Journal Homepage:
www.jcimjournal.com/jim
www.elsevier.com/locate/issn/20954964
Available also online at www.sciencedirect.com.

• Research Article

Developing the criteria for evaluating quality of individualization in homeopathic clinical trial reporting: a preliminary study

Subhranil Saha[1], Munmun Koley[1], Subhasish Ganguly[2], Prasanta Rath[3], Pulak Roy Chowdhury[4], Seikh Intaj Hossain[5]

1. Clinical Research Unit (Homeopathy), Central Council for Research in Homeopathy, Siliguri 734006, West Bengal, India
2. Department of Organon of Medicine and Homeopathic Philosophy, D N De Homeopathic Medical College and Hospital, Kolkata 700046, West Bengal, India
3. Department of Community Medicine, National Institute of Homeopathy, Kolkata 700114, West Bengal, India
4. Upasana Diagnostic Center, Salt Lake City, Kolkata 700064, West Bengal, India
5. Paikpari Block Primary Health Center, District Purba Medinipur 721134, West Bengal, India

OBJECTIVE: This study describes the development of a preliminary version of an instrument that attempts to assess the quality of reports of individualized homeopathic prescriptions in clinical trials and observational studies.

METHODS: A multidisciplinary panel of 15 judges produced an initial version of the instrument through iterative Delphi rounds and pilot-tested the instrument on five clinical trials. Later they assessed, under blind conditions, the individualization quality of 40 randomly-selected research reports. The final version of the instrument included six criteria. These items were scored consistently by all the raters regardless of background.

RESULTS: The instrument appeared to have adequate face and content validity, acceptable internal consistency or reliability (Cronbach's α 0.606 – 0.725), significant discriminant validity (F = 398.7; P < 0.000 1), moderate interrater reliability (Fleiss κ 0.533), agreeable test-retest reliability (Cohen's κ 0.765 – 0.934), moderate sensitivity (0.4; 95% confidence interval 0.253-0.566), and high specificity (1.0; 95% confidence interval 0.891-1.000).

CONCLUSION: The initial data suggest that this instrument may be a promising systematic tool amendable for further development.

KEYWORDS: clinical trials; consensus; Delphi; homeopathy; individualization; reliability; validity

http://dx.doi.org/10.1016/S2095-4964(14)60009-1
Saha S, Koley M, Ganguly S, Rath P, Roy Chowdhury P, Hossain SI. Developing the criteria for evaluating quality of individualization in homeopathic clinical trial reporting: a preliminary study. *J Integr Med*. 2014; 12(1): 13-19.
Received December 4, 2013; accepted December 20, 2013.
Correspondence: Subhranil Saha, BHMS, MSc; E-mail: drsubhranilsaha@hotmail.com

1 Introduction

Since the inception of homeopathy, individualization of remedies has been a pivotal as well as debatable concept. Individualized (classical) homeopathy is considered by its proponents as the most effective method that adheres to original tradition[1], but is also a controversial form of

therapy[2]. In individualized homeopathy, following a conventional diagnosis, the choice of remedy is based solely on matching the patient's symptom picture with the 'remedy picture'[2]. Most of the empirical literature concerning this subject matter is ambiguous and inconclusive. Consequently, subsequent development of homeopathy evolved as isopathy, clinical homeopathy, polypharmacy, complex homeopathy, and same formula in all patients[2,3]. The 1997 meta-analysis by Linde *et al*[4] found no evidence that individualized homeopathy was superior to other forms of homeopathy.

Though previously considered to be 'methodologically more difficult' to run, and to replicate trials of individualized homeopathy in a scientifically rigorous fashion[2], in the last decade, there has been a considerable increase in the number and quality of randomized controlled trials (RCTs) of individualized homeopathic practice[5]. Systematic reviews and meta-analyses have also suggested individualization as one of the key components affecting the combined estimates of study effects. Thus, clinical trials have already pointed to the need for consensus and definition regarding 'good quality homeopathy'[2,6]. Recently, some researchers have taken initiatives to develop homeopathic prescribing indicators[7] and systematic outcome assessment tool, namely Hering's Law Assessment Tool following homeopathic intervention[8,9]. However, there exists little empirical evidence that substantiates an evaluative instrument for the quality of homeopathic individualization. To our knowledge, this is the first endeavour to do so.

2 Materials and methods

The Delphi technique is an iterative and sequential, multistage, flexible, group-communication process for forecasting and decision-making purposes to obtain informed anonymous agreement and consensus among a panel of experts in the field on a particular issue or problem[10-12]. The objective is to derive quantitative estimates through qualitative assessment of evidence[13] through a structured, well-designed multiple sequential administration of survey questionnaires augmented with continuous summary feedback of aggregated responses of a panel of experts; thereby minimizing the liabilities of individual expert decision[10-12]. Experts are qualified and experienced professionals having relevant knowledge and expertise about a particular issue or problem[10-12]. The advantages include simplicity, flexibility, in-depth anonymity, controlled feedback, statistical group response, time and cost constraints, confidentiality, objectivity, and accuracy[10]. The disadvantage is indifference of experts and consequent elimination from the panel[10-12].

2.1 Design

The first round survey was planned to commence with open-ended questions inviting free-form suggestions and recommendations from the experts, followed by subsequent iterative rounds of close-ended questions, then another final round inviting suggestions to modify the provisional instrument. Extensive efforts were made to include all the probable elements identified in the literature review. Consensus to achieve was defined 'a priori' as ≥ 85% frequency of endorsement and Fleiss' kappa (κ) observed agreement between moderate to almost perfect, i.e., 0.41 and 1.00 (poor agreement: less than 0; slight agreement 0.01-0.20; fair agreement: 0.21-0.40; moderate agreement: 0.41-0.60; substantial agreement: 0.61-0.80; and almost perfect agreement: 0.81-1.00)[14].

2.2 Expert recruitment and profiles

Participants were selected for their expertise, rather than being a sampling representative for statistical purposes. Invitation and participation in the Delphi process were completed via email outlining aim, likely time commitment and processes. Those who did not respond to the initial invitation were emailed again 5, 10, and 15 d after the initial invitation. All participants were allocated with a random identification number for reporting and collation of results. Demographic data regarding the participants' profession, qualifications, employment, designation and contact details were recorded. They were invited to provide their consent to be considered as a member of Delphi panel in presentations and publications arising from this research. All participants who accepted the invitation to participate in the Delphi process were invited to complete each and every Delphi round, regardless of participation in the previous rounds unless they withdrew from the Delphi.

In order to meet the study goal, a total of 20 experts from three relevant domains – practice, teaching, and research – were invited. Experts were chosen to ensure diverse viewpoints within the scholarly, research, and clinical perspectives. Clinicians were required to have at least 10 years of practicing experience after graduating from homeopathic schools; academicians were required to have postgraduate degrees in homeopathy and at least 10 years of teaching experience in homeopathic schools; and the researchers were required to have at least ten peer-reviewed research papers published in reputed journals. Fifteen experts, five from each background, accepted the invitation and agreed to take part in the study; others remained silent in spite of repeated reminders by telephone and email (Table 1).

2.3 First-round Delphi

Following the 'Classic Delphi' method[15], the first round was qualitative in nature. A brief preamble was provided concerning the aim of the survey, definition of key terms, likely time commitment, the plan for 3-4 rounds of input as well as the necessity of completing all rounds. As circumstances might change from that of the initial recruitment, participants were asked to contact the

Table 1 Details of judges and raters

Judge No.	Gender	Background	Assessment	Comment
1	Female	Researcher	Blind	Judge & rater
2	Male	Researcher	Blind	Judge & rater
3	Male	Clinician	Blind	Judge & rater
4	Male	Academician	Blind	Judge & rater
5	Male	Academician	Blind	Judge & rater
6	Male	Researcher	Blind	Judge & rater
7	Male	Clinician	Blind	Judge & rater
8	Female	Academician	Blind	Judge & rater
9	Female	Clinician	Blind	Judge & rater
10	Male	Researcher	Blind	Judge & rater
11	Male	Academician	Blind	Judge & rater
12	Male	Researcher	Blind	Judge & rater
13	Male	Clinician	Blind	Judge & rater
14	Male	Clinician	Blind	Judge & rater
15	Male	Academician	Blind	Judge & rater

principal investigator if they wished to withdraw and be removed from the list of participants. Thus determining the willingness of individuals to serve on the panel, the first round commenced with inviting free-form opinions and suggestions, only in English, involving the experiences and judgements from the experts, and a list of predictions and recommendations. Available published literature influenced the conceptualization of the items. Reminder to reply within seven days accompanied the first round, thereby providing seven days time to respond after which the initial round was closed. All replies were analyzed to draft the initial new instrument incorporating and compiling all the suggestions from the experts (Table 2).

2.4 Second-round Delphi

Items showing high (≥ 85%) and low (≤ 15%) level of endorsement and face validity were cross-checked against the items derived from the literature review, and did not undergo subsequent Delphi rounds evaluation. The high-scoring items were automatically included in the instrument, while the low-scoring items were eliminated. A close-ended questionnaire with 7-point Likert scale of agreement (strongly disagree: 1, disagree: 2, somewhat disagree: 3, uncertain: 4, somewhat agree: 5, agree: 6, strongly agree: 7) was developed from the items attaining frequency of endorsement between 15%-85%. All the panel experts were provided with instructions to rate their level of agreement on that Likert scale. Total number of completed response sheets was compiled after one-week interval. Rating for each item by each respondent was recorded in a data extraction sheet for computing descriptive analysis (Fleiss κ agreement, endorsement frequency percentage, mean score and standard deviation). To be considered to have met the consensus criterion, the desired level of interrater agreement (Fleiss' κ) was set at 0.41-1.00.

2.5 Third-round Delphi

Items achieving Fleiss κ between 0.41-1.00 were retained and the rest were eliminated from further evaluation. Participants were invited for any further and final suggestions to modify the instrument by

Table 2 Items considered by individual judges

No.	Suggested item	Number of suggested judges
1	Single medicine prescription when required on each occasion	15
2	Medicine individualization	15
3	Proper description of approach to medicine individualization, e.g., symptoms totality, constitutional symptoms, keynote symptoms, miasmatic symptoms, repertorial totality, consensus among prescribers.	15
4	Dose individualization	14
5	Proper description of approach to dose individualization, e.g. susceptibility analysis, nature of the used medicine, consensus among prescribers.	13
6	Subsequent prescriptions as per Kent's observations and/or Hering's law	11
7	Sufficient time and concentration in prescription	5
8	Minimal influence of leading questions in clinical decision making	4
9	Minimal dominance of experience (e.g., favourite remedies, fear of remedies) on prescription	4
10	Case recording in blank sheet	1
11	Clear and easy to understand	1
12	Brief and appropriate; no unnecessary detailing	1
13	Appropriate duration of follow-up	1
14	Correct differentiation of homeopathic aggravation, disease aggravation, and medicinal aggravation	1
15	Differential selection of repertories as per need of each individual case	1

recommending any additional items or comments, if felt necessary. All the fifteen expert panelists replied in this round. All responses were compiled again after one-week interval.

2.6 Pilot testing

Pre-testing of the provisionally finalized questionnaires was conducted by experts on five randomly selected RCTs claiming to test individualized homeopathy. The criteria for pre-testing the instrument were length, clarity, language, relevance, overall adequacy, and whether the content reflected what it purports to assess.

2.7 Field testing

The final instrument containing six criteria was implemented for rating the 40 randomly selected RCTs of individualized homeopathy, which were selected randomly from different electronic databases, namely MEDLINE/PubMed, Google Scholar, EMBASE, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Cochrane, and Central Council for Research in Homoeopathy (CCRH, India), by all the 15 experts. Ratings were done under blind conditions in order to reduce or avoid the introduction of selection bias into meta-analyses and systematic reviews[16,17].

2.8 Re-testing

After obtaining the complete responses from the experts, they were requested again to rate those 40 papers to evaluate the test-retest reliability (Cohen's κ) of the said instrument. It also controlled for chance agreement between ratings. Kappa values greater than 0.75 indicate strong agreement beyond chance, and those between 0.40 and 0.74 reflect fair to good agreement.

2.9 Testing on non-individualized homeopathy trials

To test the sensitivity and specificity of the developed instrument, 40 randomly selected papers of RCTs testing non-individualized homeopathy in its various forms were also rated as 'high' (score > 3) or 'low' (score ≤ 3) quality under blind conditions.

2.10 Statistical analysis plan

The criteria identified in the analysis were tested for internal reliability (Cronbach's α coefficient), discriminant validity (by comparing mean scores obtained by the studies using one-way analysis of variance (ANOVA)), interrater reliability (Fleiss κ), test-retest reliability (Cohen's κ), sensitivity, and specificity. Different statistical websites were used for the purpose, e.g., 'Jumk.de' for mean, standard deviation, 'Wessa.Net' for Cronbach's α, 'DfreeLon.org' for κ statistics, and 'DanielSoper' for one-way ANOVA.

3 Results

3.1 Face and content validity

Face validity refers to whether the tool accurately reflects the appropriate concepts[18]. In this study, it included the questionnaires being readable, exhibiting clarity of content, language, and being unambiguous[19]. A total of 15 items as suggested criteria by the 15 experts were compiled in the Delphi round 1 (Table 2), out of which 5 items possessed high frequency of endorsement of 86.7%-100% and seemed to possess high face validity, hence automatically included in the instrument (Table 3). Six items had endorsement frequency of only 6.7% and were eliminated from further evaluation. The remaining four items achieved 26.7%-73.3% endorsement frequency and underwent evaluation in Delphi round 2 (Table 4). Out of these four items, when evaluated with a 7-point agreement Likert scale, only a single item achieved a moderate level of observed multirater agreement (Fleiss κ = 0.581) and was included in the instrument; the remaining items had slight to fair agreement (Fleiss κ = 0.124-0.314) and were eliminated. Following Delphi round 3, inviting open suggestions from the experts, no recommended changes or new items were identified and a provisional version of the instrument was finalized. Thus the instrument was content-validated: both by referring to the available literature on the topic of research and incorporating the viewpoints of 15 expert Delphi panellists through iterative rounds.

This provisional version was pilot-tested for language and relevance by the experts on five randomly selected RCTs of individualized homeopathy. The instrument was found to be clear, with acceptable language, and relevant to the topic under investigation. After minor changes, the instrument was applied on 40 RCTs.

3.2 Reliability

Cronbach's α measure of internal consistency was computed for each of the six items. All yielded good to acceptable reliability. The results are presented in Table 4 with levels

Table 3 Initial items of the instrument having high endorsement frequency

Item	Endorsement frequency (%)
1. Was single medicine prescribed when required on each occasion?	100.0
2. Was individualized homeopathic prescribed on each occasion?	100.0
3. Was the approach to individualization of medicine adequately described?	100.0
4. Was the dose of each medicine individualized on each occasion?	93.3
5. Was the approach to dose individualization adequately described?	86.7

Table 4 Items having moderate to low face validity:

Item	Endorsement frequency (%)	Fleiss κ (observed interrater agreement)
1. Were the subsequent prescriptions generated following Kent's observations and/or Hering's law?	73.3	0.581 (moderate)
2. Was sufficient time and concentration given in prescription?	33.3	0.314 (fair)
3. Was there minimal influence of leading questions in clinical decision making?	26.7	0.133 (slight)
4. Was there minimal dominance of experience (e.g. favourite remedies, fear of remedies etc.) on prescription?	26.7	0.124 (slight)

of acceptability. These results suggested that the items of the instrument were internally consistent, satisfying the minimum criteria of 0.7 for internal consistency (Table 5).

Table 5 Internal consistency or reliability of the items

Item	Cronbach's alpha	Acceptability
1	0.689	Acceptable
2	0.711	Good
3	0.683	Acceptable
4	0.684	Acceptable
5	0.725	Good
6	0.725	Good
Total	0.606	Acceptable

3.3 Discriminant validity

The ability of the instrument to accurately gauge RCTs via ascribing 'high' and 'low' homeopathic individualization qualities was established with a one-way ANOVA of mean scores of six different items, which yielded $F = 398.7$ ($P < 0.000\ 1$ two-tailed). Thus the instrument showed its discriminant validity (Table 6).

3.4 Kappa statistics

Interrater reliability (Fleiss κ) of the items altogether was 0.533, indicating moderate agreement. Test-retest reliability (Cohen's κ) for each item was greater than 0.7, indicating that the test and retest scores were highly correlated. These results show that the instrument is reliable and reproducible (Table 7).

Table 6 Discriminant validity – relationship among the items of the instrument

Item	*n*	Score (mean ± standard deviation)
1	15	0.7 ± 0.4
2	15	0.9 ± 0.3
3	15	0.5 ± 0.5
4	15	0.4 ± 0.5
5	15	0.1 ± 0.3
6	15	0.1 ± 0.3
F score		398.7
P value		<0.000 1*

*$P<0.05$ two-tailed considered as statistically significant.

Table 7 Test-retest reliability – Cohen's κ scores of the items

Item	*n*	Test (mean ± SD)	Retest (mean ± SD)	Cohen's κ (% agreement)
1	15	0.7±0.4	0.7±0.5	0.934 (97.3)
2	15	0.9±0.3	0.9±0.3	0.790 (95.7)
3	15	0.5±0.5	0.5±0.5	0.903 (95.2)
4	15	0.4±0.5	0.4±0.5	0.900 (95.2)
5	15	0.1±0.3	0.1±0.2	0.765 (97.0)
6	15	0.1±0.3	0.1±0.2	0.839 (97.8)
Total score	15	2.7±1.4	2.6±1.3	0.841 (87.5)

SD: standard deviation.

3.5 Sensitivity and specificity

Taking into account an arbitrary cut-off score of ≥ 3 and calculating the mean scores of the 40 RCTs rated by experts, 16 trials were true positive (TP), and the rest were false negative (FN). Out of the 40 trials, all those using non-individualized homeopathy were true negative (TN). Sensitivity of the instrument was calculated to be 0.4 (95% confidence interval 0.253-0.566). Specificity was TN/(FP+TN) = 1 (95% confidence interval 0.891-1.000). Thus the instrument showed moderate sensitivity, but high specificity.

4 Discussion

The chief strength of this study is that it did not use mechanical, analytical techniques, but compiled input from a collective of experts' subjective judgements. Experts in the field of homeopathy, representing diverse backgrounds with respect to experience and practice, contributed significantly to the examination of many broad and complex issues. The heterogeneity of the experts was preserved to assure validity of the results, i.e., avoidance of domination by quantity or by strength of personality

('bandwagon effect')[20], noise, and group pressure for conformity[21]. The panel reached a consensus with one common opinion. The panel viewpoint is summarized statistically rather than only in terms of a majority vote to offset the shortcomings of conventional means of pooling opinions obtained from a group interaction. And, apart from consensus, emphasis was also on identifying differing opinions and divergent responses carried out through the rounds of Delphi. Information was shared confidentially and without personal contact between respondents; thus, the biasing effects of peer pressure, seniority, or personality were minimized. Return rate was somewhat enhanced by encouragement via personal communications.

Given that our method relied on personal opinion and input, the main problem invariably came to the correct interpretation of results. Strong-willed panel members tended to hold rigidly to their views across rounds, while less opinionated members adjusted their views. This can seriously challenge the validity and reliability of Delphi findings. The Delphi process has a tendency to create convergence, usually to a single point, although there is the possibility of polarization or clustering of the results around two or more points. The definition of "expertise" was subjective and relied upon this study's leading researchers and advisors to know of the potential experts in the field. Although the group appeared to have achieved consensus, this did not necessarily mean agreement[10]. Delphi group members who were tired or bored with the process ('respondent fatigue'), especially busy experts and overworked clinicians, might shift towards consensus to stop the process.

However, although the scientific merit of this Delphi method is a subject of debate, it is often claimed to be a robust and systematic approach of data collection[10]. Chances always remain in this type of study of ignoring disagreements so that discouraged dissenters drop out and an artificial consensus is generated. Also, judgements were those of a selected group of people and might not be representative of the overall field or population[10]. So, this tool should not be used as the only, nor a complete, solution. This method also tends to eliminate extreme positions and force a middle-of-the-road consensus[10]. The Delphi method has frequently been criticized for being scientifically untenable and over-stated[22], because of the sensitivity of results to ambiguity in the questionnaire that is used for data collection in each round, and the difficulty in assessing the degree of expertise incorporated into the forecast. The chief problems are that formal, universally agreed guidelines on the use of Delphi technique do not exist, and there is also no standardized methodology. Consequently, the design and format of the technique are flexible, which then often depend on the study's aims and objectives. Adequate opportunity was provided to the experts to raise fresh issues to avoid early closure on ideas; however, this could introduce bias by making participants feel psychologically pressured to alter their views according to the recognized literature[23]. Another potential bias was the composition of the panel that could have affected the results obtained. But there is no universally accepted criterion for selection of experts, and no guidance exists on the minimum or maximum number of experts in a panel; rather it appears to be related to common sense, funding and practical logistics. The Landis and Koch's table of interpreting Fleiss' κ is by no means universally accepted, and hence debatable. Though true anonymity is not guaranteed, the principal investigator tried to maintain confidentiality by concealing every possible link from each other.

We believe that the instrument developed in this paper, though only preliminary, is the best currently available basis for assessing quality of individualized homeopathic prescriptions in clinical trials. Though this instrument seems promising and competent, with goals of achieving more in-depth assessments, this method has limitations and may seem too crude, too formal, and depend too much on reporting; thus further valid and reliable assessments of quality in homeopathy remain elusive. However, it must also be realized that in areas like science, forecasting the degree of uncertainty is so great that exact and correct predictions are often not possible. For this, our instrument should be interpreted with caution and may require further modifications in the future while being implemented in trials. Scales have the theoretical advantage over individual markers and checklists in that they provide quantitative estimates of quality that could be replicated easily and incorporated formally into the peer review process and into systematic reviews. The main disadvantage of quality scales is that there is a dearth of evidence to support the inclusion or exclusion of items and to support the numerical scores attached to each of those items[17].

The recommendations of this study may easily be extrapolated in case reports, well-planned case series, and observational studies as well as in everyday practice settings. Authors of clinical trial publications may adhere to these criteria for good quality reporting of individualization. Newer or modified criteria to obtain more convincing and consistent results are welcome.

5 Conflict of interests

The authors declare that they have no competing interests.

REFERENCES

1 Blackstone V. Single or multiple medicine prescribing – a debate. Br Hom J. 1993; 82: 37-52.

2 Linde K, Melchart D. Randomized controlled trials of individualized homeopathy: a state-of-the-art review. J Altern Complement Med. 1998; 4(4): 371-388.
3 Klejnen J, Knipschild P, ter Riet G. Clinical trials of homeopathy. BMJ. 1991; 302(6772): 316-323.
4 Linde K, Clausius N, Ramirez G, Melchart D, Eitel F, Hedges LV, Jonas WB. Are the clinical effects of homeopathy placebo effects? A meta-analysis of placebo-controlled trials. Lancet. 1997; 350(9081): 834-843.
5 Mathie RT, Hacke D, Clausen J, Nicolai T, Riley DS, Fisher P. Randomized controlled trials of homeopathy in humans: characterising the research journal literature for systematic review. Homeopathy. 2013; 102(1): 3-24.
6 Saha S, Koley M. Homeopathic treatment of headaches and migraine: a meta-analysis of the randomized controlled trials. Asian J Pharm Clin Res. 2013; 6(Suppl. 3): 194-199.
7 Koley M, Saha S, Arya JS, Choubey G, Ghosh S, Purkait R, Mondal R, Kundu B, Mukherjee R. A study on drug utilization and prescription habits of physicians in a government homeopathic hospital in West Bengal, India. J Integr Med. 2013; 11(5): 305-313.
8 Brien SB, Harrison H, Daniels J, Lewith G. Monitoring improvement in health during homeopathic intervention. Development of an assessment tool based on Hering's Law of Cure: the Hering's Law Assessment Tool (HELAT). Homeopathy. 2012; 101(1): 28-37.
9 Harrison H. The development and use of Hering's Law Assessment Tool (HELAT) in clinical trials. Homoeopathy Research Institute (HRI) Research Article. 2012; (18). [2012-12-27]. http://www.homeoinst.org/sites/default/files/newsletters/HRI_ResearchArticle_18_Winter2012_Harrison_HELAT.pdf.
10 Yousuf MI. Using experts' opinion through Delphi technique. Practical Assessment, Research and Evaluation. 2007; 12(4). [2012-12-12]. http://pareonline.net/pdf/v12n4.pdf.
11 Kalaian SA, Kasim RM. Terminating sequential Delphi survey data collection. Practical Assessment, Research and Evaluation. 2012; 17(5). [2012-12-12]. http://pareonline.net/getvn.asp?v=17&n=5.
12 Hsu CC, Sandford BA. The Delphi technique: making sense of consensus. Practical Assessment, Research and Evaluation. 2007; 12(10). [2012-12-12]. http://pareonline.net/getvn.asp?v=12&n=10.
13 Ferri CP, Prince M, Brayne C, Brodaty H, Fratiglioni L, Ganguli M, Hall K, Hasegawa K, Hendrie H, Huang Y, Jorm A, Mathers C, Menezes PR, Rimmer E, Scazufca M; Alzheimer's Disease International. Global prevalence of dementia: a Delphi consensus study. Lancet. 2005; 366(9503): 2112-2117.
14 Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33(1): 159-174.
15 Linstone HA. The Delphi technique. Fowles RB. Handbook of future research. Westport: Greenwood. 1978: 273-300.
16 Chalmers TC, Smith H, Blackburn B, Silverman B, Schroeder B, Reitman D, Ambroz A. A method for assessing the quality of a randomized control trial. ControlIed Clin Trials. 1981; 2(1): 31-49.
17 Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, McQuay HJ. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Controlled Clin Trials. 1996; 17(1): 1-12.
18 Polit DF, Beck CT. Nursing research: generating and assessing evidence for nursing practice. 8th ed. Philadelphia: Lippincott Williams & Wilkins. 2008.
19 Rose-Grippa MK, Haber J, LoBiondo-Wood G, Gorney-Moreno MJ. Nursing research: methods, critical appraisal and utilization. 4th ed. St. Louis: Mosby. 1998.
20 Linstone HA, Turoff M. The Delphi method: Techniques and applications. Reading, MA: Addison-Wesley. 1975.
21 Dalkey NC, Helmer O. An experimental application of the Delphi method to the use of experts (Report No. RM-727-PR) (Abridged). Santa Monica, CA: the RAND Corporation. 1962.
22 Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs. 2000; 32(4): 1008-1015.
23 Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. J Adv Nurs. 2006; 53(2): 205-212.