2012. Vol.33. No.4. 1-8
The Journal of Korean Oriental Medicine

Original Article

Basic Tongue Diagnosis Indicators for Pattern Identification in Stroke Using a Decision Tree Method

Ju Ah Lee[1], Jungsup Lee[2], Mi Mi Ko[1], Byoung-Kab Kang[1], Myeong Soo Lee[1]

[1]Medical Research Division, Korea Institute of Oriental Medicine, Daejeon, South Korea
[2]Korean National Rehabilitation Center, Seoul, South Korea

Objectives: The purpose of this study was to specify major tongue diagnostic indicators and evaluate their significance in discriminating pattern identification subtypes in stroke patients.
Methods: This study used a community based multi-center observational design. Participants (n=1,502) were stroke patients admitted to 11 oriental medical university hospitals between December 2006 and February 2010. To determine which tongue indicator affected each pattern identification, a decision tree analysis of the chi-square automatic interaction detector (CHAID) algorithm was performed. The chi-squared test was used as the criterion in splitting data with a p-value less than 0.05 for division, which is the main procedure for developing a decision tree. The minimum sample size for each node was specified as n =10, and branching was limited to two levels.
Results: From the 9 tongue diagnostic indicators, 6 major tongue indicators (red tongue, pale tongue, yellow fur, white fur, thick fur, and teeth-marked tongue) were identified through the decision tree analysis. Furthermore, each pattern identification was composed of specific combinations of the 6 major tongue indicators.
Conclusions: This study suggests that the 6 tongue indicators identified through the decision tree analysis can be used to discriminate pattern identification subtypes in stroke patients. However, it is still necessary to re-evaluate other pattern identification indicators to further the objectivity and reliability of traditional Korean medicine.

Key Words : stroke, pattern identification, tongue diagnosis, tongue indicators, decision tree analysis

Introduction

Stroke is a primary cause of death in Korea; in 2009, 10.5 percent of deaths were due to cerebrovascular disease[1)]. Korea has its own system of traditional medicine, known as Traditional Korean Medicine (TKM)[2)], and this is the type of treatment many stroke patients receive. TKM utilizes a unique diagnostic system called pattern identification (PI). PI is characterized by its own theoretical base and practice. This system entails the comprehensive analysis of symptoms and signs with implications for determining the cause and nature of the illness, the patient's physical condition, and the patient's treatment plan using four types of examination (inspection, listening and smelling, inquiry, and palpation)[3,4)].

The inspection process involves an examination of the patient's symptoms or disease by assessing the tongue, among other things[5)]. Tongue observation is an important procedure in diagnosing through inspection[6)]. A number of studies have shown that tongue diagnosis plays an important role in the clinical prognosis and

・Received : 2 July 2012 ・Revised : 15 October 2012 ・Accepted : 15 October 2012
・Correspondence to : Mi Mi Ko
Medical Research division, Korea Institute of Oriental Medicine, 1672 Yuseongdae-ro, Yuseong-gu, Daejeon 305-811, Republic of Korea
Tel : +82-42-868-9493, Fax : +82-42-863-9299, Email : kmm1836@kiom.re.kr

treatment[7-10], specifically in patients with a history of stroke[7,11,12]. However, the clinical value of tongue diagnoses is determined by the experience and knowledge of the clinicians performing the diagnosis. Unfortunately, the practice of traditional tongue diagnosis has not been verified scientifically (particularly quantitatively).

Since 2005, the Korea Institute of Oriental Medicine (KIOM) has been conducting a study of PI in stroke using tongue diagnosis[13-15]. To evaluate the clinical use of tongue diagnosis, tongue indicators for each PI were enumerated, and the combination of tongue indicators for each PI were determined using a decision tree method for data mining.

Methods

Ethics Statement

This study was approved by the Institutional Review Board (IRB) of each hospital. Informed consent was obtained from all study patients after a thorough explanation of the study details. The personal information of patients was managed by standard protocol, conducted by International Conference on Harmonisation (ICH) and Korea Good Clinical Practice (KGCP). There were inclusion criteria for participants.

Participants

This study used a community-based multi-center observational design. Participants were stroke patients admitted to 11 oriental medical university hospitals in the National Capital Region between December 2006 and February 2010. We examined the stroke patients within 30 days after their ictus. Inclusion criteria included acute stroke with a neurological deficit that persisted for over 24 hours; a final diagnosis of stroke confirmed by imaging, including computerized tomography and magnetic resonance imaging; and agreeing to participate within 30 days of stroke onset. In addition, there were exclusion criteria. Exclusion criteria included traumatic stroke, including epidural hemorrhage and subdural hemorrhage; degenerative brain disease; a stroke concomitant with a brain tumor; and patients unable to communicate symptoms.

Measured Variables

We used the Case Report Form (CRF) and Standard Operation Procedures (SOPs) developed by the KIOM. These included gathering general patient information, such as the presenting diagnosis, the Trial of ORG 10172 in Acute Stroke Treatment (TOAST) classification of the stroke, medical history, smoking habits, and the ratings from the four inspections. The purpose of the CRF is to discriminate between patterns of stroke in TKM. We translated TKM terms into English based on the guidelines suggested by the World Health Organization (WHO).

PI determination

Two TKM medical doctors, one from each hospital and who both had more than three years of clinical experience with stroke patients, identified the pattern information for each patient from the following patterns: fire-heat, dampness-phlegm, *qi* deficiency, *yin* deficiency and blood stasis pattern. However, blood stasis pattern was excluded because the sample size for this pattern was too small. Concordance between the two experts constituted confirmation of the pattern.

Collection of tongue indicator information

All patients were seen by two experts trained in using the SOPs regarding the tongue examination status, including tongue color (pale, red), fur color (white fur, yellow fur), fur quality (thick fur, dry fur), and special appearances of the tongue (teeth-marked, enlarged, mirror). Assessments were provided individually without discussion between experts. The tongue status was graded on the severity of appearance using the following scale: 1 = very much; 2 = somewhat; and 3 = not much. These ratings comprised the stroke PI for each patient.

Data processing and statistical analysis

Information from a total of 1,502 stroke patients was collected and analyzed for this study. To compare the demographic and clinical characteristics for PI between patients, analyses of variance (ANOVA) or Kruskal-Wallis rank-sum tests were used for continuous variables, and chi-squared tests (or Fisher's exact tests) were used for categorical

Table 1. Demographic parameters of study subjects

Characteristics	QD	DP	YD	FH	*p*
N	313	542	207	440	
anthropometric characteristics					
sex (M/F)	107/206	245/297	91/116	334/106	<.0001
age (year)	67.36±11.46	66.61±11.22	69.16±12.47	65.76±11.89	0.0030
WHR	0.93±0.11	0.95±0.08	0.93±0.15	0.95±0.10	NS
waist circumference (cm)	84.57±8.81	88.93±9.51	82.77±9.16	88.33±9.34	0.0005
smoking (none/stop/active)	232/40/41	348/65/128	135/32/40	163/107/170	<.0001
drinking (none/stop/active)	214/31/68	337/46/157	130/19/58	171/55/124	<.0001
TOAST classification					
LAA	51	114	47	117	0.0040
CE	20	26	15	34	
SVO	182	322	95	199	
SOE	7	10	7	5	
SUE	14	22	10	22	
medical history					
Hypertension (n, %)	180 (58.25)	342 (63.33)	123 (59.42)	258 (58.90)	NS
Hyperlipidemia (n, %)	35 (11.33)	64 (11.87)	18 (8.91)	54 (12.62)	NS
Diabetes (n, %)	82 (26.45)	154 (28.57)	47 (22.93)	116 (26.73)	NS
Ischemic Heart diseases (n, %)	17 (5.52)	29 (5.39)	12 (5.88)	26 (6.03)	NS
Tongue indicators					
pale (n, %)	114 (36.42)	169 (31.18)	24 (11.59)	38 (8.64)	<.0001
red (n, %)	46 (14.70)	85 (15.68)	94 (45.41)	224 (50.91)	<.0001
white fur (n, %)	175 (55.91)	357 (65.87)	69 (33.33)	136 (30.91)	<.0001
yellow fur (n, %)	48 (15.34)	134 (24.72)	53 (25.60)	216 (49.09)	<.0001
thick fur (n, %)	68 (21.73)	240 (44.28)	34 (16.43)	201 (45.68)	<.0001
dry fur (n, %)	34 (10.86)	74 (13.65)	68 (32.85)	110 (25.00)	<.0001
teeth marked (n, %)	64 (20.45)	92 (16.97)	21 (10.14)	40 (9.09)	<.0001
enlarged (n, %)	48 (15.34)	128 (23.62)	26 (12.56)	51 (11.59)	<.0001
mirror (n, %)	8 (2.56)	5 (0.92)	26 (12.56)	15 (3.41)	<.0001

All results except age, WHRandwaistcircumferenceareexpressedasfrequencies(percentages)forcategoricalvariables.QD:*Qi* deficiency pattern. DP: Dampness-phlegm pattern. YD: *Yin* deficiency pattern. FH: Fire-heat pattern. WHR:waisthipratio.TOAST: Trial of ORG10172inAcuteStrokeTreatment.LAA:large-arteryatherosclerosis.CE:cardioembolism.SVO:small-vesselocclusion.SOE:strokeofotheretiology.SUE:strokeofundeterminedetiology.P-value calculated by ANOVA test in continuous variables and chi-squared test in categorical variables. NS: not significant.

variables. Continuous variables were tested for normality using the Kolmogorov-Smirnov test. All tests of significance were 2-tailed, and *p*-values of <0.05 were considered statistically significant. To determine which tongue indicator affected each PI, a decision tree analysis of the chi-square automatic interaction detector (CHAID) algorithm was performed. The severity grading of each variable was converted to 'signs of' and 'no signs of' to be used as variables in the decision tree analysis. The chi-squared test was used as the criterion in splitting data with a p-value less than 0.05 for division, which is the main procedure for developing a decision tree. The minimum sample size for each node was specified as n =10, and branching was limited to two levels. Analyses were performed using SAS version 9.1.3 with Enterprise Miner (SAS Institute Inc., Cary, NC, USA).

Results

General characteristics of each PI

General patient characteristics are shown in Table 1. Results indicated that among the 1,502 patients, 542 had the dampness-phlegm pattern, 440 had the fire-heat pattern, 313 had the *qi* deficiency pattern, and 207 had the *yin* deficiency pattern.

The frequency of ‘red tongue’ in the fire-heat and *yin* deficiency patterns was higher than in other patterns (50.91%, 45.41%, respectively). The frequency of ‘white fur’ in the dampness-phlegm and *qi* deficiency patterns was higher than in other patterns (65.87%, 55.91%, respectively). The frequency of ‘yellow fur’ in the fire-heat pattern was higher than in other patterns (49.09%), and ‘thick fur’ in the fire-heat and dampness-phlegm patterns was higher than in other patterns (45.68%,

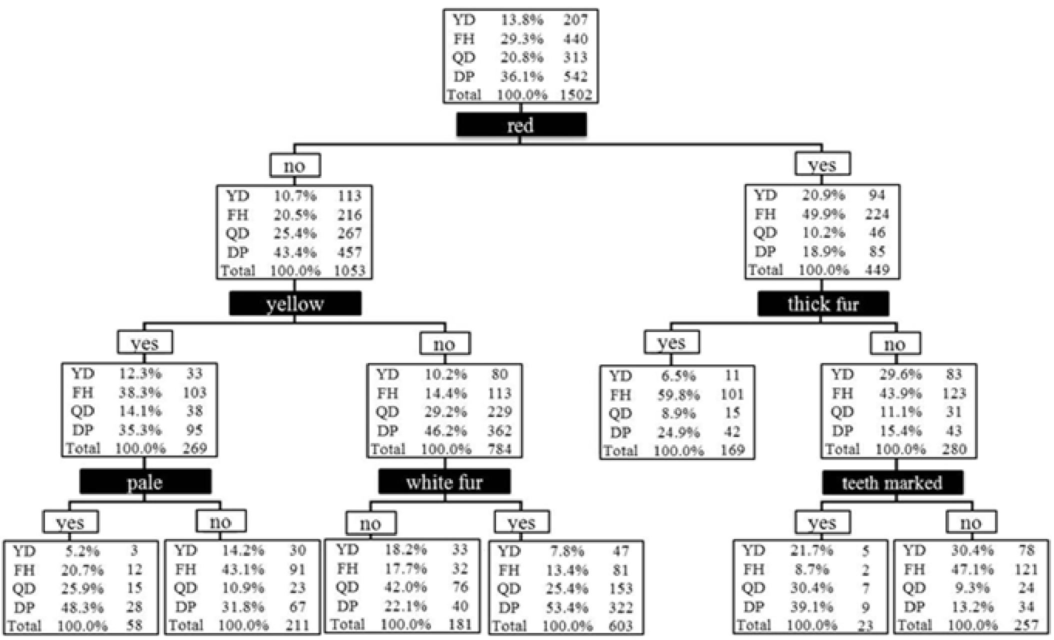

Fig. 1. Decision tree model using CHAID algorithm. FD. Fire-heat pattern; DP, Dampness-phlegm; QD, *Qi* deficiency; YD, *Yin* deficiency

Table 2. Importance of input variables

Variable	Importance	Role	Rules
Red	1.0000	input	1
White fur	0.5838	input	1
Yellow fur	0.5711	input	1
Thick fur	0.4054	input	1
Teeth marked	0.3155	input	1
Pale	0.2953	input	1

44.28%, respectively). However, the frequencies of ‘dry fur’, ‘enlarged’, ‘teeth-marked’, and ‘mirror tongue’ were relatively low.

Extraction of tongue indicators by the decision tree analysis

Among 9 tongue indicators, 6 (red tongue, pale tongue, yellow fur, white fur, thick fur, and teeth-marked tongue) were identified by the CHAID algorithms of the decision tree (Fig. 1). Using the results of the decision tree analysis, the probability of being classified as the fire-heat pattern increased from 29.3% to 49.9% if ‘red tongue’ was indicated and to 59.8% if ‘thick fur’ was indicated. In addition, the probability of being classified as the dampness-phlegm pattern increased from 36.1% to 43.4% if ‘no red tongue’ was indicated and to 53.4% if ‘no yellow fur’ and ‘white fur’ was indicated. The probability of being classified as the *qi* deficiency pattern increased from 20.8% to 25.4% if ‘no red tongue’ was indicated and to 42.0% if ‘no yellow fur’ and ‘no white fur’ was indicated. Finally, the probability of being classified as the *yin* deficiency pattern increased from 13.8% to 20.9% if ‘red tongue’ was indicated and to 30.4% if ‘no thick fur’ and ‘no teeth-marked’ was indicated.

The importance of the presence of the independent variables on the stroke PI is described in Table 2. 'Red tongue' was most important, followed by ‘white fur’, ‘yellow fur’, ‘thick fur’, ‘teeth-marked’ and ‘pale tongue’, suggesting that 'red tongue' plays an important role in building the classification model.

The classification results showing classification accuracy and error rate are summarized in Table 3. The overall classification accuracy was estimated at 40.3%. The classification accuracies of the fire-heat and dampness-phlegm patterns were 71% and 66%, respectively; therefore, these patterns presented better accuracy than the other patterns. The classification accuracies of the *qi* deficiency and *yin* deficiency patterns, however, were 24% and 0%, respectively.

Table 3. Results of classification table using decision tree analysis

		Predicted Value				
		FH	DP	QD	YD	Sum
Real Value	FH	313	95	32	0	440
	DP	143	359	40	0	542
	QD	62	175	76	0	313
	YD	119	55	33	0	207
	Sum	637	684	181	0	1502
Accuracy		71	66	24	0	40.3
Error rate		29	34	76	100	59.7

FD, Fire-heat pattern; DP, Dampness-phlegm; QD, *Qi* deficiency; YD, *Yin* deficiency; Accuracy (%) = (Predicted value/Real Value)*100.

Discussion

Tongue diagnosis is a very important method used to determine the overall condition of patients. The status of the tongue is an important indicator in diagnosing a patient's condition, including the physiological and clinicopathological changes of internal organs3. It is possible to diagnose organ disease and *qi* blood deficiency/excess using tongue inspection because the tongue is connected to organs through meridians[16]. In tongue diagnoses, it is necessary to separately observe the tongue body (tissues including muscle and blood vessels) and tongue fur (the coating on the tongue that appears fuzzy)[5]. Recently, many studies on tongue diagnosis have been conducted in China. One of these was an objectified study on the tongue images of patients with lung cancer[17]; in another study, Chinese medical specialists diagnosed tongue color using the Delphi method8. Current research is also investigating agreement between experts. However, there is little research on tongue diagnoses that affect PIs or whether discrimination between tongue indicators could help in PI.

Decision tree analysis is a method of data mining[18]. Because the process of prediction and classification is represented as a tree, researchers can easily understand and explain the comparative processes with discriminant analyses, regression, and neural networks. Decision tree analysis can be seen as a process of questioning, beginning with an initial question that assists in leading to the next question. CHAID is an example of a decision tree method that develops separate branches using chi-square statistics with categorical variables. It can be used for classification or prediction, but it is most useful in explaining the analysis process rather than the accuracy of the analysis. Thus, it is a useful tool for analyzing the process of discriminating PI subtypes using tongue indicators. In this study, all patients were seen by two experts well trained in using the SOPs to examine the tongue status, including tongue color (pale, red), fur color (white fur, yellow fur), fur quality (thick fur, dry fur), and special tongue appearances (teeth-marked, enlarged, mirror). We extracted tongue indicators using a decision tree analysis to control the combination of items.

6 tongue indicators were extracted and used as the decision tree branches (Fig. 1): red and pale tongue color, yellow and white fur color, thick fur quality, and a teeth-marked tongue in special tongue appearances. By examining the decision tree structure, many cases were classified as the fire-heat or *yin* deficiency pattern when red tongue was present. When the tongue was not red and yellow fur was not present, cases were classified as dampness-phlegm or *qi* deficiency patterns. It is thought that red tongue represents the nature of fever, which is represented in both the fire-heat and *yin* deficiency patterns. White fur combined with pale tongue or teeth-marked tongue indicates the *qi* deficiency pattern, while thick fur combined with the white fur indicates the dampness-phlegm pattern. The two patterns contain similar fur color, but there are differences in the color of the tongue body and presence of teeth-marks. Therefore, the tongue and fur color appear to be important indicators in distinguishing patterns in clinical presentations. These results lead us to conclude that tongue diagnosis can play a significant role in the differential diagnosis of several PIs. For example, diagnosing the fire-heat pattern is easier when the patient has a red tongue; the TKM doctor can ascertain the tongue and fur color. This can increase the diagnosis accuracy and simplify the complex process of determining PIs by checking important tongue indicators.

Other diagnostic methods of TKM are thought to be even more subjective than tongue

Table 4. PI of stroke algorithms using tongue indicators

PI	Algorithms	Classification prob.(%)
FH	If 'red': yes & 'thick fur': yes	59.8
	If 'red': yes	49.9
YD	If 'red' : yes & 'thick fur' : no & 'teeth marked' :no	30.4
	If 'red' : yes & 'thick fur' : no	29.6
DP	If 'red' : no & 'yellow fur' : no & 'white fur' : yes	53.4
	If 'red' : no & 'yellow fur' : yes & 'pale' : yes	48.3
QD	If 'red' : no & 'yellow fur' : no & 'white fur' : no	42.0
	If 'red' : yes & 'thick fur' : no & 'teeth marked' : yes	30.4

PI, Pattern identification; FD, Fire-heat pattern; DP, Dampness-phlegm; QD, *Qi* deficiency; YD, *Yin* deficiency

diagnosis[19]. This is only one reason that the decision tree model using tongue indicators is important. However, if further analysis is performed to collect additional information such as four types of examination (inspection, listening and smelling, inquiry, and palpation) accuracy should be improved. We developed the PIs of stroke algorithms using tongue indicators (Table 4), and with the results of this study, we were able to reduce the number of tongue indicators of stroke that should be taken into consideration from 9 to 6. In addition, we were able to obtain the major tongue indicators in PI using decision tree analysis.

It is therefore important to systematically re-evaluate other indicators of PI to further the objectivity and science of TKM.

Acknowledgement

This research was supported by a grant from the Korea Institute of Oriental Medicine (K12130).

Disclosure Statement

The authors declare that they have no financial interests to disclose.

References

1. National Statistical Office in Korea. Cause of death statistics in Korea. Seoul: National Statistical Office in Korea. 2009.
2. Kim HJ, Bae HS, Park SU, Moon SK, Park JM, Jung WS. Clinical approach to standardization of oriental medical diagnostic pattern identification in stroke patients. *eCAM*. 2011;Article ID 768492:7pages.
3. Wiseman N, Ye F. A practical dictionary of Chinese medicine. Second edition. Brookline, Massachusetts: Paradigm Publications. 1998.
4. WHO International Standard Terminologies on Traditional Medicine in the Western Pacific Region. Geneva: WHO. 2007:80.
5. Lee BK, Park YB, Kim TH. Traditional Korean medicine diagnostics. Sixth Edition. Seongnam: Seongbosa. 2004.
6. Choi DJ, Park SU, Moon SK, Jo KH, Kim YS, Bae HS, et al. The characteristics of tongue inspection and relationship between tongue inspection and differentiation of syndrome. *J Korean Oriental Med*. 1999;20(2):387-99.
7. Chen XF, Gong WJ. Analysis of 123 patients with acute stroke of tongue. *Chinese Journal of Integrative Medicine on Cardio-/ Cerebrovascular Disease*. 2007;5(12):1273.
8. Yue XQ, Liu Q, Deng WZ. Fitting study on Chinese medical specialist's diagnosis of tongue color by Delphi evaluation. *Jounal of traditional Chinese medicine*. 2007;48(7):635-7.

9. He YT, Zha QL, Yu JP, Tan Y, Lu C, Lu AP. Principal factor analysis of symptoms of rheumatoid arthritis and their correlations with efficacy of traditional Chinese medicine and Western medicine. *Zhong Xi Yi Jie He Xue Bao*. Jan 2008;6(1):32-6.
10. Anastasi JK, Currie LM, Kim GH. Understanding diagnostic reasoning in TCM practice: tongue diagnosis. *Altern Ther Health Med*. May-Jun 2009;15(3):18-28.
11. Park SW, Kang KW, Kang BK, Kim JC, Kim BY, Ko MM, et al. Study of tongue diagnosis for pattern identification in stroke patients. *Korean J of Oriental Physiology & Pathology*. 2008;22:262-6.
12. Zheng ZL, Fu YQ. Observed in patients with tongue acute cerebral hemorrhage. *HAI NAN MEDICAL JOURNAL*. 2010;21(9):94-7.
13. Kim JK, Seol IC, Lee I, Jo HK, Yu BC, Choi SM. Report on the Korean standard differentiation of the symptoms and signs for the stoke-1. *Korean J of Orient Physiology & Pathology*. 2006;20:229-234.
14. Go HY, Kim JK, Kang BK, Kim BY, Ko MM, Kang KW, et al. Report on the Korean standard differentiation of the symptoms and signs for the stoke-2. *Korean J of Orient Physiology & Pathology*. 2006;20:1789-1791.
15. Lee JA, Lee JS, Kang BK, Ko MM, Mun TU, Cho KH, et al. Report on the Korean standard pattern identifications for stroke-III. *Korean J Orient Int Med*. 2011;32(2):232-242.
16. Jiang M, Zha Q, Lu C, He Y, Lu A. Association between tongue appearance in traditional Chinese medicine and effective response in treatment of rheumatoid arthritis. *Complement Ther Med*. Jun 2011;19(3):115-121.
17. Su W, Xu ZY, Wang ZQ, Xu JT. Objectified study on tongue images of patients with lung cancer of different syndromes. *Chin J Integr Med*. Apr 2011;17(4):272-6.
18. Cao DS, Zeng MM, Yi LZ, Wang B, Xu QS, Hu QN, et al. A novel kernel Fisher discriminant analysis: Constructing informative kernel by decision tree ensemble for metabolomics data analysis. *Anal Chim Acta*. Nov 7 2011;706(1):97-104.
19. Zhang SQ. [Tongue temperature of healthy persons and patients with *yin* deficiency by using thermal video]. *Zhong Xi Yi Jie He Za Zhi*. Dec 1990;10(12):732-3, 709.