

Assessing the Quality of Randomized Controlled Trials Published in the Journal of Korean Medical Science from 1986 to 2011

Jae Hoon Chung, Dong Hyuk Kang,
Jung Ki Jo, and Seung Wook Lee

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

Received: 8 March 2012

Accepted: 7 June 2012

Address for Correspondence:

Seung Wook Lee, MD

Department of Urology, Hanyang University Guri Hospital,

153 Gyeongchun-ro, Guri 471-701, Korea

Tel: +82.31-560-2374, Fax: +82.31-560-2372

E-mail: swleepark@hanyang.ac.kr

Low quality clinical trials have a possibility to have errors in the process of deriving the results and therefore distort the study. Quality assessment of clinical trial is necessary in order to prevent any clinical application erroneous results is important. Randomized controlled trial (RCT) is a design for evaluate the effectiveness of medical procedure. This study was conducted by extracting the RCTs from the original articles published in the *Journal of Korean Medical Science* (JKMS) from 1986 to 2011 and conducting a qualitative analysis using three types of analysis tools: Jadad scale, van Tulder scale and Cochrane Collaboration risk of bias Tool. To compare the quality of articles of JKMS, quality analysis of the RCTs published in *Yonsei Medical Journal* (YMJ) and *Korean Journal of Internal Medicine* was also conducted. In the JKMS, YMJ and Korean Journal of Internal Medicine, the quantitative increase of RCT presented over time was observed but no qualitative improvement of RCT was observed over time. From the results of this study, it is required for the researchers to plan for and perform higher quality studies.

Key Words: Randomized Controlled Trial; Journal Article

INTRODUCTION

Randomized controlled trials (RCTs) have increasingly become acknowledged as important standards of evidence-based medicine. RCTs follow a study design that can reduce bias and produce the most valuable data among study methods, making them the most reliable assessments of the effectiveness of medical treatments (1). However, even the RCT study design cannot eliminate all bias, which can occur at the designing, conducting, reporting, or application phase and can lead to derivation of incorrect results (2). Peer review within journals is an indispensable prevention method against error and can assess the validity of studies so as to prevent incorrect information being used as the basis of clinical application (3). Moreover, objective assessment of articles on methodological quality is essential because it can heighten the quality of medical care (4). An article's methodological quality can represent its overall quality and should therefore be assessed at the design, conduct, and analysis levels (5, 6). In addition, throughout the complete assessment process, any unnecessary or erroneous data can be identified, thereby eliminating it from clinical relevance and saving medical expenses (7).

There are several methods to assess the methodological quality of clinical trials, including scales, individual markers, and checklists. The scales method allows for easy inter-study comparison and is more advantageous than other methods when

performing quantitative assessment on the quality of a clinical trial. Randomization, double blinding, and drop out are the three factors of scale that directly relate to reducing bias. The Jadad quality assessment scale (Jadad scale) is a representative quality assessment tool consisting of these three items (8). The Jadad scale has been widely used because of its simple assessment questions and capacity to make assessment easy, but it does not include an assessment item for allocation concealment. There is however an individual marker method that assesses allocation concealment, which is a way to randomize the allocation sequence to evade any selection bias in the allocation of patients for treatment (9). The van Tulder scale and Cochrane collaboration risk of bias tool (CCRB) include an assessment item for allocation concealment. Recent studies have examined issues in quality assessment in RCTs. For example, Kim et al. analyzed all RCTs published in five Korean medical journals (10). The *Journal of Korean Medical Science* (JKMS) is the flagship journal of the Korea Academy of Medical Sciences and is the only science citation index Medical Journal in Korea. It was launched in 1986 and includes evidence-based, scientifically written articles aimed at introducing Korean medical sciences to the world and facilitating international medical information exchanges. However, there has been no quality assessment of RCTs published in the JKMS. *Yonsei Medical Journal* (YMJ) is a clinical science citation index expanded journal which has been published since 1960 by the Yonsei University College of Medicine and it covers all

the subjects related to medicine based on either clinical or basic research. *Korean Journal of Internal Medicine* is an international medical journal published by the Korean Association of Internal Medicine. To suggest the direction of future studies and to improve medical practice in Korea, this study assessed the quality of RCTs presented in the JKMS, YMJ and *Korean Journal of Internal Medicine* using the Jadad scale, van Tulder scale, and CCRBT.

MATERIALS AND METHODS

Study cohort

A total of 2,257 original articles published in the JKMS over 26 yr from 1986 (volume 1) to 2011 (volume 27) were manually searched. A total of 2,117 and 626 original articles published in YMJ and *Korean Journal of Internal Medicine* were also searched respectively.

Selection of RCTs

Two reviewers independently determined all RCT reports published in the JKMS, YMJ, *Korean Journal of Internal Medicine* using PubMed MEDLINE database and KoreaMed. They used search limits and searched terms such as “random,” “randomized,” and “randomly” in the methods sections of these reports. The other reviewer made a final selection by adjusting the data collected by the two reviewers.

Assessment of the quality of RCTs

Quality assessment was conducted using the Jadad scale; the van Tulder scale and CCRBT were used as individual indices. All assessments were performed by two specialists in Urology, and if there were different outcomes, they adjusted the discrepancy in the results through discussion. Starting from 1986, the quality analysis of RCTs was conducted in 5-yr units. Moreover, the quality assessment was conducted by type of intervention, presence of funding and reviewed by institutional review board (IRB).

Jadad scale

The Jadad scale is also known as the Oxford quality scoring system and assesses RCT-related literature. It is composed of five points in total; two in relation to randomization, two in relation to blinding, and one in relation to the drop out rate (8). When the report includes only general comments with no detailed description of randomization and blinding, one point in each category is given. One point is added when there is a detailed description of the appropriate method. However, when the description method is inappropriate, one point is deducted. When the specified number and reasons for drop outs by each subject group are provided, one point is given. Even if there are no drop outs, this should be described in the statement. When the total is

≥ 3 points, it is assessed as high quality but when it is ≤ 2 points, it is assessed as low quality. However, if it was not possible for the design of the study to be double blinded, it is assessed as high quality when the total score is ≥ 2 points.

van Tulder scale

The van Tulder scale is designed to make assessments on 11 components including randomization, allocation concealment, baseline characteristics, patient blinding, caregiver blinding, observer blinding, co-intervention, compliance, drop out rate, end-point assessment time point, and intention-to-treat analysis (11). Its assessment method is to select ‘yes’, ‘no’, or ‘don’t know’ for each item, and when ≥ 5 items are satisfied (≥ 5 points), the quality of the report is deemed high.

CCRBT

The CCRBT assesses the quality of RCTs in six classifications: sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting, and other potential threats to validity. The assessment indicates ‘yes’, ‘no’, or ‘unclear’ for each domain, designating low, high, and unclear risk of bias, respectively. In cases where the first three questions are answered with ‘yes’ and when no important concerns related to the last three domains are identified, it is classified as having a low risk of bias, while cases where it is assessed in ≤ 2 domains with ‘unclear’ or ‘no’, it is classified as having a moderate risk of bias. The cases assessed in ≥ 3 domains with ‘unclear’ or ‘no’ as classified as having a high risk of bias.

Statistical analysis

The one way-ANOVA and Kruskal Wallis test was used to compare and analyze the respective scores obtained by each assessment tool, while a chi-square test was used to compare and analyze the ratio of the high quality articles and the quality assessment outcomes from CCRBT. The quality assessment of RCTs according to publication year was analyzed by one-way ANOVA. SPSS v.18.0 was used for all statistical analyses and a P value of < 0.05 was considered statistically significant.

RESULTS

Quantitative variation of RCTs over time

From 1986 to 2011, 44 RCT-based articles were published among 2,275 original articles in JKMS. Among them, no articles were published from 1986-1989, three articles (6.82%) from 1991-1995, two articles (4.55%) from 1996-1997, seven articles (15.91%) from 2001-2005, 26 articles (59.09%) from 2006-2010, and six articles (2.96%) in 2011 ($P = 0.025$). The number of RCT in YMJ increased over time ($P < 0.001$). However, the quantitative increase of RCT in *Korean Journal of Internal Medicine* was not statistically significant ($P = 0.276$; Table 1).

Table 1. Characteristics of RCTs according to publication year

Journals	Years	Original articles	RCT (%)	Intervention: drug (%)	Double blinding (%)	Concealment of allocation (%)
JKMS	1986-1990	80	0	0	0	0
	1991-1995	219	3 (1.37)	3 (100)	1 (33.33)	0
	1996-2000	324	2 (0.62)	1 (50.00)	0	0
	2001-2005	582	7 (1.20)	5 (71.42)	1 (14.29)	0
	2006-2010	867	26 (3.00)	18 (69.23)	10 (38.46)	6 (23.08)
	2011	203	6 (2.96)	5 (83.33)	4 (66.67)	2 (33.33)
	<i>P</i> value		0.025	0.708	0.281	0.404
	Total	2,275	44 (1.93)	32 (72.73)	16 (36.36)	8 (18.18)
YMJ	1960-1990	545	0	0	0	0
	1991-1995	201	2 (1.00)	1 (50.00)	1 (50.00)	0
	1996-2000	318	6 (1.89)	2 (33.33)	0	0
	2001-2005	456	17 (3.73)	6 (35.29)	2 (11.76)	0
	2006-2010	492	20 (4.07)	13 (65.00)	9 (45.00)	0
	2011	105	8 (7.62)	6 (75.00)	3 (37.50)	0
	<i>P</i> value		< 0.001	0.216	0.084	-
	Total	2,117	53 (2.50)	28 (52.83)	15 (28.30)	0
Korean J Intern Med	1986-1990	128	1 (0.78)	1 (100)	0	0
	1991-1995	64	4 (6.25)	4 (100)	0	0
	1996-2000	100	4 (4.00)	4 (100)	0	0
	2001-2005	149	5 (3.36)	4 (80.00)	0	0
	2006-2010	151	3 (1.99)	1 (33.33)	0	0
	2011	34	2 (5.88)	1 (50.00)	1 (50.00)	1 (50.00)
	<i>P</i> value		0.276	0.110	0.596	0.596
	Total	626	19 (3.04)	15 (78.95)	1 (5.26)	1 (5.26)

Chi-square test. JKMS, Journal of Korean Medical Science; YMJ, Yonsei Medical Journal; Korean J Intern Med, Korean Journal of Internal Medicine.

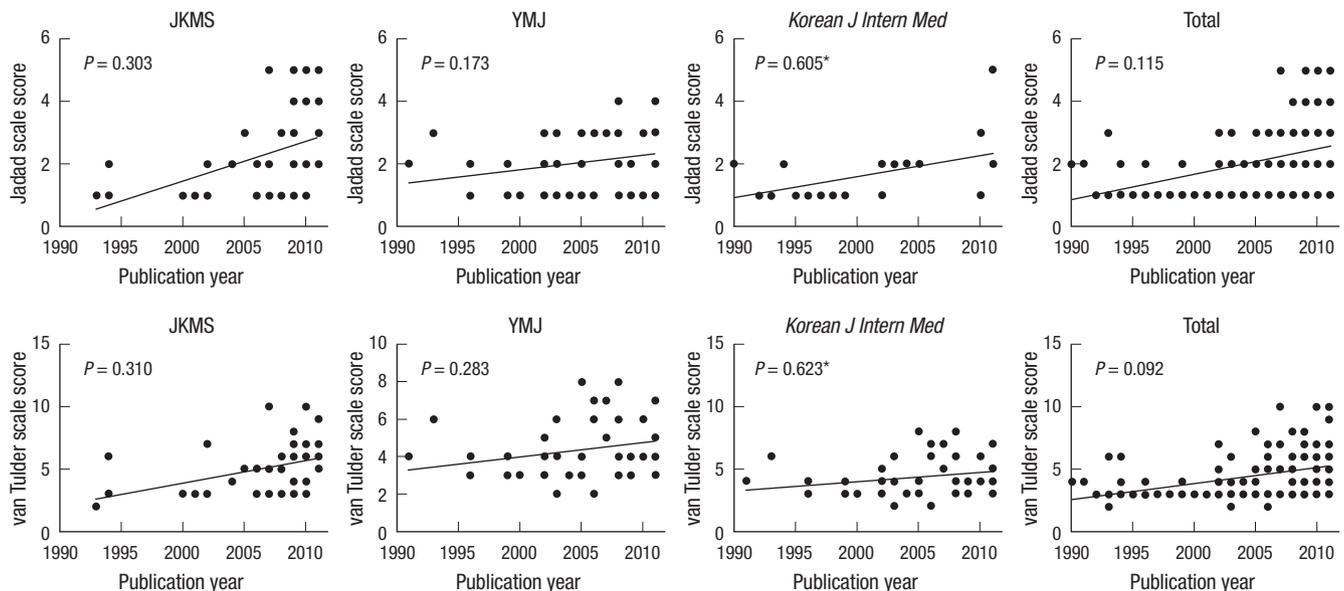


Fig. 1. Quality assessment of RCTs according to publication year. one-way ANOVA analysis, *Kruskal Wallis test.

Qualitative variation of RCTs over time

Jadad quality assessment scale

The results of quality assessment are presented in 5-yr units starting in 1986. The mean Jadad scale for RCTs published in the JKMS from 1991 to 1995 was 1.33 ± 0.58 and showed a slight increase to 3.00 ± 1.27 in 2011 ($P = 0.120$). In addition, there were no high quality articles published from 1991 to 1995, but this number increased to four (80.00%) of six RCTs in 2011 ($P = 0.111$).

All three journals, the quality of RCTs showed no statistical difference according to publication year in 1-yr units (Fig. 1).

van Tulder assessment scale

The mean van Tulder scale score of RCTs reported in the JKMS from 1991 to 1995 was 3.33 ± 2.31 , while that of RCTs reported from 2006 to 2010 was 5.12 ± 2.22 and 6.83 ± 1.84 in 2011 ($P = 0.048$). The number of high quality articles in the JKMS was one

Table 2. Quality assessment of RCTs according to publication year

Journals	Years	Jadad scale		van Tulder scale		Cochrane's assessment of risk bias		
		Score	High quality (%)	Score	High quality (%)	High risk (%)	Moderate risk (%)	Low risk (%)
JKMS	1986-1990	-	0	-	0	0	0	0
	1991-1995	1.33 ± 0.58	0	3.33 ± 2.31	1 (33.33)	3 (100.00)	0	0
	1996-2000	1.00 ± 0.00	0	3.00 ± 0.00	0	2 (100.00)	0	0
	2001-2005	1.71 ± 0.76	1 (14.29)	4.14 ± 1.46	2 (28.57)	7 (100.00)	0	0
	2006-2010	2.50 ± 1.53	12 (46.15)	5.12 ± 2.22	14 (53.85)	17 (65.38)	6 (23.08)	3 (11.54)
	2011	3.00 ± 1.27	4 (66.67)	6.83 ± 1.84	5 (83.33)	1 (16.67)	3 (50.00)	2 (33.33)
	<i>P</i> value	0.064	0.111*	0.017	0.169*			0.106*
YMJ	1960-1990	-	0	-	0	0	0	0
	1991-1995	2.50 ± 0.71	1 (50.00)	5.00 ± 1.41	1 (50.00)	2 (100.00)	0	0
	1996-2000	1.33 ± 0.52	0	3.33 ± 0.52	0	6 (100.00)	0	0
	2001-2005	1.71 ± 0.85	4 (23.53)	3.82 ± 1.51	4 (23.53)	16 (94.12)	1 (5.88)	0
	2006-2010	2.35 ± 0.88	10 (50.00)	4.95 ± 1.57	10 (50.00)	18 (90.00)	2 (10.00)	0
	2011	2.38 ± 0.92	3 (37.50)	4.75 ± 1.16	5 (62.50)	8 (100.00)	0	0
	<i>P</i> value	0.032	0.166*	0.053	0.072*			0.797*
Korean J Intern Med	1986-1990	2.00 ± 0.00	0	4.00 ± 0.00	0	1 (100.00)	0	0
	1991-1995	1.25 ± 0.50	0	3.25 ± 0.50	0	4 (100.00)	0	0
	1996-2000	1.00 ± 0.00	0	3.00 ± 0.00	0	4 (100.00)	0	0
	2001-2005	1.80 ± 0.45	0	3.80 ± 0.84	1 (20.00)	5 (100.00)	0	0
	2006-2010	1.67 ± 1.15	1 (33.33)	4.00 ± 1.00	1 (33.33)	3 (100.00)	0	0
	2011	3.50 ± 2.12	1 (50.00)	7.00 ± 4.24	1 (50.00)	1 (50.00)	0	1 (50.00)
	<i>P</i> value	0.107†	0.251*	0.140†	0.520*			0.110*
JKMS	Total	2.30 ± 1.39	17 (38.64)	4.98 ± 2.18	22 (50.00)	30 (68.18)	9 (20.45)	5 (11.36)
YMJ	Total	2.04 ± 0.90	18 (33.96)	4.38 ± 1.50	20 (37.74)	50 (94.34)	3 (5.66)	0
Korean J Intern Med	Total	1.68 ± 1.00	2 (10.53)	3.89 ± 1.63	3 (15.79)	18 (94.74)	0	1 (5.26)
	<i>P</i> value	0.138	0.081*	0.060	0.037*			0.003*

One-way ANOVA, *Chi-square test, †Kruskal Wallis test. JKMS, Journal of Korean Medical Science; YMJ, Yonsei Medical Journal; Korean J Intern Med, Korean Journal of Internal Medicine.

Table 3. Characteristics of RCTs according to subjects

Journals	Subjects, No.	Jadad scale		van Tulder scale		Cochrane's assessment of risk bias		
		Score	High quality (%)	Score	High quality (%)	High risk (%)	Moderate risk (%)	Low risk (%)
JKMS	Internal Medicine, 16	2.38 ± 1.41	5 (31.25)	5.44 ± 2.42	7 (58.82)	10 (62.50)	3 (18.75)	3 (18.75)
	Surgeries, 7	1.86 ± 1.57	2 (28.57)	4.14 ± 2.04	2 (25.00)	5 (71.43)	2 (28.57)	0
	Basic medicine, 5	2.80 ± 2.05	3 (60.00)	5.20 ± 3.11	2 (40.00)	2 (40.00)	1 (20.00)	2 (40.00)
	Others, 16	2.25 ± 1.13	7 (43.75)	4.81 ± 1.72	5 (31.25)	13 (81.25)	3 (18.75)	0
	<i>P</i> value	0.714	0.614*	0.561	0.857*			0.215*
YMJ	Internal Medicine, 10	1.90 ± 0.99	2 (20.00)	4.20 ± 1.48	3 (30.00)	9 (90.00)	1 (10.00)	0
	Surgeries, 16	2.00 ± 0.89	6 (37.50)	4.25 ± 1.69	6 (37.50)	15 (93.75)	1 (6.25)	0
	Others, 27	2.11 ± 0.89	10 (37.04)	4.52 ± 1.42	11 (40.74)	26 (96.30)	1 (3.70)	0
	<i>P</i> value	0.807	0.585*	0.787	0.836*			0.757*
Total (JKMS, YMJ, Korean J Intern Med)	Internal Medicine, 45	1.98 ± 1.18	9 (20.00)	4.53 ± 1.98	13 (51.11)	37 (82.22)	4 (8.89)	4 (8.89)
	Surgeries, 23	1.96 ± 1.11	8 (34.78)	4.22 ± 1.76	8 (34.78)	20 (86.96)	3 (13.04)	0
	Basic medicine, 5	2.80 ± 2.05	3 (60.00)	5.20 ± 3.11	2 (40.00)	2 (40.00)	1 (20.00)	2 (40.00)
	Others, 43	2.16 ± 0.97	17 (39.53)	4.63 ± 1.53	16 (37.21)	39 (90.70)	4 (9.30)	0
	<i>P</i> value	0.416	0.112*	0.694	0.848*			0.005*

One-way ANOVA, *Chi-square test. JKMS, Journal of Korean Medical Science; YMJ, Yonsei Medical Journal; Korean J Intern Med, Korean Journal of Internal Medicine.

(33.33%) from 1991 to 1995 but increased to five (83.33%) of six RCTs articles published in 2011 ($P = 0.169$). No statistical difference was seen in quality assessment of RCTs in JKMS, YMJ, and *Korean Journal of Internal Medicine* according to publication year in 1-yr units, respectively (Fig. 1).

Cochrane collaboration risk of bias tool

There were no low risk of bias articles among RCTs published

in the JKMS up to 2005 by CCRBT assessment, but this increased to three low risk of bias articles (11.54%) presented from 2006 to 2010. In addition, in 2011, there were two articles with a low risk of bias (33.33%; $P = 0.106$). The CCRBT assessment according to publication time of YMJ and *Korean Journal of Internal Medicine* showed no statistical difference ($P = 0.797$, $P = 0.110$, respectively; Table 2).

Table 4. Factors associated with quality of RCTs

Journals	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 (%)
JKMS	Intervention type								
	Drug	32 (72.73)	2.41 ± 1.41	12 (37.50)	5.25 ± 2.30	18 (56.25)	21 (65.63)	6 (18.75)	5 (15.63)
	Non-drug	12 (27.27)	2.00 ± 1.35	5 (41.67)	4.25 ± 1.71	4 (33.33)	9 (75.00)	3 (25.00)	0
	P value		0.39	0.800*	0.117	0.176*	0.341*		
	Funding source								
	Yes	20 (45.45)	2.45 ± 1.54	9 (45.00)	5.55 ± 2.56	11 (55.00)	12 (60.00)	4 (20.00)	4 (20.00)
	No	24 (54.55)	2.17 ± 1.27	8 (33.33)	4.50 ± 1.72	11 (45.83)	18 (75.00)	5 (20.83)	1 (4.17)
	P value		0.515	0.429*	0.113	0.545*	0.250*		
	Reviewed by IRB								
Yes	28 (63.64)	2.57 ± 1.50	14 (50.00)	5.32 ± 2.34	15 (53.57)	17 (60.71)	6 (21.43)	5 (17.86)	
No	16 (36.36)	1.81 ± 1.05	3 (18.75)	4.38 ± 1.78	7 (43.75)	13 (81.25)	3 (18.75)	0	
P value		0.078	0.041*	0.143	0.531*	0.172*			
YMJ	Intervention type								
	Drug	28 (52.83)	2.39 ± 0.88	14 (26.42)	5.04 ± 1.60	16 (30.19)	25 (89.29)	3 (10.71)	0
	Non-drug	25 (47.17)	1.64 ± 0.76	4 (7.55)	3.64 ± 0.95	4 (7.55)	25 (100)	0	0
	P value		0.001	0.009*	< 0.001	0.002*	0.092*		
	Funding source								
	Yes	19 (35.85)	2.32 ± 0.82	8 (15.09)	4.74 ± 1.15	8 (15.09)	18 (94.74)	1 (5.26)	0
	No	34 (64.15)	1.88 ± 0.91	10 (18.87)	4.18 ± 1.64	12 (22.64)	32 (94.12)	2 (5.88)	0
	P value		0.084	0.349*	0.152	0.624*	0.925*		
	Reviewed by IRB								
Yes	32 (60.38)	2.34 ± 0.87	15 (28.30)	4.84 ± 1.51	17 (32.08)	29 (90.63)	3 (9.38)	0	
No	21 (39.62)	1.57 ± 0.75	3 (5.66)	3.67 ± 1.20	3 (5.66)	21 (100)	0	0	
P value		0.001	0.014*	0.003	0.004*	0.149*			
Korean J Intern Med	Intervention type								
	Drug	15 (78.95)	1.60 ± 1.06	1 (5.26)	3.93 ± 1.83	3 (15.79)	14 (93.33)	0	1 (6.67)
	Non-drug	4 (21.05)	2.00 ± 0.82	1 (5.26)	3.75 ± 0.50	0	4 (100)	0	0
	P value		0.200†	0.288*	0.509†	0.330*	0.596*		
	Funding source								
	Yes	8 (42.11)	2.00 ± 1.41	2 (10.53)	4.50 ± 2.33	2 (10.53)	7 (87.50)	0	1 (12.50)
	No	11 (51.89)	1.45 ± 0.52	0	3.45 ± 0.69	1 (5.26)	11 (100)	0	0
	P value		0.520†	0.080*	0.219†	0.348*	0.228*		
	Reviewed by IRB								
Yes	7 (36.84)	2.14 ± 1.46	2 (10.53)	4.71 ± 2.43	2 (10.53)	6 (85.71)	0	1 (14.29)	
No	12 (63.16)	1.42 ± 0.52	0	3.42 ± 0.67	1 (5.26)	12 (100)	0	0	
P value		0.279†	0.050*	0.094†	0.243*	0.179*			
Total	Intervention type								
	Drug	75 (64.66)	2.24 ± 1.20	27 (23.28)	4.92 ± 1.99	37 (31.90)	60 (80.00)	9 (12.00)	6 (8.00)
	Non-drug	41 (35.34)	1.78 ± 0.96	10 (8.62)	3.83 ± 1.20	8 (6.90)	38 (92.68)	3 (7.32)	0
	P value		0.027	0.200*	0.002	0.002*	0.114*		
	Funding source								
	Yes	47 (40.05)	2.32 ± 1.25	19 (16.38)	5.04 ± 2.06	21 (18.10)	37 (78.72)	5 (10.64)	5 (10.64)
	No	69 (59.48)	1.91 ± 1.03	18 (15.52)	4.19 ± 1.57	24 (20.69)	61 (88.41)	7 (10.14)	1 (1.45)
	P value		0.069	0.104*	0.018	0.283*	0.087*		
	Reviewed by IRB								
Yes	67 (57.76)	2.42 ± 1.22	31 (26.72)	5.03 ± 1.98	34 (29.31)	52 (77.61)	9 (13.43)	6 (8.96)	
No	49 (42.24)	1.61 ± 0.81	6 (5.12)	3.86 ± 1.34	11 (9.48)	46 (93.88)	3 (6.12)	0	
P value		< 0.001	< 0.001*	< 0.001	0.002*	0.034*			

Student-t test, *Chi-square test, †Mann-Whitney test. JKMS, Journal of Korean Medical Science; YMJ, Yonsei Medical Journal; Korean J Intern Med, Korean Journal of Internal Medicine.

Analysis of RCT quality by medical subject

Among the RCT-based articles presented in the JKMS during 26 yr (1986-2011), 16 articles were in internal medicine, seven in surgery, five in basic medicine, and 16 in other specialties including pediatrics, anesthesiology, and rehabilitation medicine. In the quality assessment by subject, there were no statistically significant differences observed between groups. From the results of CCRBT, two articles (40%) from basic medicine and three

(18.75%) from internal medicine were identified as having low risks of bias. In addition, the quality of overall RCTs in three journals according medical subjects showed no significant difference statistically (Table 3).

Analysis of factors related to the quality of the articles

There was no statistically significant difference in quality observed between drug and non-drug studies in the JKMS. How-

ever, in YMJ the quality of RCTs showed significant difference between drug and non-drug studies (Jadad scale; $P = 0.001$, van Tulder scale; $P < 0.001$). No significant differences were found between studies with and without funding in all journals. The quality assessment of overall RCTs in three journals using van Tulder scale presented significant difference according existence of funding ($P = 0.018$). From the Jadad scale assessment, the ratio of high quality articles was higher in studies reviewed by IRB than studies with no such review ($P = 0.041$). The quality assessment of total RCTs in all of three journals showed higher scores in studies reviewed by IRB than no IRB reviewed studies (Table 4).

DISCUSSION

In this study, the quality of original RCT reports published in the JKMS, YMJ and *Korean Journal of Internal Medicine* was assessed. While the number of RCTs published has gradually increased, no statistically significant increase in the quality of the articles was observed. The small number of double blinded reports and the absence of methodological details described for concealment of allocation have hindered high quality assessment.

There had been three previous analyses on the quality of RCTs published in Korean journals. Kim et al. performed an analysis on RCTs published in five different national academic journals: the *Korean Journal of Internal Medicine*, the *Journal of the Korean Surgical Society*, the *Korean Journal of Obstetrics and Gynecology*, the *Korean Journal of Pediatrics*, and the *Korean Journal of Family Medicine* (10). According to that report, the number of RCTs with a Jadad score of ≥ 2 points increased in the 1990s compared to the 1980s. In addition, the number of papers with high quality articles increased from two to seven during the same time period. In another study, Chung et al. analyzed RCTs published in the *Korean Journal of Family Medicine* from 1980 to 2005 (4). Among the 1,290 original articles published, there were 23 RCT articles (1.8%), and their mean Jadad scale score was 1.87. The authors reported that the proportion of RCTs increased from 1.09% of the original articles in the 1980s to 2.63% in the 2000s, with a corresponding 1.17-point increase in the Jadad scale from 1.00 in the 1980s to 2.17 points in the 2000s. These two studies clearly indicate that the number and quality of RCT articles increased over time. This is considered a result of the growing influence of evidence-based medicine. Therefore, the overall number of RCTs that provides a high quality level of evidence is increasing (12). In the present study, the number of RCTs increased as in these previous reports. However, the quality status of the RCTs analyzed here did not significantly change over time. Specifically, the quality status did quantitatively increase over time, but a significant difference was observed only from the assessment by the van Tulder scale in terms of quality in the JKMS. However, when the qualities of articles were compared in 1-yr

units, rather than 5-yr units, no statistically significant differences were observed.

Recently, Lee et al. (13) analyzed the quality of RCTs published in *the Korean Journal of Urology* using the Jadad scale. Their study showed that there were 28 RCT articles (0.89%) out of 3,156 original articles presented since 1991 and during the following 20 yr. The mean Jadad scale score of those RCT articles was 1.75. In addition, eight high quality articles were found out of those 28 articles. The number of RCT articles, which totaled five before 2000, increased to 23 by 2000, and their quality improved over time. Moreover, Lee et al. (13) showed that only one article in *the Korean Journal of Urology* had adequate allocation concealment. In the present study, among the RCTs published by the JKMS, the majority did not even comment on concealment of allocation. In 2006, an article contained the context related to the concealment of allocation for the first time, but only eight articles (18.18%) in total provided details. There is the only one article contained the context related to the concealment of allocation in *the Korean Journal of Internal Medicine*. Schulz et al. (14) explained that without concealment of allocation, randomization tends to be damaged in the process of study performance, even if the randomization is well conducted. This omission could distort the effects of intervention by $\geq 40\%$. Hewitt et al. (15) presented that cases showing inappropriate or uncertain concealment of allocations constituted almost 46% of RCTs published in four different major medical journals (*British Medical Journal*, *the Journal of the American Medical Association*, *Lancet*, and *New England Journal of Medicine*) in 2000. If appropriate planning and preventions are taken by the medical academics of Korea on double blind and concealment of allocation in RCTs, the quality of such reports would improve.

There have been few studies conducted worldwide on qualitative analysis of RCTs. Uetani et al. (1) analyzed whether RCTs conducted in Japan and published in medical journals met the consolidated standards for reporting of trials (CONSORT) statements, which provide guidance on reporting RCTs and is comprised of a checklist, flow diagram, explanations, and extensions (16). They showed that 98 RCT articles had been published in various academic journals from January to March in 2004, and only 11 RCT articles were in compliance with the CONSORT statements. However, since the CONSORT statement was not a quality assessment tool, it was not possible to digitize the qualitative analysis.

There are various types of qualitative assessment tools for RCTs including Campell, Moher, Chalmers, Jadad, van Tulder, Newell's, and Cochrane. The interesting point in this study is that by using three different tools, we found differences in the qualitative analysis outcomes of RCTs. The assessment of the quality of trials remains controversial, and there is no consensus on highly accurate and valid tools (17). However, in this study, efforts were made to overcome such limitations by using three

different tools: the Jadad scale, van Tulder scale, and CCRBT. These are representative assessment tools used most commonly nationwide and worldwide. In particular, the Jadad scale has advantages in the simplicity of the assessment questions and ease of assessment performance, but it does not include assessment items for the most important item of RCT assessment: concealment of allocation. Therefore, additional analyses were performed using the van Tulder scale and CCRBT to supplement in this regard.

Furthermore, in this study, articles that were reviewed and approved by an institutional review board (IRB) were, on average, of higher quality. To the best of our knowledge, no previous studies have been conducted in connection to the association of IRB review with the quality of articles. The review of IRB serves to obtain acknowledgement of the feasibility of the study design and performance in the study protocol. The recent system to validate study protocols in RCTs in order to obtain IRB approval has played an important role in raising the quality of articles since IRB review is considered an international quality standard. Studies that used placebo and drug products easy for double blinding were considered high quality in comparison to the studies with non-drugs. However, there were no significant intergroup differences observed from the results of this study in JKMS. However, overall quality assessment of RCTs in three journal, there was statistically difference according intervention type. Among studies that analyzed RCTs previously published by *the Korean Journal of Urology* and *the Journal of Korean Academy of Family Medicine*, those receiving financial support were able to establish well-organized study designs and perform orderly research, resulting in many high quality articles (4, 13). By contrast, Clifford et al. analyzed 100 RCT articles published in five different peer-reviewed, high impact, general medical journals (*Annals of Internal Medicine*, *the British Medical Journal*, *the Journal of the American Medical Association*, *Lancet*, and *New England Journal of Medicine*) and found no association between the funding source and the quality of the article (18). This study also did not reveal significant differences in the qualitative assessment of articles depending on the presence/absence of funding except van Tulder scale scoring. This result is not considered to be an accurate conclusion, however, because the number of analyzed articles was small in this report. Therefore, re-assessment through analyses of more RCTs is required.

The limitations of this study include a probability for subjective judgment of the researcher likely to intervene in extraction of RCTs and their quality assessments. Therefore, two medical doctors independently extracted RCTs while the assessments were performed independently by two reviewers; thereafter, the outcomes were adjusted to secure objectivity and reliability. The assessment of RCTs published by the JKMS, YMJ, and *Korean Journal of Internal Medicine* was implemented for the first time in this study. Suggestions for the qualitative improvement

of medical research in the Republic of Korea would be a significant contribution of this study.

While the number of RCTs published in the JKMS, YMJ, and *Korean Journal of Internal Medicine* has gradually increased according to publication time, the quality of these reports has remained unchanged. Therefore, national medical academics should focus more efforts in performing high quality studies to ensure appropriate randomization, reviews by IRB, financial support, and inclusion of allocation concealment during study performance.

REFERENCES

1. 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.
2. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D et al. *Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA* 1996; 276: 637-9.
3. Jackson JL, Srinivasan M, Rea J, Fletcher KE, Kravitz RL. *The validity of peer review in a general medicine journal. PLoS One* 2011; 6: e22475.
4. Chung W, Lee KW, Hwang IH, Lee DH, Kim SY. *Quality assessment of randomized controlled trials in the Journal of the Korean Academy of Family Medicine. Korean J Fam Med* 2009; 30: 626-31.
5. Liberati A, Himel HN, Chalmers TC. *A quality assessment of randomized control trials of primary treatment of breast cancer. J Clin Oncol* 1986; 4: 942-51.
6. Autorino R, Borges C, White MA, Altunrende F, Perdona S, Haber GP, De Sio M, Khanna R, Stein RJ, Kaouk JH. *Randomized clinical trials presented at the World Congress of Endourology: how is the quality of reporting? J Endourol* 2010; 24: 2067-73.
7. Lim SM, Shin ES, Lee SH, Seo KH, Jung YM, Jang JE. *Tools for assessing quality and risk of bias by levels of evidence. J Korean Med Assoc* 2011; 54: 419-29.
8. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ. *Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials* 1996; 17: 1-12.
9. Moher D, Cook DJ, Jadad AR, Tugwell P, Moher M, Jones A, Pham B, Klassen TP. *Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses. Health Technol Assess* 1999; 3: i-iv, 1-98.
10. Kim SW, Choi YS, Ahn HS, Lee HY, Ahn DS, Lee YM. *Quantity and quality assessment of randomized controlled trials published in five Korean medical journals, from 1980 to 2000. J Korean Acad Fam Med* 2004; 25: 118-25.
11. van Tulder M, Furlan A, Bombardier C, Bouter L. *Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. Spine (Phila Pa 1976)* 2003; 28: 1290-9.
12. Keech AC, Pike R, Granger RE, GebSKI VJ. *Interpreting the results of a clinical trial. Med J Aust* 2007; 186: 318-9.
13. Lee JY, Chung JH, Kang DH, Lee JW, Moon HS, Yoo TK, Choi HY, Lee SW. *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.
14. Schulz KF, Grimes DA. *Allocation concealment in randomised trials: defending against deciphering. Lancet* 2002; 359: 614-8.
 15. 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.
 16. Moher D, Schulz KF, Altman D. *The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA* 2001; 285: 1987-91.
 17. Armijo-Olivo S, Stiles CR, Hagen NA, Biondo PD, Cummings GG. *Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. J Eval Clin Pract* 2012; 18: 12-8.
 18. 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.