ARTICLE

Randomized controlled trials on erectile dysfunction: quality assessment and relevant clinical impact (2007–2018)

Kyu Shik Kim[1] · Jae Hoon Chung[1] · Seung Wook Lee[1]

Abstract

The aim of this study was to assess the quality of randomized controlled trials (RCTs) on erectile dysfunction (ED) conducted from 2007 to 2018. We searched for RCT original articles on ED published between 2007 and 2018 using PubMed, Embase, and Cochrane Library databases. RCT quality assessment was performed using Jadad scale, van Tulder scale, and Cochrane Collaboration's Risk of Bias Tool. The effects on RCT quality of including treatment methods, funding sources, institutional review board (IRB) approval statements, and intervention description to the studies were assessed. Blinding and allocation concealment were described in 67.9 and 8.7% of the RCTs, respectively. Blinding tended to decrease, but a sharp rise in blinding was observed in 2011–2012 and allocation in 2017–2018. Funding statement inclusion (60.3% overall) and intervention description (96.4% overall) tended to increase steadily. IRB statement inclusion (78.3% overall) increased ($p = 0.05$). Jadad scores rose significantly until 2011–2012 but decreased thereafter except 2017–2018 ($p = 0.09$). RCTs with funding statements had higher Jadad and van Tulder scores than unfunded RCTs ($p < 0.01$ and 0.02, respectively). Quality improvement has observed from 2007 to 2012 and 2017 to 2018 with Jadad scale because of increased funding, multicenter studies, and intervention description.

Introduction

Evidence-based medicine (EBM) has been increasing in importance and gaining popularity in the medical field. EBM entails the application of high-quality results to clinical practice using an integrated scientific process based on clinical evidence [1]. With the growing importance of EBM in modern medicine, the use of randomized controlled trials (RCTs) has become more significant even as the number of RCTs increases globally [2]. This is because RCT is the highest-rated study method of the many models as it minimizes study design bias [3]. Bias abounds if, from the design to the implementation of a study, the basic elements of RCT such as randomization, blinding, and allocation concealment are not properly performed [4]. To minimize mistakes and improve the quality of articles, it is necessary to objectively assess the methodological quality of articles [5]. Currently,

as a guideline for quality improvement of RCTs, the CONSORT statement is recommended by the International Committee of Medical Journal Editors [6–8]. The CONSORT statement guideline helps medical doctors conduct RCTs with minimal bias [9]. There is a limitation, however, that the CONSORT statement is merely a guideline for RCTs and not a tool for assessing the quality of articles. There are individual markers, checklists, and scales used for RCT methodological quality assessment. Of these, scales have the advantage of comparison ease in quantitative assessment of clinical trial quality between studies [7, 8].

Representative quality assessment tools using scoring systems include the Jadad scale, the van Tulder scale, and Cochrane Collaboration Risk of Bias tool (CCRBT). Jadad scale is a tool for assessing randomization, double blinding, and drop-out items related to bias reduction [10]. Although the Jadad scale has the advantage of being simple and easy to assess, it does not include the individual marker index of allocation concealment, making it difficult to assess selection bias in patient allocation for treatment. The van Tulder scale [11] and CCRBT [12] contain assessment items for allocation concealment with the advantage of selection bias assessment.

Erectile dysfunction (ED) refers to a state characterized by inability to develop or maintain proper erection of the

✉ Seung Wook Lee
  swleepark@hanyang.ac.kr

1 Department of Urology, Hanyang University College of Medicine, Seoul, Korea

penis for sexual intercourse [13]. The incidence and severity of ED increase with age. It occurs in 52% of men aged 40–70 with various degrees of severity (mild, moderate or severe), with an incidence of 8% in men in their 40 s to 15% in men in their 70 s, according to the Massachusetts Male Aging Study (MMAS) [14]. ED treatment varies from the administration of oral phosphodiesterase type 5 (PDE-5) inhibitors to implants, but the choice of treatment should consider factors such as patient age and general patient health. Many RCTs have published studies on the efficacy of ED treatments. There have heretofore been no studies analyzing the quality of RCTs on ED.

Thus, the purpose of this study is to assess the quality of RCTs on ED published in the last 10 years using 3 representative RCT quality assessment tools, namely Jadad scale, van Tulder scale, and CCRBT, and to propose directions for further studies.

Materials and methods

The subjects of analysis

The subjects of the analysis were RCT studies that were identified from searches in the PubMed, Embase, and Cochrane Library databases using the keyword "Erectile Dysfunction." RCT original articles on ED published from 2007 to 2018 and identified from the searches were selected.

The quality assessment was conducted after dividing the period under study into 6, 2-year periods, namely 2007–2008, 2009–2010, 2011–2012, 2013–2014, 2015–2016, and 2017–2018.

Selection of RCTs

Two reviewers independently searched for RCTs using PubMed. Then, to find RCTs that may have been missed, the reviewers modified their search by using the keywords "randomized," "randomization," "randomly," and "erectile dysfunction." The different search outcomes of extracted articles obtained by each reviewer were adjusted and fine-tuned by a third reviewer (Fig. 1).

Assessment method using quality assessment tools

The 2 reviewers analyzed the RCTs using the Jadad scale, the van Tulder scale, and CCRBT. If there were differences in data obtained by the 2 reviewers, the third reviewer modulated them.

Fig. 1 Flowchart of selected data
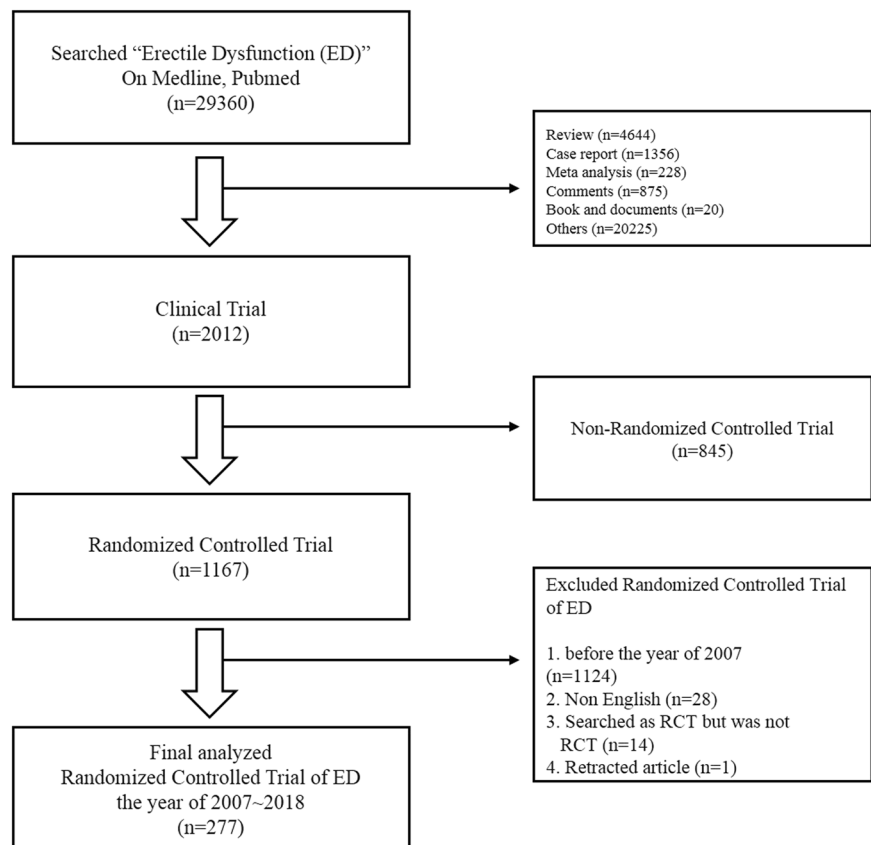
Jadad scale

The Jadad scale, also known as the Oxford quality scoring system, is a tool for assessing the quality of RCTs and it consists of a total of 3 graded questions. A maximum of 2 points can be awarded for a randomization question, 2 points for a blinding question, and one point for a drop-out question. For the randomization question, a point is awarded when the RCT article includes a description of randomization. An additional point is assigned if the article includes appropriate description on randomization, but 1 point is deducted for inappropriate description. Thus, Jadad scale distributes 0–2 points in total. For the blinding question, a point is assigned when double blinding is mentioned in the article. An additional point is assigned when appropriate blinding is included, but 1 point is deducted for inappropriate blinding. For the drop-out question, a point is awarded when a drop-out is mentioned in the article. The quality of the final RCT article is classified as low quality when 0–2 points are accrued, and high quality when 3–5 points are awarded in the assessment for the allocation of a possible maximum of 5 points [15].

van Tulder scale

Van Tulder scale is one of the most appropriate tools for RCT quality assessment. It includes 11 factors, namely: randomization, allocation concealment, baseline characteristics, patient blinding, care provider blinding, observer blinding, co-intervention, compliance, drop-out rate, end-point assessment time point, and intention-to-treat analysis. Each item is assessed by a response of "yes," "no" or "I do not know" and the RCT article is regarded as high quality if the score is ≥5 points [11].

Cochrane collaboration risk of bias tool (CCRBT)

The Cochrane assessment is divided into 6 domains: sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting, and other potential threats to validity. Each item is assessed by a response of "yes," "no" or "unclear." The reviewers judge and assess according to the detailed criteria for "yes," "no" or "unclear" responses for each of the 6 domains, which indicate low, high or uncertain risks of bias, respectively. If the responses for the first 3 domains are all "yes" and no important concerns are identified in relation to the last 3 domains, the RCT article is considered to have a low risk of bias. If the response in ≤2 domains is "unclear" or "no," the study is classified as having a moderate risk of bias. If the response in ≥3 domains is "unclear" or "no," the RCT is considered to have a high risk of bias [12].

Analysis of RCT quality according to other factors

In the RCT quality assessment in this study, the presence of intervention, funding, and institutional review board (IRB) approval were also taken into account.

Statistical analysis methods

The score of each assessment tool was compared and analyzed using one-way analysis of variance test, while chi-squared test was used to compare and analyze the ratio of the high-quality articles and the quality assessment outcomes from CCRBT. The student's t-test was used for comparison based on the presence or absence of IRB approval, funding, and intervention. All statistical analyses were performed using SPSS v.22.0, and a p-value < 0.05 was considered statistically significant.

Results

Quantitative changes of RCTs over time

From 2007 to 2018, there were a total of 277 RCT original articles related to ED. RCT studies were responsible for the increase in the number of articles (from 58 to 67 articles) from 2007 to 2010, with a subsequent decrease (from 48 to 30 articles) from the year 2015 to 2018 (Table 1).

Qualitative changes of RCTs over time

1. Jadad assessment scale: From 2007 to 2018, the mean score of the Jadad scale for the RCTs was 2.80 ± 1.24. In the period 2007–2008, the score was 2.60 ± 1.30 and 3.27 ± 1.31 in 2011–2012, meaning that the mean score increased. However, it decreased to 2.73 ± 1.22 in 2015–2016 ($p = 0.09$). Also, the number of high-quality articles was 30 (51.7%) in 2007–2008, 28 (75.7%) in 2011–2012, and 27 (56.3%) in the period 2015–2016 ($p = 0.06$). In 2017–2018, the score and the number of high-quality articles were increased compare to 2015–2016. The total number of high-quality articles was 174 (62.8%) (Table 1).
2. van Tulder assessment scale: The mean scores of the van Tulder scale at 2-year intervals for the periods 2007–2008, 2009–2010, 2011–2012, 2013–2014, 2015–2016, and 2017-2018 were 6.05 ± 1.48, 5.72 ± 1.47, 6.35 ± 1.57, 6.14 ± 1.57, 5.83 ± 1.66, and 6.33 ± 1.61, respectively ($p = 0.27$). The numbers of high-quality articles were 49 (84.5%), 54 (80.6%), 31 (83.8%), 30(81.1%), 37 (77.1%), and 24 (80.0%),

Table 1 Characteristics of RCTs by publication year with quality assessment of RCTs

	2007~2008	2009~2010	2011~2012	2013~2014	2015~2016	Total	*p* value
Original articles	58	67	37	37	24	223	<0.01
Blinding (%)	35(60.3%)	45(67.2%)	32(86.5%)	28(75.7%)	11(45.8%)	151(67.7%)	<0.01[a]
Concealment of allocation(%)	6(10.3%)	5(7.5%)	3(8.1%)	4(10.8%)	2(8.3%)	20(9.0%)	0.97[a]
Multicenter(%)	29(50.0%)	24(35.8%)	14(37.8%)	14(37.8%)	8(33.8%)	89(39.92%)	0.48[a]
Funding(%)	36(62.1%)	42(62.2%)	23(62.2%)	20(54.1%)	12(50.0%)	133(59.6%)	0.75[a]
IRB(%)	42(72.4%)	46(68.7%)	32(86.5%)	33(89.2%)	18(75.0%)	171(76.7%)	0.08[a]
Intervention(%)	58(100.0%)	63(94.0%)	36(97.3%)	35(94.6%)	23(95.8%)	215(96.4%)	0.45[a]
Jadad scale							
Score	2.60 ± 1.30	2.61 ± 1.04	3.27 ± 1.31	2.92 ± 1.28	2.50 ± 1.29	2.76 ± 1.24	0.04
High quality	30(51.7%)	44(65.7%)	28(75.7%)	24(64.9%)	11(45.8%)	137(61.4%)	0.07[a]
VTC							
Score	6.05 ± 1.48	5.72 ± 1.47	6.35 ± 1.57	6.14 ± 1.57	5.58 ± 1.95	5.96 ± 1.57	0.09
High quality	49(84.5%)	54(80.6%)	31(83.8%)	30(81.1%)	16(66.7%)	180(80.7%)	0.43[a]
CCRBT							
High risk(%)	35(60.3%)	45(67.2%)	17(45.9%)	22(59.5%)	16(66.7%)	135(60.5%)	0.60[a]
Moderate risk(%)	18(31.0%)	19(28.4%)	17(45.9%)	11(29.7%)	6(25.0%)	71(31.8%)	
Low risk(%)	5(8.6%)	3(4.5%)	3(8.1%)	4(10.8%)	2(8.3%)	17(7.6%)	

Analysis of variance (ANOVA) tests [a]chi-square tests were used

IRB institutional review board, *RCT* randomized controlled trial, *VTC* van Tulder scale, *CCRBT* Cochrane Collaboration's assessment of risk bias tool

respectively ($p = 0.95$) among the total 225 high-quality articles (81.2%).

3. CCRBT: There were 5 (8.6%), 3 (8.1%), 2 (4.2%), and 5(16.7%) articles with a low risk of bias for the periods 2007–2008, 2011–2012, 2015–2016 and 2017–2018, respectively based on the CCRBT ($p = 0.08$; Table 1).

Analysis of factors related to the quality of the articles

In the assessments of Jadad and van Tulder scales, the number of high-quality articles was significantly higher when there were funding, IRB approval, interventions, or single-center studies, and the comparison is as follows: funding (yes: 66.7%, $p < 0.01$; no: 36.4%, $p = 0.02$), IRB (yes: 83.3%, $p < 0.01$; no: 18.7%, $p = 0.01$), intervention (yes: 98.9%, $p < 0.01$; no: 1.3%, $p < 0.01$), and single-center studies (yes: 55.2%, $p < 0.01$; multicenter: 43.6%, $p = 0.01$). For CCRBT, the number of studies with low risk of bias was of higher statistical significance when there was IRB approval, funding, single-center, or interventions (Table 2).

Analysis of subject related to the quality of the articles

Many RCT studies on PDE-5 inhibitors such as sildenafil, tadalafil, vardenafil, and udenafil have been published (Table 3). In the analysis of the quality of RCTs according to the ED treatment drug, the Jadad scale scores were 2.73 ± 1.10, 2.88 ± 1.27, 3.63 ± 1.15, and 2.69 ± 1.23 for sildenafil, tadalafil, vardenafil, and others, respectively. The numbers of high-quality articles obtained using the Jadad scale for sildenafil, tadalafil, vardenafil, and others were 34 (69.4%), 27 (65.9%), 14 (87.5%) and 52 (55.9%), respectively ($p < 0.01$). The van Tulder scores for sildenafil, tadalafil, vardenafil, and others were 5.96 ± 1.35, 6.27 ± 1.47, 6.81 ± 1.47, and 5.74 ± 1.52, respectively. There were also 43 (87.8%), 36 (87.8%), 16 (100.0%), and 70 (75.3%) high-quality articles for sildenafil, tadalafil, vardenafil, and others, respectively as assessed using the van Tulden scale ($p = 0.04$). The numbers of RCTs assessed by CCBRT to have a low risk of bias for sildenafil, tadalafil, vardenafil, and others were 3 (6.1%), 3 (7.3%), 1 (6.3%) and 9 (9.7%) respectively ($p = 0.10$; Table 3).

Distribution of RCTs of ED with other diseases and Journals

The frequency of ED as a topic was reported when it appeared as a single subject, associated with diabetes, benign prostatic hyperplasia or other diseases, and prostate cancer (Fig. 2). The distribution of journals with publications on RCTs related to ED include The Journal of Sexual Medicine, International Journal of Impotence Research, British Journal of Urology International, and Asian Journal of Andrology (Fig. 3).

Table 2 Factors associated with the quality of RCTs

Factors	No. of RCTs(%)	Jadad scale		Van Tulder scale		Cochrane's assessment of risk bias		
		Score	High quality	Score	High quality	High risk(%)	Moderate risk(%)	Low risk(%)
Funding source								
Yes	133(59.6%)	2.99 ± 1.25	92(69.2%)	6.32 ± 1.57	113(85.0%)	13(9.8%)	50(37.6%)	70(52.6%)
No	90(40.4%)	2.41 ± 1.14	45(50.0%)	5.43 ± 1.42	67(74.4%)	4(4.4%)	21(23.3%)	65(72.2%)
p value		0.68	<0.01[a]	0.25	0.05[a]	0.01[a]		
Reviewed by IRB								
Yes	171(76.7%)	2.96 ± 1.21	114(66.7%)	6.23 ± 1.56	144(84.2%)	15(8.8%)	67(39.2%)	89(52.0%)
No	52(23.3%)	2.10 ± 1.07	23(44.2%)	5.10 ± 1.26	36(69.2%)	2(3.8%)	4(7.7%)	46(88.5%)
p value		0.91	<0.01[a]	<0.01	0.02[a]	<0.01[a]		
Intervention								
Yes	215(96.4%)	2.81 ± 1.22	136(96.4%)	6.05 ± 1.52	179(83.3%)	17(7.9%)	71(33.0%)	127(59.1%)
No	8(3.6%)	1.25 ± 0.71	1(12.5%)	3.75 ± 1.17	1(12.5%)	0(0.0%)	0(0.0%)	8(5.9%)
p value		0.02	<0.01[a]	0.2	<0.01[a]	0.07[a]		
Multicenter								
Yes	89(39.9%)	3.13 ± 1.27	64(71.9%)	6.58 ± 1.52	80(89.9%)	8(9.0%)	43(48.3%)	38(42.7%)
No (single center)	134(60.1%)	2.51 ± 1.16	73(54.5%)	5.55 ± 1.46	100(74.6%)	9(6.7%)	28(20.9%)	97(72.4%)
p value		0.66	<0.01[a]	.055	<0.01[a]	<0.01[a]		

Student's t-test and [a]chi-square test were used

IRB institutional review board, *RCT* randomized clinical trial

Table 3 Assessment of RCT quality according to subject of erectile dysfunction research

| Subject	Jadad scale		Van Tulder scale		Cochrane's assessment of risk bias		
	Score	High quality	Score	High quality	High risk(%)	Moderate risk(%)	Low risk(%)
Sildenafil($n = 49$)	2.73 ± 1.10	34(69.4%)	5.96 ± 1.35	43(87.8%)	31(63.3%)	15(30.6%)	3(6.1%)
Tadalafil($n = 41$)	2.88 ± 1.27	27(65.9%)	6.27 ± 1.47	36(87.8%)	21(51.2%)	17(41.5%)	3(7.3%)
Udenafil($n = 11$)	3.36 ± 1.21	9(81.8%)	6.91 ± 1.51	10(90.9%)	4(36.4%)	6(54.5%)	1(9.1%)
Vardenafil($n = 16$)	3.63 ± 1.15	14(87.5%)	6.81 ± 1.47	16(100.0%)	5(31.3%)	10(62.5%)	1(6.3%)
Avanafil($n = 8$)	3.75 ± 0.71	8(100%)	7.00 ± 1.20	8(100.0%)	2(25.0%)	6(75.0%)	0(0.0%)
Mirodenalfil($n = 4$)	4.00 ± 0.82	4(100%)	8.00 ± 1.16	4(100.0%)	1(25.0%)	2(50.0%)	1(25.0%)
Sildenafil with other($n = 17$)	2.00 ± 1.28	4(23.5%)	5.41 ± 1.42	13(76.5%)	13(76.5%)	3(17.6%)	1(5.9%)
Tadalafil with other($n = 9$)	3.00 ± 1.32	7(77.8%)	5.89 ± 1.69	7(77.8%)	6(66.7%)	3(33.3%)	0(0.0%)
Vardenafil with other($n = 4$)	2.75 ± 1.26	3(75.0%)	6.00 ± 1.63	3(75.0%)	2(50.0%)	1(25.0%)	1(25.0%)
ESWL($n = 8$)	3.13 ± 1.46	6(75.0%)	6.00 ± 2.45	6(75.0%)	2(25.0%)	4(50.0%)	2(25.0%)
PDE 5i's($n = 12$)	1.92 ± 1.00	3(25.0%)	5.08 ± 1.51	6(50.0%)	10(83.3%)	2(16.7%)	0(0.0%)
Exercise($n = 5$)	2.40 ± 1.34	3(60.0%)	5.20 ± 1.92	3(60.0%)	4(80.0%)	1(20.0%)	0(0.0%)
Surgery($n = 9$)	2.22 ± 0.83	4(44.4%)	5.22 ± 0.83	7(77.8%)	8(88.9%)	1(11.1%)	0(0.0%)
Others($n = 84$)	2.70 ± 1.26	48(57.1%)	5.76 ± 1.55	62(89.9%)	46(54.8%)	28(33.3%)	10(11.9%)
p-value	<0.01	<0.01[a]	<0.01	0.04	0.10[a]		
Total	2.80 ± 1.24	274(62.8%)	6.01 ± 1.55	225(81.2%)	157(56.7%)	98(35.4%)	22(7.9%)

Analysis of variance (ANOVA) and [a]chi-square test were used

RCT randomized clinical trial

Discussion

In the quality assessment of RCTs conducted on ED as a subject, it was observed that the quality steadily improved over time from 2007 to 2012 and from 2017 to 2018, except for the assessment using CCRBT. However, the number of high-quality RCTs declined from 2012 to 2016. The quality of the RCTs was higher when there were IRB approval,
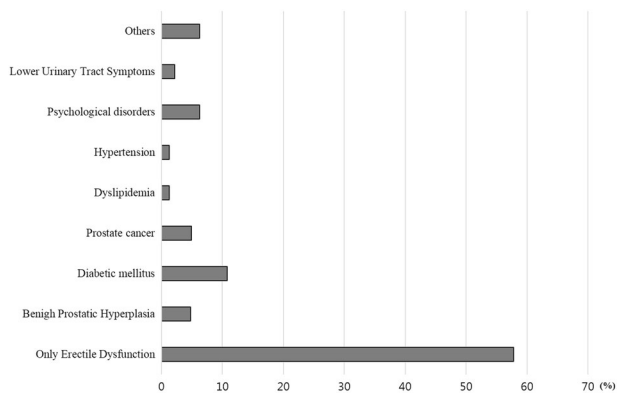
Fig. 2 Distribution of randomized controlled trials on erectile dysfunction with other diseases
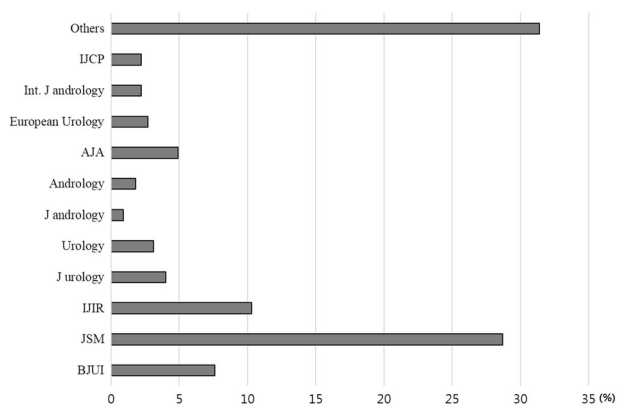
Fig. 3 Distribution of journals publishing randomized controlled trials on erectile dysfunction. *BJUI* British Journal of Urology International, *JSM* Journal of Sexual Medicine, *IJIR* International Journal of Impotence Research, *J Urology*: Journal of Urology, *J Andrology*: Journal of Andrology, *AJA* Asian Journal of Andrology, *Int. J Andrology* International Journal of Andrology, *IJCP* International Journal of Clinical Practice

funding, interventions, and single-center studies. Most of the RCT studies on ED was associated with PDE-5 inhibitors, and published in the Journal of Sexual Medicine. This study found that the number of RCTs on ED has been constantly increasing as evidenced by recent publications in many international journals. Scales et al. reported that the number and percentage of RCTs have increased over time because of a comparison of RCTs published in journals such as The Journal of Urology, Urology, European Urology, and BJU International from 1996 to 2004 [16]. Lee et al. reported an increase in the number of RCTs published by Korean Journal of Urology over the past 20 years [17]. This increase in RCTs is considered the reason for the growing importance of EBM.

As the value of EBM to clinical medicine increases, a systematic and scientific approach to RCT is emphasized, with RCT quality also taking on greater importance [18]. Studies on EBM and their findings have been consistently

collected in databases by Cochrane Collaboration with the help of clinical epidemiologists working around the world. Data compiled by meta-analyzing and collecting RCTs in health and medicine constitute the Cochrane Library, which is published through various media forms and widely used as a guideline for medical doctors around the world [19]. In recent years, journals require authors to ensure that RCT studies are conducted in accordance with CONSORT statement, by completing tasks in a checklist, before submission for publication. However, it is hard to quantitatively assess RCT quality using the CONSORT statement as the CONSORT statement has no weight on items and is an unspecified guideline for high quality. RCT quality assessment is important for evaluating the bias occurring in the study process, the validity of study conclusions, and the necessity of further studies. Thus, RCT quality assessment should be conducted [20, 21].

There are many tools for quantitatively assessing RCT quality such as Campell, Moher, Chalmers, Jadad, van Tulder, Newell's, and Cochrane [5]. In this study, we used 3 tools, namely Jadad scale, van Tulder scale, and CCRBT, to comprehensively analyze the various elements of the CONSORT statement for RCT quality assessment. RCT quality assessment using these tools has recently been conducted in several places. Kim Lee et al. analyzed RCTs published in several urology journals (International Journal of Impotence Research, Journal of Endourology, neurourology and urodynamics) [22–24]. According to the study, an increase was observed in the mean Jadad scale scores of the RCTs and in the number of high-quality articles over time. Lee et al. analyzed the RCTs published in the journal Neurourology and Urodynamics from 1993 to 2012 using Jadad scale, Van Tulder scale, and CCRBT and concluded that the number of RCTs increased but the quality did not improve [24]. In our study, the overall quality improved over time, but no qualitative improvement was observed during certain periods. This revealed a qualitative decline caused by a decrease in blinding. Therefore, there is a need to complement these points in future ED studies.

In this study, the quality differences based on IRB approval were also investigated. Bridoux et al. concluded that high-quality RCTs were approved by IRB [25]. The IRB review is a step that is recognized for the validity of design and implementation in the research planning phase of a study, and is considered an international standard. In recent RCTs, a valid study plan for obtaining approval at the IRB review plays a key role in improving the rate of high-quality articles. Schulz et al. reported that if the concealment of allocation was not performed properly, randomization in clinical research may be compromised, and even with initial randomization, the effect of intervention could be distorted by more than 40% [26]. Hewitt et al. assessed the quality of RCTs published in the New England

Journal of Medicine, BMJ, JAMA, and Lancet, and reported that 46% of the studies were conducted with uncertain concealment of allocation [27]. In this study, the percentage of RCT studies on ED that correctly conducted concealment of allocation was low. Concealment of allocation is a major component of improving the quality of study. Therefore, it is important to ensure that concealment of allocation in further studies is properly implemented. Clifford et al. analyzed 100 RCTs published in five different peer-reviewed, general medical journals with high impact factor and concluded that there was no correlation between funding sources and quality of articles [28]. However, Lee et al. reported that financially supported RCTs had many high-quality articles because they made well-designed and large-scale studies possible [17]. In this study, the quality assessment scores of financially supported articles evaluated using Jadad and van Tulder scales were higher than those of articles which were not financially supported. In the quality assessment based on the implementation of interventions, the number of high-quality studies, evaluated with Jadad and van Tulder scales, was greater when interventions were conducted. Assessment using CCRBT also revealed a higher risk ratio in RCTs with interventions. Computer generated randomization, allocated concealment, and double blinding are performed to reduce the risk of bias in a large number of RCTs with interventions [29]. For this reason, RCT studies with intervention are considered of high quality due to the increased use of objective methods. In the RCTs on ED, the topics most studied were PDE-5 inhibitor-related, and the quality assessment was relatively high. The PDE-5 inhibitor, sildenafil, was first developed and later, tadalafil, vardenafil, and others were developed [30]. During this process, a large number of RCT studies were conducted in connection with various diseases such as benign prostatic hyperplasia and pulmonary hypertension [30]. The quality of PDE-5 inhibitor-related RCTs was the highest of the RCTs on treatment.

One limitation of this study is that the subjective opinion of the reviewers may play a part in the assessment process due to the manual nature of the research and evaluation. To minimize this limitation, 2 reviewers participated independently in RCT sampling and quality assessment. When there were differences in the results, a third reviewer was consulted for correction to ensure the objectivity and reliability of the study. Another limitation of this study is that there are no formal representative tools among the currently available quality assessment tools in RCT quality analysis for assessing all the items represented in the CONSORT statement. However, we tackled this limitation by using 3 of the most widely used representative tools for RCT quality assessment. This study is of great significance because it shows the quantitative and qualitative changes of RCTs on ED over time, and suggests ways to improve the quality of further studies by analyzing factors affecting the quality of RCTs.

Conclusion

The number of RCT original articles on ED published increase from 2007 to 2018. However, it starts to decrease from 2013 and no significant increase in the quality of RCTs is observed from 2013. The number of high-quality articles increased when IRB approval is granted, when funding is provided, or when RCTs involved interventions or single-center studies. Researchers will do well to focus efforts on conducting high-quality studies.

Compliance with ethical standards

References

1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312:71–2.
2. Gluud C, Nikolova D. Likely country of origin in publications on randomised controlled trials and controlled clinical trials during the last 60 years. Trials. 2007;8:7.
3. Uetani K, Nakayama T, Ikai H, Yonemoto N, Moher D. Quality of reports on randomized controlled trials conducted in Japan: evaluation of adherence to the CONSORT statement. Intern Med. 2009;48:307–13.
4. Featherstone K, Donovan JL. Random allocation or allocation at random? Patients' perspectives of participation in a randomised controlled trial. BMJ . 1998;317:1177–80.
5. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA. 1996;276:637–9.
6. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg. 2012;10:28–55.
7. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ. 2001;323:42–6.
8. Hill J, Bullock I, alderson P. A summary of the methods that the national clinical guideline centre uses to produce clinical guidelines for the national Institute for Health and clinical excellence. Ann Intern Med. 2000;154:752–7.
9. Gagnier J, Boon H, Rochon P, Barnes J, Moher D, Bombardier C, CONSORT Group. Improving the quality of reporting of randomized controlled trials evaluating herbal interventions: implementing the CONSORT statement [corrected]. Explore. 2006;2:143–9.
10. Clark HD, Wells GA, Huët C, McAlister FA, Salmi LR, Fergusson D, et al. Assessing the quality of randomized trials: reliability of the Jadad scale. Control Clin Trials. 1999;20:448–52.

11. van Tulder M, Furlan A, Bombardier C, Bouter L. Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. Spine. 2003;28:1290–9.

12. Higgins JP, Green S (eds). Cochrane handbook for systematic reviews of interventions. Ver. 5.1.0. The Cochrane Collaboration; 2011. Available at: http://www.cochrane-handbook.org. Accessed 7 Jan 2011.

13. NIH Consensus Conference. Impotence. NIH consensus development panel on impotence. JAMA. 1993;270:83–90.

14. Feldman HA, Goldstein I, Hatzichristou DG, Krane RJ, McKinlay JB. Impotence and its medical and psychosocial correlates: results of the Massachusetts Male Aging Study. J Urol. 1994;151:54–61.

15. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin trials. 1996;17:1–12.

16. Scales CD Jr, Norris RD, Keitz SA, Peterson BL, Preminger GM, Vieweg J, et al. A critical assessment of the quality of reporting of randomized, controlled trials in the urology literature. J Urol. 2007;177:1090–4.

17. Lee JY, Chung JH, Kang DH, Lee JW, Moon HS, Yoo TK, et al. Quality assessment of randomized controlled trials published in the korean journal of urology over the past 20 years. Korean J Urol. 2011;52:642–6.

18. Sutton AJ, Cooper NJ, Jones DR. Evidence synthesis as the key to more coherent and efficient research. BMC Med Res Methodol. 2009;9:29.

19. Gill P, Dowell AC, Neal RD, Smith N, Heywood P, Wilson AE. Evidence based general practice: a retrospective study of interventions in one training practice. BMJ . 1996;312:819–21.

20. Chalmers TC, Celano P, Sacks HS, Smith H Jr. Bias in treatment assignment in controlled clinical trials. N Engl J Med. 1983;309:1358–61.

21. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA. 1995;273:408–12.

22. Kim KS, Jo JK, Chung JH, Kim JH, Choi HY, Lee SW. Quality analysis of randomized controlled trials in the International Journal of Impotence Research: quality assessment and relevant clinical impact. Int J Impot Res. 2017;29:65–69.

23. Jo JK, Autorino R, Chung JH, Kim KS, Lee JW, Baek EJ, et al. Randomized controlled trials in endourology: a quality assessment. J Endourol. 2013;27:1055–60.

24. Lee JW, Chung JH, Jo JK, Lee SW. Assessing the quality of randomized controlled trials published in neurourology and urodynamics from 1993 to 2012. Neurourol Urodyn. 2014;33:472–4.

25. Bridoux V, Moutel G, Roman H, Kianifard B, Michot F, Herve C, et al. Methodological and ethical quality of randomized controlled clinical trials in gastrointestinal surgery. J Gastrointest Surg. 2012;16:1758–67.

26. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. Lancet. 2002;359:614–8.

27. Hewitt C, Hahn S, Torgerson DJ, Watson J, Bland JM. Adequacy and reporting of allocation concealment: review of recent trials published in four general medical journals. BMJ. 2005;330:1057–8.

28. Clifford TJ, Barrowman NJ, Moher D. Funding source, trial outcome and reporting quality: are they related? Results of a pilot study. BMC Health Serv Res. 2002;2:18.

29. Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. Bmj. 2005;330:753.

30. Hong JH, Kwon YS, Kim IY. Pharmacodynamics, pharmacokinetics and clinical efficacy of phosphodiesterase-5 inhibitors. Expert Opin Drug Metab Toxicol. 2017;13:183–92.