs-12: Transparent descriptions of analysis | 54 (16.6) | 27 (9.1) | 27 (93.1) |
| QHEs-13: Statement and justification of models and assumptions | 54 (16.6) | 27 (9.1) | 27 (93.1) |
| QHEs-14: Bias | 89 (27.7) | 65 (22) | 24 (99.3) |
| QHEs-15: Accurate conclusions and recommendations of the study | 325 (100) | 296 (100) | 29 (86.2) |
| QHEs-16: Funding | 40 (12.3) | 14 (4.7) | 26 (89.7) |

Note: The QHEs terms are simplified for the better presentation.



2.4.2 Comparative analysis of literature quality

A comparative analysis was carried out on the quality score of the included literature according to the language of publication, research units and study design.

The average QHES score of 296 Chinese literature was 54.06 (SD, 15.74), while for 29 English literature was 90.24 (SD, 8.19), which also indicated that the quality of English literature was significantly better than that of Chinese literature ($P < 0.001$), as shown in Table 8. The rate of compliance with items such as “research perspective” “subgroup analysis described” and “bias” in English literature was higher than that of Chinese literature by 50%, the reporting quality of this items of English literature was superior to that of Chinese literature. The reporting quality of both Chinese and English literature for items such as “best available source for variable estimates” and “appropriate measurement of costs” was close and the compliance rates were relatively low. One important reason for the generally higher quality of English literature may be that most English literature used model-based method, and some QHES items may be more applicable to model research study (such as QHES-12 and QHES-

13). In addition, English journals may have stricter standards for research quality than Chinese journals, and researchers naturally pay more attention to research quality and report in more standardization way in order to be successfully published.

As shown in Table 8, 56 articles conducted model research, and the QHES score of the model research method was higher than that of non-model research literature, with a mean value of 50.92 (SD, 12.28). There was a significant difference in quality scores between the two group of literature ($P < 0.001$). One reason for this might be that some QHES items were more suitable for model research methods. In addition, most of the model research studies came from universities (33 studies, 58.9%), with more comprehensive report content and more standardized reporting format than non-model research studies.

The systematic review results showed that the research units of the economic evaluation literature on drugs for T2DM included in this study were mainly hospitals. The quality of literature from universities was significantly better than that of hospitals and others ($P < 0.001$), as shown in Table 8.

Table 8 Comparative analysis of literatures quality

| Comparative analysis | Number of literature (%) | QHES score (mean) | QHES score (SD) | <i>P</i> |
|----------------------|--------------------------|-------------------|-----------------|----------|
| Publication language | | | | |
| Chinese | 296 | 54.06 | 15.74 | < 0.001 |
| English | 29 | 90.24 | 8.19 | |
| Design type | | | | |
| Using model | 56 | 50.92 | 12.28 | < 0.001 |
| Not using model | 269 | 87.85 | 10.37 | |
| Research units | | | | |
| Hospital institutes | 247 | 53.41 | 15.91 | < 0.001 |
| Academic institutes | 58 | 73.69 | 18.34 | |
| Others | 20 | 57.60 | 21.59 | |

3 Discussion

3.1 Deficient economic evaluation of drugs for T2DM

The systematic review results showed that the

included literature on economic evaluation of drugs for T2DM had no clear study perspectives (82.2%) and single evaluation methods (90.8% were CEA). Besides, most literature lacked sensitivity analysis (40.3%) or incremental analysis (45.2%). It can be



seen that the number of papers that did not report or misuse of relevant contents and methods accounted for a large number. This result was generally consistent with the previous studies which focused on the quality of economic evaluation literatures in different disease type or medications. Xu Ailin, et al.^[12] mainly investigated the literature about children's drugs. Feng Yiman, et al.^[13] studied the antitumor drugs, and Lin Jie, et al.^[14] focused on the traditional Chinese medicine (TCM). Their study results showed that the above key elements were missing to varying degrees in the literatures included, resulting in highly inconsistent research contents of the economic evaluation. Such incompleteness in content and misuse of methods not only affected the quality and reference value of a single economic evaluating study, but also led to the failure to integrate existing studies for comprehensive analysis, limiting the possibility of obtaining higher level evidence based on more data.

A scientific and reasonable economic evaluation needs to follow certain steps, including clarifying research questions and expected objectives, clarifying service objects, selecting alternative programs, determining appropriate evaluation indicators and methods, identifying and measuring costs and benefits, comparing output results, and carrying out uncertainty analysis^[6]. However, the results of this study showed that the included literature had not given sufficient consideration in the aspects of clarifying the research object, determining the appropriate evaluation method, and carrying out incremental analysis and sensitivity analysis. In the future, more attention should be paid to those key elements in the research design and formal implementation of economic evaluation. Each research element should be analyzed comprehensively, deeply and in detail as far as possible and the results should be reported and displayed in a standardized manner, which can ensure that the evaluation results are more scientific and objective, thus promoting the wider and deeper application of drug economic evaluation results.

3.2 Low quality of economic evaluation of drugs for T2DM

The mean QHES score of the 325 included

economic evaluation literature on drugs for T2DM based on the Chinese population was 57.29. Although there were some high-quality studies, but the overall study quality was low. This was similar to previous research results. Zhang Di, et al.^[11] assessed the quality of 29 studies on the economic evaluation of opioid analgesics for cancer pain in China and found that 28 studies had QHES scores below 75 points, indicating poor research quality. Sun Mao, et al.^[15] evaluated the quality of 132 published economic evaluation literature using the QHES scale in 2013, and found a poor quality of these studies with a QHES score of (62.90 ± 8.50) . Chen Chaoxin, et al.^[9] found that the QHES score of the economic evaluation of generic drugs that passed through the consistency evaluation was low, indicating further standardization and improvement of research quality. The comparison analysis of the quality of the included literature indicated that the overall quality of studies published in English journals was higher than that of Chinese literature, while studies using model method and those whose first authors were from universities had relatively higher quality.

Considering that the assessment conducted in this paper was only based on journal literature, instead of a complete economic evaluation report, some researchers may have considered relevant contents that did not appear in their papers due to journals requirements^[16], which finally resulted in lower quality. In addition, a total of 12 literature evaluating the economic efficiency of TCM for T2DM were included in this study. Due to TCM's different mechanisms of action compared to Western medicine, there were some differences in the research content and focus of economic evaluation of TCM and Western medicine. Therefore, the use of the QHES scale for quality assessment of economic evaluation literature on TCM drugs^[14] may not reflect the true quality of these papers, which may influence our results. However, relative researchers and scholars need to clarify and pay attention to key content in the economic evaluation of drugs, take measures to fully conduct all of the content of economic evaluation of drugs, and gradually improve overall study quality and level.



This paper has certain limitations because it only uses the QHES scale to evaluate the quality of included literature. Due to the limitations of the scale itself, combined with different understanding of reviewers, there is some subjectivity in the scoring results. However, this paper strictly follows the review mechanism to minimize subjectivity. Future research can use multiple evaluation scales to comprehensively evaluate the quality of economic evaluation study. In addition, there is significant heterogeneity in the included literature. Therefore, only the basic characteristics and quality of these papers are summarized, and further comparison of the economic evaluation of various drugs for type 2 diabetes is needed for future research.

References

- [1] GBD 2016 disease and injury incidence and prevalence collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: A systematic analysis for the Global Burden of Disease Study 2016 [J]. *Lancet*, 2017, 390: 1211-1259.
- [2] Sun H, Saeedi P, Karuranga S, et al. IDF diabetes atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045 [J]. *Diabetes Research Clinical Practice*, 2022, 183: 109119.
- [3] Chinese Diabetes Society. Guideline for the prevention and treatment of type 2 diabetes mellitus in China (2020 edition) [J]. *Chinese Journal of Diabetes Mellitus*, 2021, 13 (4): 315-409.
- [4] Liu Yahong, Chang Hongsheng, Zhu Yan, et al. Rational use of medication and progress in type 2 diabetes [J]. *Sichuan Medical Journal*, 2016, 37 (10): 1179-1183.
- [5] American Diabetes Association Professional Practice Committee. Pharmacologic approaches to glycemic treatment: Standards of medical care in diabetes-2022 [J]. *Diabetes Care*, 2022, 45 (Supplement 1): S125.
- [6] Sun Lihua. *Pharmacoeconomics* [M]. Beijing: China Medical Science Press, 2019.
- [7] Walley T, Haycox A. Pharmacoeconomics: Basic concepts and terminology [J]. *British Journal of Clinical Pharmacology*, 1997, 43: 343-348.
- [8] Tao Libo, Wang Fangxu. Closed-loop management mechanism of pharmacoeconomics in decision-making application [J]. *China Journal of Pharmaceutical Economics*, 2021, 16 (4): 31-33+44.
- [9] Chen Chaoxin, Zheng Bin, Zheng Huanrui, et al. Evaluation of systematic evaluation of economics of generic drugs through the quality and efficacy consistency of generic drugs [J]. *China Journal of Pharmaceutical Economics*, 2020, 15 (7): 10-19.
- [10] Chiou CF, Hay JW, Wallace JF, et al. Development and validation of a grading system for the quality of cost-effectiveness studies [J]. *Med Care*, 2003, 41 (1): 32-44.
- [11] Zhang Di, Zou Yuanjun, Jin Taowei, et al. Quality of economic evaluation of opioid drugs for cancer pain treatment in China [J]. *Chinese Journal of Gerontology*, 2021, 41 (10): 2048-2052.
- [12] Xu Ailin, Li Wenjun, Lu Mengqing. Quality appraisal of pediatric pharmacoeconomic research literatures in China from 2009 to 2018 [J]. *China Pharmacy*, 2020, 31 (6): 718-723.
- [13] Feng Yiman, Ke Xiatong, Tang Wenxi. Evaluation of the quality of Chinese literature in the pharmacoeconomic evaluation of antitumor drugs [J]. *China Journal of Pharmaceutical Economics*, 2021, 16 (4): 11-15.
- [14] Lin Jie, Wang Shuo, Guo Dongmei. Literature of Chinese herbal medicine economics research from 2014 to 2018: A quality assessment [J]. *Clinical Medication Journal*, 2019, 17 (11): 49-52.
- [15] Sun Mao, Feng Xiaoyuan, Gao Peng, et al. Quality evaluation of domestic pharmacoeconomics literatures in 2013 [J]. *China Pharmacy*, 2015, 26 (17): 2308-2311.
- [16] Yan Juntao, Bao Shiyi, Liu Liu, et al. Reporting quality of economic evaluations of the negotiated traditional Chinese medicines in national reimbursement drug list of China: A systematic review [J]. *Integrative Medicine Research*, 2023, 12 (1): 100915.