Interobserver Reliability of Pulse Diagnosis Using Traditional Korean Medicine for Stroke Patients

Mi Mi Ko, MS,* Tae-Yong Park, OMD,* Ju Ah Lee, OMD, Tae-Young Choi, MD,
Byoung-Kab Kang, MS, and Myeong Soo Lee, PhD

Abstract

Objectives: The aim of this study was to evaluate the reliability between observers with regard to pulse signs that are observed by Traditional Korean Medicine (TKM) clinicians.

Methods: A total 658 patients with stroke who were admitted into Oriental medical university hospitals from February 2010 through December 2010 were included in this study. Each patient was seen independently by 2 experts from the same department for an examination of the pulse signs. Interobserver reliability was measured using three methods: simple percentage agreement, the κ value, and the AC₁ statistic.

Results: The κ value indicated that the interobserver reliability in evaluating the pulse signs of the subjects ranged from poor to moderate, whereas the AC₁ analysis revealed that agreement between the 2 experts was generally high (with the exception of slippery pulse). The κ value indicated that the interobserver reliability for assessing subjects who garnered the same opinion between the raters was generally moderate to good (with the exceptions of rough pulse and sunken pulse) and that the AC₁ measure of agreement between the 2 experts was generally high.

Conclusions: Pulse diagnosis is regarded as one of the most important procedures in TKM, despite the aforementioned limitations. This study reveals that the interobserver reliability in making a pulse diagnosis in stroke patients is not particularly high when objectively quantified. Additional research is needed to help reduce this lack of reliability for various portions of the pulse diagnosis.

Introduction

SIMILAR TO TRADITIONAL CHINESE medical diagnosis, traditional Korean medical diagnoses employ a unique diagnostic system, namely, pattern identification (PI) that is based on the information obtained from the following four diagnostic processes: inspection, listening and smelling, inquiry, and palpation.¹ PI is a diagnostic system that entails the comprehensive analysis of symptoms and signs with implications for determining the cause, nature, and location of the illness, the patient's physical condition, and the patient's treatment.^{2,3}

Of the four diagnostic processes, pulse diagnosis, which belongs to the palpation diagnostic processes and has been practiced for more than 2000 years, is widely regarded as a core component of the diagnostic framework of Traditional East Asian Medicine (TEAM), including Traditional Korean Medicine (TKM).^{4–6} This high regard for pulse diagnosis is rooted in the premise that the pulse assessment method is

clinically reliable for formulating a diagnosis. Using the pulse diagnosis, pathological changes in a person's body condition can be detected, after which the clinician can treat the patient.⁶ However, there is currently no direct evidence to either support or refute the aforementioned premise. This lack of evidence is because the clinical competence of performing a pulse diagnosis is dependent on the experience and knowledge of the clinician; moreover, the pulse diagnosis can be affected by a variety of factors such as emotion, activity, diet, biorhythm, and season.⁷ Thus, although many experimental studies have detected the pulse using mechanic tools in an attempt to obtain an accurate pulse diagnosis, and although studies have attempted to assess the reliability in order to standardize and objectify the pulse diagnosis,^{8–11} the majority of these studies cannot be considered scientifically and/or quantitatively reliable.

Stroke is the second most common cause of death in Korea.¹² In Korea, many stroke patients receive traditional medical care, as Korea has its own system of traditional

Brain Disease Research Center, Korea Institute of Oriental Medicine, Daejeon, Republic of Korea.

*These authors contributed equally to this work.

alternative medicine, called TKM, the role of which has been emphasized in stroke management.³ As a part of the fundamental study for the standardization and objectification of PI in TKM for Stroke (SOPi-Stroke) to develop Korean standard differentiation of the symptoms and signs (KSDSS) of stroke, a committee that is composed of experts at Oriental medicine hospitals and researchers at the brain disease research center of the Korea Institute of Oriental Medicine (KIOM) has participated in the study since 2005.^{13–15} A number of studies have shown that pulse diagnosis in particular plays important roles in both the treatment and prognosis for patients with stroke.^{16–21} The reliability of TKM pulse diagnosis was planned as one of the subdivisions of the KSDSS of stroke.

In this study, the reliability of TKM pulse diagnosis was investigated in stroke patients by evaluating the interobserver reliability in measuring the pulse sign by TKM practitioners.

Methods

The data for this study were collected as part of a multicenter study for the standardization of stroke diagnosis in Korea. Stroke patients who were admitted to nine Oriental medical university hospitals participated in this study from February 2010 through December 2010. Each patient provided informed consent to undergo procedures that were approved by the respective institutions' Institutional Review Boards (IRB). Stroke patients were enrolled within 30 days of the onset of their symptoms, provided their diagnosis was confirmed by an imaging diagnosis such as computerized tomography or magnetic resonance imaging. Patients with traumatic stroke such as subarachnoid, subdural, and epidural hemorrhage were excluded from the study. This study was approved by the IRB of the KIOM and by each Oriental medical university hospital's IRB.

Each patient was seen by 2 experts at the same department within each site. A total of 18 experts who were well trained in standard operation procedures (Appendix) were participating in this study. The experts had at least 3 years of clinical experiences with stroke after finishing regular college education in the subject of TKM for 6 years. Each patient received an examination of the status of the pulse, pulse location (i.e., floating pulse or sunken pulse), pulse rate (i.e., slow or rapid), pulse force (i.e., strong or weak), and pulse shape (e.g., string-like, slippery, fine, rough, or surging). The examination parameters were extracted from portions of a case report form for the standardization of stroke diagnosis that was developed by an expert committee organized by the KIOM. These assessments were conducted individually without discussion between the 2 experts and made on the same day without delay to minimize the time difference between the former and the latter diagnosis. Early morning was chosen to be the best time for pulse-taking, and the patients were allowed to rest at least 10 minutes for a stable pulse-taking. The grading of the severity of each variable was based on the following scores: 1=very significant, 2=significant, 3=not significant. Furthermore, as suggested by the KIOM, the clinicians were required to measure the stroke PI of each patient according to the Fire-Heat pattern, the Dampness-Phlegm pattern, the Blood Stasis pattern, the Qi Deficiency pattern, or the Yin Deficiency pattern.^{3,13–15}

A total of 658 stroke patients were enrolled in the study. Thirty (30) patients were excluded from the analysis because of a PI that was omitted by one of the 2 TKM clinicians. A total of 452 stroke patients received a PI assessment with the same resulting opinions by the raters with the following distribution: Fire-Heat pattern ($n=147$), Dampness-Phlegm pattern ($n=158$), Yin Deficiency pattern ($n=80$), and Qi Deficiency pattern ($n=66$). The Blood Stasis pattern was excluded because the sample size for this PI was too small ($n=1$) (Fig. 1).

Interobserver reliability was measured using the following three methods: simple percentage agreement, Cohen's κ coefficient and Gwet's AC₁ statistic (and their corresponding confidence intervals). The κ value is typically used to measure the level of agreement beyond that which would be expected by chance and provides a measure of interobserver reliability.¹¹ In general, definitive κ interpretations have been proposed.^{22–27} However, for most purposes, a value ≤ 0.40 represents "poor" agreement, a value between 0.40 and 0.75 represents "moderate" to "good" agreement, and a value ≥ 0.75 indicates "excellent" agreement.²⁷ The AC₁ statistic is not vulnerable to the well-known paradoxes that plague κ .^{28–30} First, interobserver reliability for pulse signs among all of the subjects was calculated using simple percentage agreement, Cohen's κ coefficient and Gwet's AC₁ statistic. Next, interobserver reliability regarding PI that had the same opinion

FIG. 1. Flow diagram of patients enrolled in the study. PI, pattern identification.

between the raters was calculated in the same way. The data were statistically analyzed with SAS software, version 9.1.3 (SAS Institute Inc., Cary, NC).

Results

The interobserver reliability results with regard to pulse signs for all of the subjects ($n=628$) are presented in Table 1. The κ value measure of agreement between the 2 experts ranged from “poor” ($\kappa=0.19$) to “moderate” ($\kappa=0.49$). In contrast, the AC_1 measure of agreement between the 2 experts was generally high for pulse signs, ranging from 0.65 to 0.93 (with the exception of slippery pulse, which had an AC_1 of 0.38). In most cases, agreement as assessed by the κ values was considerably lower than agreement as assessed by the AC_1 values.

The results of the interobserver reliability for subjects with a pattern that yielded the same opinion between the raters are presented in Table 2. The κ measure of agreement for pulse signs for the subjects of pattern generally ranged from moderate to good, with κ values ranging from 0.40 to 0.49; two exceptions were rough pulse and sunken pulse, which yielded κ values of $\kappa=0.17$ and 0.34, respectively. Moreover, the AC_1 measure of agreement between the 2 experts was generally high for pulse signs, ranging from “moderate” ($AC_1=0.41$) to “excellent” ($AC_1=0.94$).

The interobserver reliability results for the subjects of each PI are presented in Table 3. The κ measure of agreement for the subjects with the Fire-Heat pattern was generally low ($\kappa=0.18$ –0.39) with regard to pulse location (i.e., floating pulse or sunken pulse) and pulse shape, with the exception of fine pulse ($\kappa=0.55$). However, the AC_1 measure of agreement between the 2 experts was generally quite high for pulse signs, ranging from “moderate” ($AC_1=0.45$) to “excellent” ($AC_1=0.94$). In addition, the κ measure of agreement for the subjects with the Dampness-Phlegm pattern was generally low ($\kappa=0.04$ –0.38), with the exception of surging pulse ($\kappa=0.56$). However, the AC_1 measure of agreement between the 2 experts was generally quite high for pulse signs, ranging from “moderate” ($AC_1=0.46$) to “excellent” ($AC_1=0.98$). The AC_1 measure of agreement for

the subjects with the Qi Deficiency pattern was generally quite high ($AC_1=0.51$ –0.97), with the exception of weak pulse ($AC_1=0.31$).

Discussion

In TEAM, including TKM, it is generally believed that the wrist pulse conveys important information regarding an individual's health status, and the pulse diagnosis has long been used. However, the practice of pulse diagnosis has caused confusion in the modern context because there is little evidence that is based in clinical fact, and there is a shortage of precision with regard to the historical pulse literature as a reliable means for the interpretation of pulse.

Pulse diagnosis has played a prominent role in the diagnosis and subsequent treatment of stroke and has attracted increasing attention in Oriental medicine. In China, a review by Su¹⁶ discussed the important role that pulse diagnosis plays in the diagnosis and treatment of stroke by explaining the string-like pulse, slippery pulse, and fine pulse in the diagnosis of stroke. Liu¹⁷ and Cui¹⁸ reported the frequency of several pulse types in patients with stroke, including string-like pulse; string-like pulse plus moderate, fine, rapid, or slippery pulse; intermittent plus bound pulse; slippery pulse; and sunken plus fine pulse. In Korea, Cho et al.,¹⁹ in seeking important factors that affect the prognosis of stroke, observed the pulse location, pulse rate, and irregularity in 132 stroke patients within 30 days of onset. Shin et al.²⁰ used a pulse analyzer in an attempt to objectively classify pulse signs by analyzing the pulse wave in 43 stroke patients within 7 days of onset. Lee et al.²¹ analyzed the distribution of pulse indicators with regard to PI in 764 stroke patients to evaluate the value of using pulse diagnosis as an indicator for the classification of the PI in stroke patients. These results revealed a meaningful relationship between the pulse diagnosis and the PI of stroke. Kim et al.³ attempted to standardize the Oriental medical PI for stroke patients using logistic regression. Interestingly, they found that all of the patterns in their study essentially included pulse and tongue diagnosis in their final equations.

TABLE 1. AGREEMENT BETWEEN RATERS FOR ALL SUBJECTS

Variable	% Agreement	Kappa (κ)	CI of κ	AC_1	CI of AC_1
Pulse location					
Floating	77.66	0.36	(0.28, 0.44)	0.66	(0.60, 0.72)
Sunken	82.67	0.30	(0.21, 0.40)	0.77	(0.73, 0.82)
Pulse rate					
Slow	90.43	0.36	(0.24, 0.49)	0.89	(0.86, 0.92)
Rapid	80.85	0.46	(0.38, 0.53)	0.71	(0.66, 0.76)
Pulse force					
Strong	79.03	0.47	(0.40, 0.55)	0.65	(0.60, 0.71)
Weak	84.19	0.49	(0.40, 0.57)	0.77	(0.73, 0.82)
Pulse shape					
String-like pulse	78.43	0.37	(0.29, 0.46)	0.67	(0.62, 0.73)
Slippery pulse	69.00	0.38	(0.31, 0.45)	0.38	(0.31, 0.45)
Fine pulse	85.14	0.46	(0.37, 0.56)	0.79	(0.75, 0.84)
Rough pulse	93.29	0.19	(0.04, 0.34)	0.93	(0.90, 0.95)
Surging pulse	91.19	0.39	(0.26, 0.52)	0.9	(0.87, 0.93)

CI, 95% confidence interval.

TABLE 2. AGREEMENT MEASURES IN PATTERN IDENTIFICATION WITH THE SAME OPINION BETWEEN THE RATERS

Variable	% Agreement	Kappa (κ)	CI of κ	AC ₁	CI of AC ₁
Pulse location					
Floating	77.38	0.4	(0.30, 0.49)	0.64	(0.57, 0.71)
Sunken	83.81	0.34	(0.23, 0.46)	0.79	(0.74, 0.84)
Pulse rate					
Slow	91.57	0.46	(0.31, 0.60)	0.9	(0.87, 0.93)
Rapid	79.82	0.48	(0.39, 0.57)	0.68	(0.61, 0.74)
Pulse force					
Strong	78.93	0.48	(0.39, 0.57)	0.65	(0.58, 0.72)
Weak	83.59	0.49	(0.39, 0.59)	0.76	(0.70, 0.82)
Pulse shape					
String-like pulse	81.25	0.43	(0.33, 0.53)	0.72	(0.66, 0.78)
Slippery pulse	70.28	0.4	(0.32, 0.49)	0.41	(0.32, 0.49)
Fine pulse	84.15	0.47	(0.36, 0.57)	0.77	(0.72, 0.83)
Rough pulse	94.41	0.17	(0, 0.36)	0.94	(0.92, 0.96)
Surging pulse	92.15	0.47	(0.33, 0.62)	0.91	(0.88, 0.94)

CI, 95% confidence interval.

However, traditional pulse diagnosis has many limitations that stem from the clinical skill of pulse diagnosis, which depends on the clinician's experience and knowledge, and environmental factors have a large influence on the clinician's willingness to obtain diagnostic results from the pulse, which is more heavily affected by ephemeral influences than either the tongue diagnosis or other forms of diagnosis. Specifically, the pulse diagnosis can be transiently affected by emotion, pernicious influences, acute illness, severe activity, medication, diet, a full bladder, an imminent or current menstrual flow, biorhythm, the season of the year, and even the time of day. Therefore, it is essential to establish an objective diagnostic standard for pulse diagnosis among clinicians. However, there is currently little agreement among clinicians with regard to analysis.

Cole et al.* asserted that the reliability and validity of pulse diagnosis are generally poor. In contrast, King et al.¹⁰ found that when using a standardized pulse-taking procedure with clear operational definitions, the agreement between 2 practitioners was higher than 80% for 10 of 16 pulse categories. Similarly, the conclusions of various studies vary widely and include a low level of reliability of pulse diagnosis, moderate agreement, or extremely high agreement. In a general review, O'Brien et al.³¹ suggested that as the level of complexity of pulse detection increases, the reliability of pulse diagnosis decreases. The subjects in the studies regarding interobserver reliability of pulse diagnosis that were reviewed included hypercholesterolemia and cystic fibrosis patients and primarily normal groups.^{11,31} However, few studies have investigated the reliability of pulse diagnosis in patients with stroke.

The data for the present analysis were collected as part of a multicenter study of the standardization of stroke diagnosis in Korea. In this study, to evaluate interobserver reliability in the pulse status in stroke patients (which was assessed by TKM clinicians), interobserver reliability in total subjects (or subjects of pattern with the same opinion be-

tween the 2 raters) was calculated as a simple percentage agreement, κ value and AC₁ statistic. When investigating agreement between observers, clinicians have long used κ and other chance-adjusted measures together with a commonly used scale to interpret κ that was derived by Landis and Koch in 1977.²³ However, the suitability of κ as a measure of agreement has recently been debated.^{29,30} The AC₁ statistic is a relatively new measure that has been suggested by Gwet to adjust for chance in agreement studies.^{28,32}

As a result of interobserver agreement among all of the subjects, we determined that seven items had poor values, whereas four items had moderate to good values (Table 1). However, five of the seven items—including floating, slow, string-like, slippery, and surging pulse—were close to a κ value of 0.4. In particular, rough pulse had an extremely poor κ value relative to the other items but did not have a poor value in the agreement percentage or AC₁ value. It was determined that many of the clinicians checked “3=not significant” because of the difficulty in detecting a rough pulse and its low-frequency appearance. Therefore, contrary to the κ value, the agreement percentage and AC₁ values were high (93.29% and 0.93, respectively). In the interobserver agreement for the subjects who were classified into the same pattern between the raters, slightly higher κ values were observed. Only 2 of the 11 items had poor values, and the others had moderate to good values. However, the rough pulse still had an extremely poor value (Table 2). In the interobserver agreement of each pattern, the 4 of the 11 items that had moderate to good values included the Fire-Heat pattern, one Dampness-Phlegm pattern, six Qi Deficiency patterns, and five Yin Deficiency patterns (Table 3).

Pulse diagnosis has historically been regarded as one of the most important procedures in TKM, despite the limitations discussed above. The current study shows that the interobserver reliability of pulse diagnosis in stroke patients is poor when objectively quantified. Additional research is needed to help reduce this lack of reliability for various portions of the pulse diagnosis through detailed-oriented criteria and better training of the clinicians. The authors believe that the results of this study will be useful to clinicians in diagnosing stroke.

*Cole P. Acupuncture and pulse diagnosis in Great Britain [unpublished Ph. D. thesis]. University of Sussex, 1975.

TABLE 3. AGREEMENT MEASURES IN EACH PATTERN WITH THE SAME OPINION BETWEEN THE RATERS

Variables	PI	% Agreement	Kappa (κ)	CI of κ	AC ₁	CI of AC ₁	PI	% Agreement	Kappa	CI of κ	AC ₁	CI of AC ₁
Pulse location	FH (N=147)						DP (N=158)					
Floating		72.11	0.34	(0.19, 0.50)	0.52	(0.38, 0.67)		79.11	0.36	(0.19, 0.53)	0.69	(0.58, 0.80)
Sunken		89.12	0.33	(0.08, 0.58)	0.87	(0.80, 0.94)		80.38	0.22	(0.06, 0.38)	0.74	(0.65, 0.84)
Pulse rate												
Slow		93.88	0.37	(0.04, 0.70)	0.93	(0.89, 0.98)		1.14	0.32	(0.06, 0.59)	0.9	(0.84, 0.95)
Rapid		77.55	0.52	(0.37, 0.66)	0.58	(0.45, 0.72)		79.74	0.25	(0.09, 0.42)	0.73	(0.63, 0.83)
Pulse force												
Strong		73.47	0.47	(0.32, 0.61)	0.47	(0.33, 0.62)		76.58	0.38	(0.21, 0.54)	0.63	(0.50, 0.75)
Weak		93.88	0.61	(0.37, 0.84)	0.93	(0.88, 0.98)		84.18	0.3	(0.1, 0.51)	0.8	(0.71, 0.88)
Pulse shape												
String-like pulse		73.79	0.39	(0.24, 0.55)	0.55	(0.40, 0.69)		84.81	0.29	(0.08, 0.50)	0.81	(0.73, 0.89)
Slippery pulse		70.75	0.38	(0.23, 0.53)	0.45	(0.30, 0.60)		67.09	0.19	(0.03, 0.34)	0.46	(0.31, 0.61)
Fine pulse		93.1	0.55	(0.30, 0.80)	0.92	(0.87, 0.97)		88.61	0.04	(0, 0.23)	0.87	(0.81, 0.93)
Rough pulse		94.48	0.18	(0, 0.51)	0.94	(0.90, 0.98)		94.3	0.15	(0, 0.47)	0.94	(0.90, 0.98)
Surging pulse		79.31	0.36	(0.18, 0.54)	0.7	(0.58, 0.81)		98.09	0.56	(0.11, 1)	0.98	(0.96, 1)
Pulse location	QD (N=66)						YD (N=80)					
Floating		81.82	0.49	(0.24, 0.74)	0.72	(0.55, 0.89)		80	0.47	(0.26, 0.68)	0.69	(0.53, 0.85)
Sunken		80.3	0.51	(0.28, 0.73)	0.68	(0.50, 0.86)		83.75	0.34	(0.05, 0.63)	0.78	(0.66, 0.91)
Pulse rate												
Slow		86.36	0.6	(0.37, 0.84)	0.79	(0.65, 0.93)		92.5	0.36	(0, 0.77)	0.92	(0.84, 0.99)
Rapid		86.36	0.5	(0.21, 0.78)	0.81	(0.69, 0.94)		78.75	0.53	(0.35, 0.72)	0.62	(0.44, 0.80)
Pulse force												
Strong		87.88	0.36	(0.004, 0.7)	0.85	(0.74, 0.96)		86.25	0.49	(0.23, 0.74)	0.81	(0.70, 0.93)
Weak		65.15	0.3	(0.07, 0.53)	0.31	(0.08, 0.55)		78.75	0.46	(0.24, 0.68)	0.65	(0.48, 0.82)
Pulse shape												
String-like pulse		86.15	0.49	(0.20, 0.78)	0.81	(0.68, 0.94)		83.75	0.55	(0.35, 0.76)	0.75	(0.61, 0.89)
Slippery pulse		72.73	0.39	(0.15, 0.63)	0.51	(0.29, 0.73)		73.75	0.32	(0.11, 0.53)	0.58	(0.4, 0.77)
Fine pulse		73.85	0.45	(0.24, 0.67)	0.51	(0.29, 0.73)		67.5	0.37	(0.19, 0.56)	0.36	(0.15, 0.57)
Rough pulse		92.31	-0.02	(0, 0.02)	0.92	(0.84, 0.99)		96.25	0.39	(0, 0.93)	0.96	(0.91, 1)
Surging pulse		96.92	0	(0, 0)	0.97	(0.92, 1)		100	1	(1, 1)	1	(1, 1)

PI, pattern identification; CI, 95% confidence interval; FH, Fire-Heat pattern; DP, Dampness-Phlegm pattern; QD, Qi Deficiency pattern; YD, Yin Deficiency pattern.

Conclusions

Pulse diagnosis is regarded as one of the most important procedures in TKM, despite the aforementioned limitations. In this study, to evaluate interobserver reliability in the pulse status in stroke patients who were assessed by TKM clinicians, interobserver reliability was calculated as a simple percentage agreement, κ value, and AC₁ statistic. This study reveals that the interobserver reliability in making a pulse diagnosis in stroke patients is not particularly high when objectively quantified. Additional research is needed to help reduce this lack of reliability for various portions of the pulse diagnosis.

Acknowledgment

This research was supported by a grant from the Korea Institute of Oriental Medicine (K11131).

Disclosure Statement

No competing financial interests exist.

References

1. Teng TT, Deng T, Ergil K. Practical Diagnosis in Traditional Chinese Medicine. Elsevier Health Sciences, Edinburgh: Churchill Livingstone, 1999.
2. World Health Organization, ed. WHO International Standard Terminologies on Traditional Medicine in the Western Pacific Region. Geneva: WHO Western Pacific Regional Office, 2007.
3. Kim HJ, Bae HS, Park SU, et al. Clinical Approach to the standardization of Oriental medical diagnostic pattern identification in stroke patients. *Evid Based Complement Altern Med* 2011;2011:768492.
4. Li SZ, Seifert GM. Pulse Diagnosis. Brookline, MA: Paradigm Publications, 1985.
5. Flaws B. The Secret of Chinese Pulse Diagnosis. Boulder, CO: Blue Poppy Press, 1995.
6. Walsh S, King E, Simpson G. Pulse diagnosis: An introductory guide for the experienced practitioner. *J Chin Med* 2009;89:5-12.
7. Hammer L. Contemporary pulse diagnosis: Introduction to an evolving method for learning an ancient art. Part 1. *Am J Acupunct* 1993;21:123-139.
8. Lee J. The study on the intellectual analysis algorithm for oriental pulse parameters. *J Med Syst* 2007;31:345-349.
9. Lee J. The systematical analysis of oriental pulse waveform: A practical approach. *J Med Syst* 2008;32:9-15.
10. King E, Cobbin D, Walsh S, et al. The reliable measurement of radial pulse characteristics. *Acupunct Med* 2002;20:150-159.
11. O'Brien KA, Abbas E, Zhang J, et al. Understanding the reliability of diagnostic variables in a Chinese medicine examination. *J Altern Complement Med* 2009;15:727-734.
12. National Statistical Office in Korea. Annual Report on the Cause of Death Statistics (based on vital registration). Daejeon: National Statistical Office in Korea, 2009.
13. Lee JA, Lee JS, Kang BK, et al. Report on the Korean standard pattern identifications for stroke. *Korean J Orient Int* 2011;32:232-242.
14. Go HY, Kim YK, Kang BK, et al. Report on the Korean standard differentiation of the symptoms and signs for the stroke-2. *Korean J Orient Physiol Pathol* 2006;20:1789-1791.
15. Kim JK, Seol IC, Lee I, et al. Report on the Korean standard differentiation of the symptoms and signs for the stroke-1. *Korean J Orient Physiol Pathol* 2006;20:229-234.
16. Su N. Pulse in diagnosis and treatment of guiding significance for stroke. *J Liaoning Univ TCM* 2010;12:46-47.
17. Liu DH. Observed of pulse for acute stroke of 273 cases. *Fujian J Trad Chin Med* 1989;20:50-52.
18. Cui JC. The principle of stroke pulse holographic. *Gianjin J Trad Chin Med* 1990;4:31-32.
19. Cho EH, Kwon JN, Kim YK. A clinical study on the functional outcome in stroke. *Korean J Orient Int* 2001;22:647-657.
20. Shin BC, Lee SH. Clinical study on CVA patients by electro-pulse graph. *Korean J Orient Med Pathol* 1999;13:123-127.
21. Lee JS, Kang BK, Ko MM, et al. Study on the pulse diagnosis for pattern identifications in stroke patients. *Korean J Orient Physiol Pathol* 2008;22:1378-1382.
22. Fleiss JL. Statistical Methods for Rates and Proportions. 2nd ed. New York: Wiley, 1981.
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-174.
24. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess* 1994;6:284-290.
25. Altman DG. Practical Statistics for Medical Research. London: Chapman & Hall/CRC, 1991.
26. Shrout PE. Measurement reliability and agreement in psychiatry. *Stat Methods Med Res* 1998;7:301-317.
27. Jelles F, Van Bennekom CA, Lankhorst GJ, et al. Inter-and intra-rater agreement of the Rehabilitation Activities Profile. *J Clin Epidemiol* 1995;48:407-416.
28. Gwet K. Computing inter-rater reliability with the SAS system. *Stat Methods Inter-rater Reliability Assess* 2002;3:1-16.
29. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990;43:551-558.
30. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990;43:543-549.
31. O'Brien KA, Birch S. A review of the reliability of traditional East Asian medicine diagnoses. *J Altern Complement Med* 2009;15:353-366.
32. Gwet K. Handbook of Inter-rater Reliability. Gaithersburg, MD: STATAXIS Publishing Company, 2001:223-246.

Address correspondence to:

Myeong Soo Lee, PhD

Brain Disease Research Center

Korea Institute of Oriental Medicine

1672 Yuseongdae-ro

Yuseong-gu Daejeon 305-811

Korea

E-mail: drmslee@gmail.com