Quality Evaluation of Randomized Controlled Trials on Complementary and Alternative Medicine

Kasil Oh[1], PhD, Keum-Soon Kim[2], PhD, Jee-Won Park[3]*, PhD, Jiyeon Kang[4], PhD

[1]College of Nursing, Yonsei University, Seoul, Korea
[2]College of Nursing, Seoul National University, Seoul, Korea
[3]College of Nursing, Ajou University, Suwon, Korea
[4]Department of Nursing, Dong-A University, Busan, Korea

Purpose   This study aims to describe the research characteristics and analyze the methodological quality of randomized clinical trials on complementary and alternative medicine (CAM).

Methods   A total of 76 studies using randomized controlled trials (RCTs) on CAM (16 by Koreans, 60 by internationals) published in the 6 years from 2000 to 2005 were reviewed systematically and analyzed with assessment criteria developed by the researchers on the basis of Jadad guidelines.

Results   Most of the studies were carried out in the area of medicine, nursing and CAM. More than 80% of the study subjects were patients. CAM modalities for independent variables were mainly on energy medicine, mind-body medicine and manipulative and body-based practices, while dependent variables were mostly physiological and psychological indexes. Most of the studies utilized randomization (93.4%) and identified the dropout rate (90.8%), whereas allocation concealment (49.3%) and double-blinding (18.9%) were specified in a small number of studies. The overall quality of RCTs based on the assessment criteria of this study was satisfactory. However, the quality score of the Korean studies (2.87) was lower than that of the international studies (3.37).

Conclusion   The methodological objectivity of CAM studies has been improving in spite of controversy over the scientific bases of CAM. More Korean studies with rigorous experimental design are needed to build up the evidence-based practice of CAM. [*Asian Nursing Research* 2007;1(3):153–164]

Key Words   complementary and alternative medicine, randomized controlled trials, research methodology, quality

INTRODUCTION

Complementary and alternative medicine (CAM) has become popular for the purposes of health promotion, disease prevention and treatment. As a result, many studies on CAM (Barnes & Bero, 1997; Eisenberg et al., 1998; Menniti-Ippolito, Gargiulo, Bologna, Forcella, & Raschetti, 2002; Ni, Simile, & Hardy, 2002)

*Correspondence to: Jee-Won Park, PhD, College of Nursing, Ajou University, San 5 Wonchun Dong, Yeongtonggu, Suwon 443-749, Korea.
E-mail: pjwon@ajou.ac.kr

have been published worldwide. Under the influence of this trend, CAM research has also grown rapidly in health care disciplines including nursing and medicine. According to Barnes, Abbot, Harkness, and Ernst (1999), there were about 33,000 articles on CAM published in medical journals from 1966 to 1996. For the period between 1997 and 2002, a total of 20,209 CAM studies were published, where various disciplines including medicine showed their interests in CAM (Raschetti, Menniti-Ippolito, Forcella, & Bianchi, 2005).

However, it has been pointed out that the research quality of studies on CAM was, in some cases, not sufficient to meet the standard of scientific evidence (Oh, Kim, Kwon, & Park, 2006). Randomized controlled trials (RCTs) are widely accepted as the gold standard for providing evidence on the effectiveness of interventions (Wang et al., 2007). Among the CAM studies conducted from 1997 to 2002 (Raschetti et al., 2005), only 7.5% used RCTs, while the rest focused on the growing tendency of using CAM from the view point of sociology, history and anthropology. The reasons for the limited use of RCTs can be explained as follows: CAM practitioners were not much interested in CAM method; researchers were not well trained in designing or conducting RCTs; some scientific communities did not voluntarily recognize CAM as an acceptable intervention (Raschetti et al.). A number of studies using RCTs also failed to specify randomization, dropout rate, sample size, and adverse events (Linde, Jonas, Melchart, & Willich, 2001; Tang, Zhan, & Ernst, 1999).

Raschetti et al. pointed out in 2005 that more than half of the articles on RCTs were published in journals with no impact factor. Although 67% of 351 controlled clinical trials reported statistically significant effects of CAM, the scores of method quality remained in the range of 23–51% of maximum score (Pittler, Abbot, Harkness, & Ernst, 2000). This led to questions being raised on the effectiveness of CAM due to the lack of scientific evidence. Accordingly, the evidence should be found through systematic reviews.

Every year, numerous articles on CAM in the fields of medicine and nursing are published in Korea,

reflecting the extended interests in CAM (Kim, 2003; Lee et al., 2002; Seol, Choi, & Jong, 2002; Son, 2002). Most of the CAM studies in Korea have tended to focus on quasi- or pre-experimental designs, while RCTs have been frequently used internationally. Therefore, there are few published investigations that systematically evaluate the quality of RCTs of CAM in Korea. But it is still important to prove the scientific attributes of CAM with rigorous research methods to ensure the possible use of CAM as a nursing intervention (Oh et al., 2006). The purposes of this study were to: (a) describe the general characteristics of RCTs; (b) evaluate the research method of RCTs; and (c) compare the quality of RCT studies published in Korean journals with that of international studies.

Definitions of CAM

The National Center for Complementary and Alternative Medicine (NCCAM) of the USA (2002) defines CAM as "a group of diverse medical and health care systems, practices and products that are not presently considered to be part of conventional medicine" and classifies the types of CAM into the following: (a) biologically-based practices; (b) energy medicine; (c) manipulative and body-based practices; (d) mind-body medicine; and (e) whole medical systems.

In this study, CAM is defined as treatments and therapies that are not part of traditional Western medicine or oriental medicine and include energy medicine, manipulative and body-based practices and mind-body medicine as search terms. Biologically-based practices such as herbal medicines or vitamins and whole medical systems such as traditional Chinese medicine or Ayurveda were excluded because it is difficult to utilize these methods for nursing interventions.

METHODS

Research data selection

We searched the following databases to find eligible studies published between January 2000 and June 2005: PubMed was searched for English-language

publications, and KoreaMed, Riss4U and RICHIS were searched for those published in Korean. The key words used in the search were "CAM" in English and the 76 therapies classified as CAM by NCCAM in both Korean and English.

Searching was restricted to English-language journals (in PubMed), human studies and types of CAM intervention (energy medicine, manipulative practice, mind-body medicine).

The total number of eligible samples of 301 abstracts of experimental studies on CAM (202 by Koreans, 99 by internationals) from health-related scientific journals between 2000 and 2005, excluding those published in China, were selected for preview. Of those 301 studies, 223 had full-text articles available for review (161 by Koreans, 62 by internationals). As a final step, 76 RCT studies (16 by Koreans, 60 by internationals) were chosen for analysis (Figure 1). These data were part of the study conducted by Oh et al. in 2006.

Analysis tools

The researchers developed a descriptive recording form to investigate research characteristics and methodological quality of research. The form included
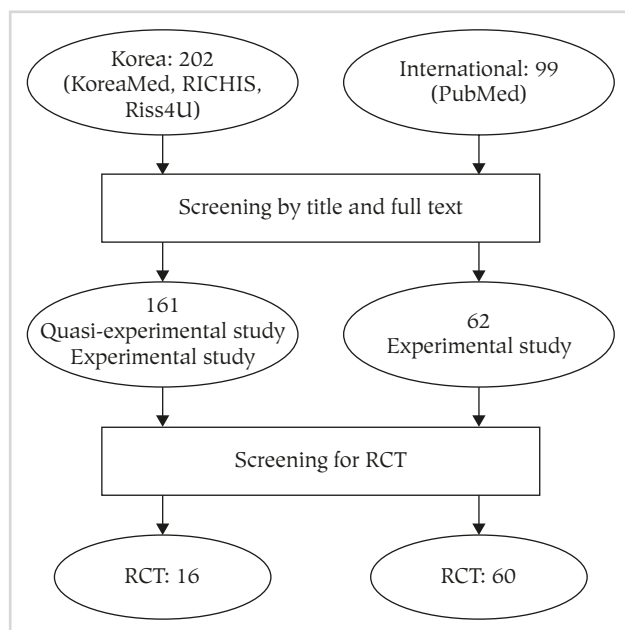

Figure 1. Literature screening process.

conditions for experimental design (randomization, double-blinding, dropout rates, allocation concealment) and quality of sample and intervention. In addition to the descriptive recording form, Jadad's scale was used to quantify the quality of the studies. The highest possible score of Jadad's scale is 5 points and its sub-items and their score are as follows: the study is described as randomized (+1); the means of carrying out randomization is described and appropriate (+1); the study is described as double-blind (+1); the means of double blinding is described and appropriate (+1); there is a description of dropouts giving number and reason in both groups (+1).

Many scales and checklists are available to assess the quality of RCTs. However, the Jadad scale is known as a standard quality measure and is the only known scale developed with standard scale development techniques (Jadad et al., 1996). It was also systematically developed and underwent an empirical validation process, so the scale has better discriminative power (Linde et al., 2001).

Data analysis

Two doctoral students reviewed the full texts of the final chosen publications. The reviewers were qualified to comprehend and analyze research papers and trained for coding the detailed record forms that were developed by the researchers prior to the commencement of the review.

The process of data extraction and the evaluation of quality were performed independently by two reviewers, and discrepancies were resolved through discussion by the research team.

The data were statistically analyzed using SPSS version 12.0 (SPSS Inc., Chicago, IL, USA). In many studies using the Jadad scale, RCTs were classified as higher quality if their scores were ≥3 and as lower quality if their scores were ≤2 on the 5-point scale (Trinh, Phillips, Ho, & Damsma, 2004). This criterion was recommended as the standard by Khan, Daya, and Jadad in 1996. So, only a study that scored ≥3 points on Jadad's scale was classified as a high quality study. The differences in methodological quality according to the general characteristics of the studies were analyzed by the χ^2 test.

RESULTS

General characteristics of RCTs on CAM

Table 1 shows the overall research characteristics of the 76 selected articles. More than half of those were published in 2003 (22.4%) and 2004 (29.0%). American principal investigators (PI) made up 39.5%, and 46.1% of the articles came from authors affiliated to educational institutes. The most active participants were from medicine (30.3%), CAM-related fields (22.4%) and nursing (21.1%). The energy medicine area comprised 38.2% of the studies that utilized CAM interventions, and the mind-body medicine area comprised 36.8%. The majority of study subjects were patients (84.2%), and 61.8% of interventions were tried in hospitals. The outcome indexes used to measure the effects of CAM intervention were physiological index (47.4%), psychological function (36.8%), body function (34.2%) and the level of pain (30.3%).

Quality of RCTs on CAM
Quality of experimental design

Table 2 shows the results of quality assessment. Most of the studies clearly stated their randomization process, and almost two thirds (63.4%) of them used randomization properly. With regard to blinding, 22 studies (28.9%) applied double-blinding. Of these 22 double-blinded studies, 20 (90.9%) employed a relevant blinding technique. Of the 69 (90.8%) studies that reported the dropout rate, 32 studies (42.1%) reported that there was no dropout. Allocation concealment was well described in only 35 randomized studies (49.3%). The mean Jadad score of primary studies was 3.37 out of 5; 31 studies (40.8%) scored 4 points and 13 (17.1%) scored 5 points. Meanwhile, 23 studies (30.3%) were of low quality (< 3 points) and 53 (69.7%) studies were classified as high quality.

Quality of sample and intervention program

In addition to the sub-items of Jadad's guideline, the quality of study sample and intervention program was analyzed (Table 3). The total number of study subjects (number in experimental and control groups) was counted. The description of selection criteria, homogeneity test, and power analysis were also included in the analysis. The minimum sample size for normal distribution to satisfy the hypothesis of a statistical test is 30 for each experimental and control group. Twenty-nine studies (38.2%) included more than 30 experimental subjects and 23 studies (30.3%) included more than 30 control subjects. Therefore, 26 studies (34.2%) had more than 60 subjects in total. The majority of studies (85.5%) described the criteria used for sample selection, and homogeneity tests for demographic characteristics and diagnoses were done in 63 studies (82.9%). Scientific bases for the sample size, such as power analysis, were carried out in 60 studies (78.9%).

When an analysis was performed for the use of the intervention protocols (frequency, period, intensity), 72 studies (94.7%) reported that they had protocols for their interventions. Intervention frequency (85.5%), period (92.1%), and intensity (69.7%) were specified in more than half of the studies, whereas clear statements on adverse events (6.6%), qualification of the intervention provider (36.8%) and consistency (27.6%) of intervention were relatively low.

Evaluation criteria by quality and modalities of CAM

Comparisons per evaluation criterion are shown in Table 4. All of the high-quality studies had a clear statement on randomization and 85.7% had applied relevant randomization methods. While there was no low-quality study that applied double-blinding, 44.9% of the high-quality studies did apply double-blinding. The dropout rates were clearly stated in 85.2% of low-quality studies and in 93.9% of high-quality studies. Only one low-quality study described allocation concealment while 69.4% of the high-quality studies did so.

The frequencies of randomization and double-blinding were also high in energy medicine studies (Table 4). More than 79% and 44% of energy medicine studies received 2 points in randomization and double-blinding, respectively. The occurrence of clear statements on dropout rates was highest in the studies on manipulative and body-based practices (94.7%). The rate of relevant allocation concealment was high in the energy medicine studies (69.0%).

Table 1					
General Research Characteristics					
Characteristics of principal investigator	Category	Total studies (*N* = 76)		Korean studies (*n* = 16)	
		n	%	*n*	%
Nationality	Korea	16	21.1		
	USA	30	39.5		
	England	16	21.1		
	Germany	6	7.9		
	Etc.	8	10.5		
Affiliation	Educational institution	35	46.1	5	31.2
	Health institution	14	22.4	4	25.0
	Research center	15	19.7	0	0
	Etc.	9	11.8	7	43.8
Specialty area	Nursing	16	21.1	7	43.8
	Medicine	23	30.3	7	43.8
	CAM-related	17	22.4	0	0
	Etc.	20	26.3	2	12.5
Year of publication	2000	5	6.6	1	6.2
	2001	9	11.8	0	0
	2002	14	18.4	1	6.2
	2003	17	22.4	5	31.3
	2004	22	29.0	5	31.3
	2005 (January–June)	9	11.8	4	25.0
Modality of CAM	Energy medicine	29	38.2	3	18.7
	Manipulative practices	19	25.0	5	31.3
	Mind-body medicine	28	36.8	8	50.0
Study subject	Patient	64	84.2	14	87.5
	Non-patient	12	15.8	2	12.5
Study setting	Hospital	47	61.8	9	56.3
	Community	22	29.0	3	18.7
	Not specified	7	9.2	4	25.0
Dependent[a] variables	Psychological index	28	36.8	3	18.7
	Cognitive function	11	14.5	4	25.0
	Behavior	7	9.2	0	0
	Quality of life	12	15.8	1	6.2
	Pain	23	30.3	2	12.5
	Disease state	20	26.3	4	25.0
	Physiological index	36	47.4	6	37.5
	Body function	26	34.2	4	25.0

[a]Multiple answers were allowed.

Table 2

Quality of Experimental Design

Criteria	Category	Total studies (N = 76)		Korean studies (n = 16)	
		n	%	*n*	%
Randomization					
Specified	Yes	71	93.4	13	81.3
	No	5	6.6	3	18.7
Adequacy of method	Yes	45	63.4	9	69.2
	No	26	36.6	4	30.8
Double-blinding					
Specified	Yes	22	28.9	3	18.7
	No	54	71.1	13	81.3
Adequacy of method	Yes	20	90.9	2	66.7
	No	2	9.1	1	33.3
	Total	22	100.0		
Dropout rate					
Specified	Yes	69	90.8	12	75.0
	No	7	9.2	4	25.0
Distribution of rate	0%	32	42.1	10	90.9
	1–10%	11	14.5	0	0
	11–20%	11	14.5	0	0
	21–30%	7	9.2	0	0
	> 30%	8	10.5	1	9.1
	Not specified	7	9.2	0	0
Allocation concealment					
Adequacy of method	Adequate	35	49.3	6	37.5
	Inadequate/unclear	36	50.7	10	62.5
Jadad scale					
Distribution of scores	1	6	7.9	3	18.7
	2	17	22.4	3	18.7
	3	9	11.8	3	18.7
	4	31	40.8	7	43.8
	5	13	17.1	0	0
	Mean	3.37 ± 1.23		2.87 ± 1.20	

Quality assessment by Jadad scale

The mean Jadad score of the 76 studies was 3.37 out of 5 points. Thirty-one studies (40.8%) scored 4 points and 17 obtained 2 points. Thirteen studies (17.1%) scored the highest possible score of 5 points. In this study, 53 (69.7%) of the 76 studies were classified as high quality (≥ 3 points), and 23 (30.3%) studies were classified as low quality (≤ 2 points).

Research characteristics and quality of RCTs on CAM in Korea

An analysis of the research characteristics and method quality of 16 Korean RCTs on CAM was done. Nine of 16 PIs belonged to educational or medical institutes. Seven PIs specialized in nursing and another seven specialized in medicine. Fourteen studies were published between 2003 and 2005. Mind-body

Table 3

Quality of Sample and Intervention

Criteria	Category	Subcategory	Total studies ($N = 76$)		Korean studies ($n = 16$)	
			n	%	*n*	%
Quality of sample						
Total number	< 60		38	50.0	13	81.3
	≥ 60		26	34.2	3	18.7
	Not specified		12	15.8		
Number of	< 30		43	52.6	13	81.3
experimental	≥ 30		29	38.2	3	18.7
subjects	Not specified		7	9.2		
	Mean		49.1 ± 75.7		24.4 ± 13.3	
Number of	< 30		41	53.9	14	87.5
control subjects	≥ 30		23	30.3	2	12.5
	Not specified		12	15.8		
	Mean		42.5 ± 61.6		23.4 ± 9.4	
Inclusion criteria	Yes		65	85.5	14	87.5
	No		11	14.5	2	12.5
Homogeneity test	Yes		63	82.9	12	75.0
for demographic		Homogeneous	*49*		*9*	
characteristics		Non-homogeneous	*14*		*3*	
	No		13	17.1	4	25.0
Homogeneity test	Yes		63	82.9	13	81.3
for diagnostic		Homogeneous	*53*		*12*	
characteristics		Non-homogeneous	*10*		*1*	
	No		13	17.1	3	18.7
Rational for	Yes		60	78.9	11	68.8
sample size		Power analysis	*18*		*3*	
		No power analysis	*42*		*8*	
	No		16	21.1	5	31.2
Quality of intervention						
Protocol	Yes		72	94.7	15	93.8
	No		4	5.3	1	6.2
Frequency	Yes		65	85.5	13	81.3
	No		11	14.5	3	18.7
Period	Yes		70	92.1	16	100.0
	No		6	7.9	0	
Intensity	Yes		53	69.7	15	93.8
	No		23	30.3	1	6.2
Adverse events	Yes		5	6.6	1	6.2
	No		71	93.4	15	93.8
Qualification of	Yes		28	36.8	3	18.7
providers	No		48	63.2	13	81.3
Consistency of	Yes		21	27.6	3	18.7
intervention	No		55	72.4	13	81.3

Table 4

Evaluation Criteria by Quality and Modalities of CAM

Criteria	Methodological quality				Modalities of CAM					
	Low quality[a]		High quality[b]		Energy medicine		Manipulative & body-based practices		Mind-body medicine	
	f	%	*f*	%	*f*	%	*f*	%	*f*	%
1. Randomization										
Not specified (0)	5	18.5	0	0.0	1	3.5	1	5.2	3	10.7
Specified (1)	19	70.4	7	14.3	5	17.2	9	47.4	12	42.9
Specified + relevant method (2)	3	11.1	42	85.7	23	79.3	9	47.4	13	46.4
Total	27	100.0	49	100.0	29	100.0	19	100.0	28	100.0
2. Double-blind										
Not specified (0)	27	100.0	27	55.1	15	51.7	17	89.5	22	78.6
Specified (1)	0	0.0	2	4.1	1	3.5	0	0.0	1	3.6
Specified + relevant method (2)	0	0.0	20	40.8	13	44.8	2	10.5	5	17.9
Total	27	100.0	49	100.0	29	100.0	19	100.0	28	100.0
3. Dropout rate										
Not specified (0)	4	14.8	3	6.1	3	10.3	1	5.3	3	10.7
Specified (1)	23	85.2	46	93.9	26	89.7	18	94.7	25	89.3
Total	27	100.0	49	100.0	29	100.0	19	100.0	28	100.0
4. Allocation concealment										
Irrelevant/not specified	26	96.3	15	30.6	9	31.0	14	73.7	18	64.3
Relevant	1	3.7	34	69.4	20	69.0	5	26.3	10	35.7
Total	27	100.0	49	100.0	29	100.0	19	100.0	28	100.0

[a]Jadad score < 3; [b]Jadad score ≥ 3.

medicine was the most frequently used CAM modality. Nine studies were implemented in hospitals and six studies measured physiological index as a dependent variable (Table 1).

Thirteen studies clearly described the randomization process, and nine of these studies applied relevant randomization methods. Only six randomized studies carried out appropriate allocation concealment. With regard to blinding technique, only three studies employed double-blinding, while a single-blind method was employed in five studies. Twelve studies reported the dropout rate and 10 of these studies reported that there was no dropout.

The mean Jadad score of the Korean studies was 2.87. Of the 16 studies, seven scored 4 points but none scored 5 points (Table 2).

Three studies recruited more than 30 experimental subjects and two studies recruited more than 30 control subjects (Table 3). The inclusion criteria were indicated in 14 studies. Homogeneity tests on the demographic and diagnostic characteristics were implemented in 12 and 13 studies, respectively. Eleven studies presented their rationale for the sample size while only three analyzed the statistical power of the sample. Fifteen studies reported to have protocols for the intervention (Table 3).

DISCUSSION

The number of people using CAM is continuously increasing even though the scientific proof is not yet strong. In order to ensure that clients and health care providers can safely choose CAM as an alternative, studies on scientific proof are much needed.

Regarding the method quality of RCTs on CAM, further development is still required, and the method quality varied by the type of CAM provided. Sample size was one of the important factors that determined research quality. RCTs on CAM in this study showed the same results as previous studies in terms of method quality. When evaluating the quality of the articles for the study with the Jadad scale, the mean score was 3.37 out of 5 points and only 13 (17.1%) studies were rated 5. The mean Jadad score of 3.37 is higher than that of 2.5 obtained for 36 CAM studies in 1996. This implies that the quality of research design and reporting has been improving over the past few years (Lee, Schotland, Bacchetti, & Bero, 2002).

The Jadad scale employed in this study has been widely used as an evaluation standard (Moja et al., 2005). Linde et al. (2001) reported that the Jadad scale had better specificity than any other scale in evaluating the method quality of a research study. However, evaluation of the level of method quality by score remains controversial. For this reason, an analysis was made for each sub-item as well as the total Jadad score, and allocation concealment was added in order to supplement the limit of the Jadad scale.

Since the keen relationships between allocation concealment/double-blinding and the effect of interventions have been reported in many cases (Balk et al., 2002) and some researchers have emphasized the importance of assessing blinding methods for the purpose of decreasing selection bias (Chalmers et al., 1981), the randomization, double-blind and dropout rates and allocation concealment of studies were investigated. As a result, the randomization and dropout rates of the articles in this study were 93.4% and 90.9%, respectively. On the other hand, allocation concealment and double-blinding were 49.3% and 28.9%, respectively. Meanwhile, Linde et al.

(2001) assessed the method quality of 297 studies and reported that most of the studies did not specify the randomization, allocation concealment and dropout rates. Egger, Juni, Bartlett, Holenstein, and Sterne (2003) also analyzed 128 CAM-related studies and reported that the rates of randomization and double-blinding were higher than in Linde et al.'s study, but they were lower than in this study. There is some positive evidence to support the role of randomization and double-blinding in reducing bias (Jadad et al., 1996). Schulz, Chalmers, Hayes, and Altman (1995) analyzed 33 meta-analyses of 250 studies. Their results showed that RCTs with inappropriate concealment of therapeutic backgrounds or unclear allocation concealment tended to exaggerate the effects of therapies compared to studies with relevant concealment. Also, studies without double-blinding tended to overstate the effects of therapies compared to studies with double-blinding.

On the other hand, assessment of the quality of the research method may depend on the researcher's ability to describe the results. In some cases, limited journal space restricts the detailed description of methodology, resulting in a poor evaluation despite relevant design, implementation and analysis (Jadad et al., 1996). Therefore, it should be checked whether or not there is a tendency for well written papers to score higher for method quality in avoiding evaluation bias.

The results of this study show that the RCTs were methodologically insufficient as those in previous studies. However, most of the previous reports dealt with Western medical studies. Therefore, a direct comparison of this study with previous ones seems meaningless. An interesting trend is that recent CAM studies tend to widely use the method of double-blinding and allocation concealment, and their qualities are relatively higher than those of Western medical studies. Moher, Jadad, and Tugwell (1996) reported that the mean quality scores of CAM studies were very similar to that of Western medical studies. Pittler et al. (2000) analyzed 351 clinical controlled trials (CCT) and their results showed that the method quality scores of CAM and Western medical studies were still low, rating from 23%

to 51% of the best score. An analysis of the method quality of 484 Western medical studies and 128 CAM studies showed that randomization was higher in the Western medical studies (90%) than in the CAM studies (81%). However, the dropout rate (Western 45%, CAM 58%), double-blinding (Western 56%, CAM 90%) and allocation concealment (Western 13%, CAM 28%) were higher in the CAM-related studies compared to the Western medical studies (Egger et al., 2003).

The above discussion hardly represents the method quality of Korean CAM studies. The ratio of RCTs shows that Korean studies on CAM are still limited in supporting the effects of CAM. Of 161 Korean experimental studies, only 16 (9.9%) were RCTs and the rest were single or control group quasi-experimental designs. The mean Jadad score of the studies done in Korea was 2.87, which is lower than that of studies done in other countries. The vulnerable points in the Korean studies are low rates of randomization or double-blinding and sample sizes that are too small to satisfy the condition of normal distribution. It may be that English-language journals require a higher level of quality for CAM RCTs, so CAM researchers submit their highest quality work to English-language journals. According to Klassen, Pham, Lawson, and Moher (2005), the RCTs published in English appear to be better than RCTs published in other languages. Tang et al. (1999), who evaluated the method quality of Chinese-language RCTs, also reported many methodological problems including inadequate method of randomization, lack of blinding, small sample size and inappropriate controls.

This study had some limitations. First, the study was carried out on a sample of CAM RCTs rather than the entire population of published papers. Second, we did not include all types of CAM reports, and third, we did not include studies published prior to 2000.

In summary, the reliability of research methods has improved in spite of controversies over the scientific evidence for CAM. Our findings imply that more nursing research studies with rigorous experimental design are necessary to build up the scientific bases of CAM and to build up evidence-based nursing practice.

CONCLUSION

Studies on CAM have been increasing in number, but the research methods have many limitations, rendering study results less likely to be accepted as scientific evidence. This study analyzed 76 RCTs on CAM published between 2000 and 2005 to assess the quality of the methodology employed. Overall, the quality of the studies reviewed was generally good, but descriptions of double-blinding and allocation concealment were inadequate. According to Egger et al. (2003), allocation concealment and double-blinding are strongly related to intervention effects. Consequently, the CAM studies present insufficient scientific evidence.

Designing a RCT on CAM is so challenging that research methodologists need to be involved in the whole research process (Ernst, 2001). RCTs represent the gold standard by which health care professionals make decisions about treatment effectiveness (Wang et al., 2007). However, CAM practitioners tend to have little interest and capability in applying this scientific research method and expert researchers tend to be rather conservative when it comes to accepting CAM as a scientific intervention. Therefore, there are few RCTs on CAM with a high quality of methodology. If health providers are to make decisions about the utility of CAM interventions, they will need more information about the efficacy of the treatments. Nursing has already used various CAMs as independent nursing interventions. Thus, more high method quality RCTs on CAM in the field of nursing is needed to build up the scientific foundation for CAM. This study's results deserve further attention to improve the design and reporting of key methodological aspects of future CAM RCTs.

It is recommended that, first, CAM practitioners and nursing researchers be involved in conducting RCTs on CAM, so that nurses, patients and health providers can implement appropriate, realistic and safe treatment interventions. Second, valid and reliable

criteria for assessing the method quality of RCTs should be developed. Third, more studies on the evaluation of method quality should be conducted, and fourth, the research environment should be cultivated to recognize the value of CAM as a nursing intervention.

REFERENCES

Balk, E. M., Bonis, P. A., Moskowitz, H., Schmid, C. H., Ioannidis, J. P., Wang, C., et al. (2002). Correlation of quality measures with estimates of treatment effect in meta-analyses of randomized controlled trials. *The Journal of the American Medical Association, 287,* 2973–2982.

Barnes, D. E., & Bero, L. A. (1997). Scientific quality of original research articles on environmental tobacco smoke. *Tobacco Control, 6,* 19–26.

Barnes, J., Abbot, N. C., Harkness, E. F., & Ernst, E. (1999). Articles on complementary medicine in the mainstream medical literature: An investigation of MEDLINE, 1966 through 1996. *Archives of Internal Medicine, 159,* 1721–1725.

Chalmers, T. C., Smith, H., Blackburn, B., Silverman, B., Schroeder, B., Reitman, D., et al. (1981). A method for assessing the quality of a randomized control trial. *Controlled Clinical Trials, 2,* 31–49.

Egger, M., Juni, P., Bartlett, C., Holenstein, F., & Sterne, J. (2003). How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technology Assessment, 7,* 1–68.

Eisenberg, D. M., Davis, R. B., Ettner, S. L., Appel, S., Rompay, M. V., & Kessler, R. C. (1998). Trends in alternative medicine use in the United States, 1990–1997: Results of a follow up national survey. *The Journal of the American Medical Association, 280,* 1569–1575.

Ernst, E. (2001). Science friction—CAM on the stony road from opinion to evidence [Commentary]. *International Journal of Epidemiology, 30,* 531–532.

Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C. J., Reynolds, D. J., Gavaghan, D. J., et al. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials, 17,* 1–12.

Khan, K. S., Daya, S., & Jadad, A. R. (1996). The importance of quality of primary studies in producing unbiased systematic reviews. *Archives of Internal Medicine, 156,* 661–666.

Kim, M. H. (2003). *Use of complementary and alternative medicine and its related factors in cancer patients.* Unpublished master's thesis, Young Nam University, Deagu, Korea.

Klassen, T. P., Pham, B., Lawson, M. L., & Moher, D. (2005). For randomized controlled trials, the quality of reports of complementary and alternative medicine was as good as reports of conventional medicine. *Journal of Clinical Epidemiology, 58,* 763–768.

Lee, E. I., Shin, Y. C., Lee, J. Y., Kim, S. D., Kim, H. J., & Cho, M. S. (2002). Use of complementary and alternative medicine in cancer patients at 7 general hospitals in Seoul. *Journal of the Korean Public Health Association, 28,* 225–238.

Lee, K. P., Schotland, M., Bacchetti, P., & Bero, L. A. (2002). Association of journal quality indicators with methodological quality of clinical research articles. *The Journal of the American Medical Association, 287,* 2805–2808.

Linde, K., Jonas, W. B., Melchart, D., & Willich, S. (2001). The methodological quality of randomized controlled trials of homeopathy, herbal medicines and acupuncture. *International Journal of Epidemiology, 30,* 526–531.

Menniti-Ippolito, F., Gargiulo, L., Bologna, E., Forcella, E., & Raschetti, R. (2002). Use of unconventional medicine in Italy: A nation wide survey. *European Journal of Clinical Pharmacology, 58,* 61–64.

Moher, D., Jadad, A. R., & Tugwell, P. (1996). Assessing the quality of randomized controlled trials. Current issues and future directions. *International Journal of Technology Assessment in Health Care, 12,* 195–208.

Moja, L. P., Telaro, E., D'Amico, R., Moschetti, I., Coe, L., & Liberati, A. (2005). Assessment of methodological quality of primary studies by systematic reviews: Results of the metaquality cross sectional study. *British Medical Journal, 330,* 1053.

National Center for Complementary and Alternative Medicine. (2002). *What is CAM?* Retrieved August 2, 2007, from http://nccam.nih.gov/health/whatiscam/

Ni, H., Simile, C., & Hardy, A. M. (2002). Utilization of complementary and alternative medicine by United States adults: Results from the 1999 national health interview survey. *Medical Care, 40,* 353–358.

Oh, K. S., Kim, K. S., Kwon, S. H., & Park, J. W. (2006). Research trend of complementary and alternative medicine. *Taehan Kanho Hakhoe Chi, 36,* 721–731.

Pittler, M. H., Abbot, N. C., Harkness, E. F., & Ernst, E. (2000). Location bias in controlled clinical trials of complementary/alternative therapies. *Journal of Clinical Epidemiology, 53*, 485–489.

Raschetti, R., Menniti-Ippolito, F., Forcella, E., & Bianchi, C. (2005). Complementary and alternative medicine in the scientific literature. *Journal of Alternative and Complementary Medicine, 11*, 209–212.

Seol, K. L., Choi, S. Y., & Jong, I. L. (2002). A study on the use, understanding and satisfaction with alternative therapy for hospitalized cancer patients. *Journal of the Korean Public Health Association, 28*, 198–211.

Schulz, K. F., Chalmers, I., Hayes, R. J., & Altman, D. G. (1995). Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *The Journal of the American Medical Association, 273*, 408–412.

Son, H. M. (2002). Nurses' attitudes toward complementary and alternative therapies. *Journal of the Korean Academy of Adult Nursing, 14*, 62–72.

Tang, J. L., Zhan, S. Y., & Ernst, E. (1999). Review of randomised controlled trials of traditional Chinese medicine. *British Medical Journal, 319*, 160–161.

Trinh, K. V., Phillips, S.-D., Ho, E., & Damsma, K. (2004). Acupuncture for the alleviation of lateral epicondyle pain: A systematic review. *Rheumatology, 43*, 1085–1090.

Wang, G., Mao, B., Xiong, Z. Y., Fan, T., Chen, X. D., Wang, L., et al; CONSORT Group for Traditional Chinese Medicine. (2007). The quality of reporting of randomized controlled trials of traditional Chinese medicine: A survey of 13 randomly selected journals from mainland China. *Clinical Therapeutics, 29*, 1456–1467.