Integrative Medicine Research 14 (2025) 101115

Contents lists available at ScienceDirect

Integrative Medicine Research

journal homepage: www.elsevier.com/locate/imr

Original Article

Discovering the key symptoms for identifying patterns in functional dyspepsia patients: Doctor's decision and machine learning

Da-Eun Yoon [a], Heeyoung Moon [a,b], In-Seon Lee [a], Younbyoung Chae [a,*]

[a] Department of Science in Korean Medicine, College of Korean Medicine, Kyung Hee University, Seoul, Republic of Korea
[b] Department of Meridian and Acupoints, College of Korean Medicine, Semyung University, Jecheon, Republic of Korea

A R T I C L E I N F O

Keywords:
Functional dyspepsia
Pattern identification
Syndrome differentiation
Machine learning
Symptom importance

A B S T R A C T

Background: Pattern identification is a crucial diagnostic process in Traditional East Asian Medicine, classifying patients with similar symptom patterns. This study aims to identify key symptoms for distinguishing patterns in patients with functional dyspepsia (FD) using explicit (doctor's decision-based) and implicit (computational model-based) approaches.

Methods: Data from twenty-one FD patients were collected from local clinics of traditional Korean Medicine and provided to three doctors in a standardized format. Each doctor identified patterns among three types: spleen-stomach weakness, spleen deficiency with qi stagnation/liver-stomach disharmony, and food retention. Doctors evaluated the importance of the symptoms indicated by items in the Standard Tool for Pattern Identification of Functional Dyspepsia questionnaire. Explicit importance was determined through doctors' survey by general evaluation and by selecting specific information used for the diagnosis of patient cases. Implicit importance was assessed by feature importance from the random forest classification models, which classify three types for general differentiation and perform binary classification for specific types.

Results: Key symptoms for distinguishing FD patterns were identified using two approaches. Explicit importance highlighted dietary and nausea-related symptoms, while implicit importance identified complexion or chest tightness as generally crucial. Specific symptoms important for particular pattern types were also identified, and significant correlation between implicit and explicit importance scores was observed for types 1 and 3.

Conclusion: This study showed important clinical information for differentiating FD patients using real patient data. Our findings suggest that these approaches can contribute to developing tools for pattern identification with enhanced accuracy and reliability.

1. Introduction

Pattern identification, also known as syndrome differentiation, is a diagnostic procedure wherein a doctor collects and assesses a patient's symptoms or signs to assign a specific pattern. This process is crucial in Traditional East Asian Medicine (TEAM), as treatment decisions and subsequent clinical outcomes can be influenced by the identified patterns or syndromes.[1,2] For instance, functional dyspepsia (FD) has been classified into several phenotypes based on TEAM principles, with distinct treatment methods applied accordingly.[3,4] A meta-analysis has demonstrated improved outcomes when treatments are tailored to these phenotypes.[5]

Patterns are differentiated based on a set of symptoms and signs, with certain symptoms being particularly crucial for determining the pattern.[6] While some symptoms are useful for differentiating patterns in general, specific symptoms can be key for defining certain patterns. Traditionally, questionnaires for pattern identification have been developed, where weights are usually assigned to each symptom, reflecting its importance.[7,8] Although some symptoms commonly have high weights in several types, indicating general importance, different weights correspond to distinct pattern types, emphasizing that specific symptoms are key to particular patterns. These questionnaires have been developed using the Delphi method, which is based on the knowledge of experts. However, this approach faces challenges such as issues with inter-rater reliability,[9,10] and the lack of a reliable tool for guiding TCM pattern diagnosis has been reported in conditions like FD.[11]

With advancements in machine learning, recent studies have attempted to uncover the relationships between clinical signs and differentiated patterns and to identify critical information using diverse algorithms.[12-14] These approaches were discussed in terms of implicit and explicit knowledge in Park's study.[15] Explicit knowledge, such as the evaluated results of questionnaires, can be codified and expressed by a doctor's judgment. Implicit knowledge, on the other hand, refers to intuitive knowledge acquired through a doctor's experience that is chal-

* Corresponding author at: Acupuncture and Meridian Science Research Center, Kyung Hee University, 1 Hoegi-dong, Dongdaemun-gu, Seoul, Republic of Korea.
E-mail address: ybchae@khu.ac.kr (Y. Chae).

https://doi.org/10.1016/j.imr.2024.101115
Received 7 October 2024; Received in revised form 25 November 2024; Accepted 11 December 2024
Available online 12 December 2024

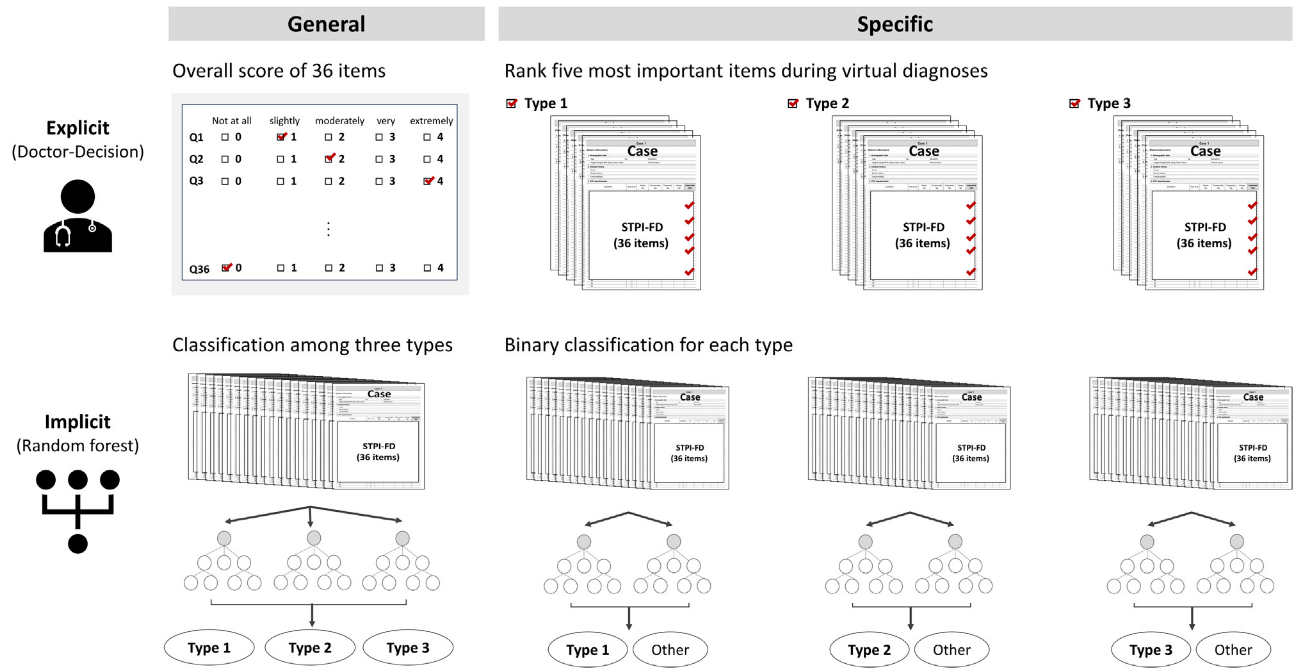

Fig. 1. Study procedure. Three Korean Medicine doctors completed the survey, which included scoring the overall importance of each questionnaire item on the Standard Tool for Pattern Identification of Functional Dyspepsia (STPI-FD), identifying the patterns of twenty-one FD patients, and ranking the five most important symptoms for identification. Overall importance scores indicate general explicit importance, and ranked symptoms were converted into scores to indicate specific explicit importance. For implicit knowledge, a random forest model was applied using variables from the STPI-FD questionnaire. General importance was derived from feature weights used to classify the three patterns, and specific importance was obtained from feature weights used in the binary classification for each pattern.

lenging to articulate but may be extracted as high-weighted features in pattern classifying models. The critical information might differ between implicit and explicit approaches. It has been shown that key symptoms can be identified by using these approaches in general, however, these alone are insufficient for determining a specific pattern type. To practically guide clinicians in deciding on a pattern type, it is essential to identify symptoms that are directly associated with it.

In this study, we evaluated the importance of symptoms both generally and specifically, using explicit (doctor's decision-based) and implicit (computational model-based) approaches. We identified generally important clinical signs or symptoms for defining patterns in patients with FD and identified the specifically important information for the particular pattern type. We further explored how these symptoms contribute to the identification in each pattern type. For the explicit approach, we surveyed doctors to obtain direct information underlying diagnostic decisions. For the implicit approach, we applied the random forest model to obtain calculated information regarding feature importance for classification.

2. Methods

2.1. Study design

This study aims to identify important medical information for making clinical decisions in TEAM regarding pattern identification of patients with FD. The FD patient data were collected from an observational study conducted in fifteen Korean medical clinics.[16] Three Korean Medicine doctors were asked to identify the patterns of FD patients and select the critical medical information for the identification while reviewing the patient data presented in a standardized format. The explicit importance of symptoms was calculated based on the doctor's survey, while the implicit importance was derived from the feature extraction in a random forest model. This concept of explicit and implicit importance was adopted from Park's study.[15] The importance of symptoms was evaluated from both general and specific perspectives. This study was approved by the Institutional Review Board (IRB) of Kyung Hee University (KHSIRB-22-074RA). Written consent was obtained from all participants. The schematic diagram of the experiment is described in Fig. 1.

2.2. Patient data acquisition

In the observational study, outpatients from fifteen Korean medical clinics who met the inclusion criteria were enrolled in the registry, and their data were collected during visits. Adults aged 18 to 70 years diagnosed with FD via Rome IV Criteria were included. Patients with diseases that induce gastrointestinal dysfunction or who had received active treatment within three months prior to visiting the clinic were excluded. Detailed inclusion and exclusion criteria are described in the protocol paper.[16] For this study, we selected twenty-one patient cases based on the type of FD.

Patients' clinical information, including demographic data (age, sex, education, height, weight, body mass index, marital status), medical history (onset, family history, comorbidities), and symptoms were extracted and provided to the doctors in a standardized survey form. Symptoms were obtained using the Standard Tool for Pattern Identification of Functional Dyspepsia (STPI-FD) questionnaire, which consists of 36 items reflecting the degree of various symptoms of FD and was developed based on the Delphi technique.[17]

2.3. Medical decision-making by doctors

Three doctors with at least four years of experience were invited to complete the survey. For each case, they identified the pattern among three options: spleen-stomach weakness pattern (type 1), spleen deficiency with qi stagnation/liver-stomach disharmony pattern (type 2), and food retention (type 3) pattern. After deciding the pattern, they ranked five important questionnaire items for identifying the pattern in the case from the STPI-FD questionnaire (5 = Most important, 4 = second, 3 = third, 2 = fourth, 1 = fifth) for the specific information.

Lastly, after reviewing all twenty-one patients, they scored the overall importance of each questionnaire item on the STPI-FD questionnaire for diagnosing FD patterns using a 5-point Likert scale (0 = not at all important, 1 = slightly important, 2 = moderately important, 3 = very important, 4 = extremely important) for the general information.

2.4. Evaluation of the important medical information

Importance scores of symptoms were assessed through scoring and ranking the questionnaire items for explicit information and using feature weights from the random forest model for implicit information.

Importance scores of general explicit information for each of the 36 items were obtained by averaging the overall importance scores provided by three doctors. Meanwhile, importance scores of specific explicit information were calculated using the rankings of the five most important symptoms identified by the doctors during pattern diagnoses for each patient. Based on the doctors' diagnoses for each patient, the average ranking scores for the 36 items by each doctor were first calculated and then averaged across the three doctors for each pattern. Here, we did not account for disagreements among the three doctors' pattern diagnoses and treated each doctor's judgment separately, as the important information chosen by the doctors could differ depending on the pattern type. The main difference between the general and specific importance scores is that the specific scores were obtained during the diagnosis of individual patient cases, while the general scores were based on the doctors' prior knowledge without reference to specific cases.

For implicit importance, we employed a random forest model to predict patterns using variables from the STPI-FD questionnaire. The random forest model is widely applied in TEAM pattern diagnosis due to its superior performance and reduced risk of overfitting.[18,19,20] The model has the advantage of providing useful internal estimates, such as feature selection, which aligns with the purpose of this study in evaluating the importance of information. General implicit importance scores were calculated based on the feature weights from the random forest model for classification among three pattern types. To evaluate specific implicit information, we modified the random forest model to perform binary classification and extracted feature weights for the three pattern types separately. Consistent with explicit importance, we labeled each doctor's pattern diagnoses for each case, regardless of their disagreements. The average feature weights from the 1000 random forest models were used as importance scores. Due to the limited dataset size, we applied leave-one-out cross-validation (LOOCV) for all random forest models.

The explicit and implicit importance scores were normalized to a scale of 0 to 5 for comparison and visualized using a heatmap for both general and specific types.

2.5. Severities of symptoms in three patterns

To investigate the details of the important information—specifically whether the significance arises from the severity or the absence of a symptom—we calculated the average score for each questionnaire item by patterns and normalized the scores to a scale of 0 to 5. A high score indicates that the symptom described in the questionnaire is severe. Symptom severity for each questionnaire item was presented by patterns using a heatmap, along with normalized importance scores.

2.6. Correlation analysis between explicit and implicit importance

Pearson's correlation analyses were conducted to examine the correlations between explicit and implicit importance within each pattern type. Normalized importance scores were analyzed with a significance level set at $p < 0.05$. A radar chart was utilized to compare the importance scores of explicit and implicit information across the pattern types.

3. Results

3.1. Key symptoms for distinguishing patterns identified through the explicit approach

The explicit importance was calculated based on doctors' judgment. Symptoms of Q18 "I do not feel like eating", Q13 "I sometimes feel nauseous or vomit", and Q14 "I sometimes feel nauseous or vomit, and the symptoms reduce after vomiting" were recognized as the most important symptoms in general. For specific symptoms by pattern, Q17 "I eat small amounts and feel full easily" was the most important for type 1, while Q13 "I sometimes feel nauseous or vomit" was the most important for both type 2 and type 3 during the diagnosis of patient cases (Fig. 2A).

3.2. Key symptoms for distinguishing patterns identified through the implicit approach

The implicit importance was evaluated by extracting feature importance from the random forest model. A high score of implicit importance indicates that the presence or severity of the symptom is determinative for diagnosing a specific pattern. Symptoms of Q25 "My complexion is pale and sometimes turns yellow", Q8 "My chest feels tight", and Q17 "I eat small amounts and feel full easily" were identified as the most important symptoms in general. Specifically, Q8 "My chest feels tight" was significant for identifying type 1, Q25 "My complexion is pale and sometimes turns yellow" for type 2, and Q14 "I sometimes feel nauseous or vomit, and the symptoms reduce after vomiting" for type 3. Symptoms that were highly important in specific patterns also tended to be highly important in general (Fig. 2B).

3.3. Distribution of symptom severities in three pattern types

Based on the patterns identified by the doctors, the average scores for each of the 36 questionnaire items from patient data were calculated for each pattern and described as "symptom severity". For example, the score for Q8 "My chest feels tight" was low in type 1 (2.48) but comparatively higher in type 2 (3.45) and type 3 (4.16). The score for Q14 "I sometimes feel nauseous or vomit, and the symptoms reduce after vomiting" was low in type 1 (0.38) and type 2 (0) but comparatively higher in type 3 (2.90). The score for Q25 "My complexion is pale and sometimes turns yellow" was high in type 1 (2.63) and type 3 (3.11) but comparatively lower in type 2 (0.92) (Fig. 2C).

3.4. Correlations between explicit and implicit importance scores

There was generally an overlap between the implicit and explicit approaches (Fig. 3). The importance scores were correlated between the implicit and explicit approaches in type 1 ($r = 0.34$, $p = 0.04$) and type 3 ($r = 0.41$, $p = 0.01$), but not in type 2 ($r = 0.04$, $p = 0.84$).

4. Discussion

In this study, critical symptoms for distinguishing patterns of FD were identified using both explicit and implicit approaches. We extracted important symptoms that contribute to classifying patterns in general, as well as symptoms that are significant to specific patterns. Explicit importance emphasized symptoms associated with diet and nausea, whereas implicit importance highlighted symptoms related to complexion and chest tightness as generally critical. Distinct symptoms crucial to specific pattern types were identified, and a significant correlation was observed between implicit and explicit importance scores for these symptoms.

In this study, we demonstrated the important symptoms for identifying patterns of FD both generally and specifically. Based on doctors' judgment, key symptoms may differ when they considered the pattern identification process in general versus when they prioritize symptoms

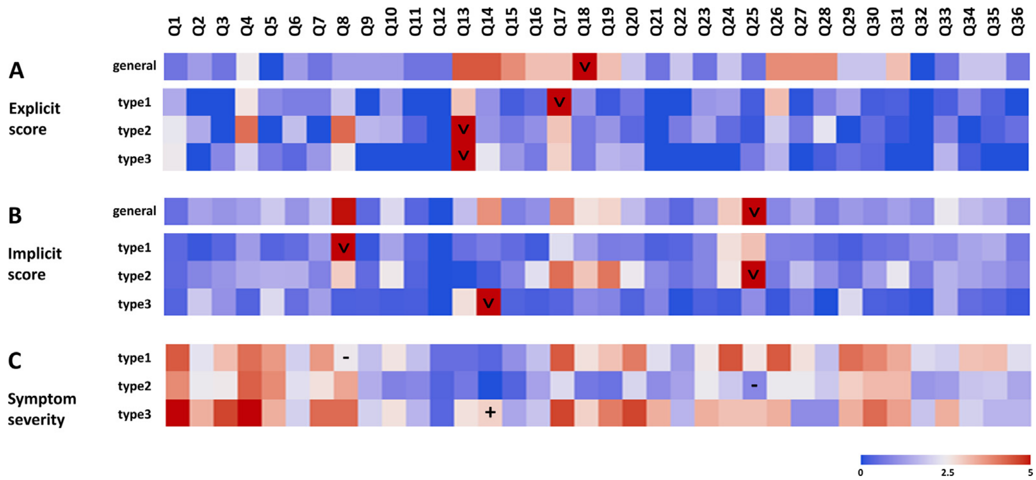

Fig. 2. Explicit and implicit importance scores for pattern identification and the patient-reported scores of symptoms. Importance scores for distinguishing three types of patterns in patients with FD are visualized using a heatmap. Red indicates high and blue indicates low scores. In each questionnaire item, A) explicit importance scores for general differentiations and specific to three pattern types, B) implicit importance scores for general differentiations and specific to three pattern types and, C) the mean score of symptoms from questionnaire for each pattern were presented. "v" represents the highest score among the 36 questionnaire items. "–" represents a symptom showing comparatively lower symptom severity with the highest implicit importance score, while "+" represents a symptom showing comparatively higher symptom severity with the highest implicit importance score.

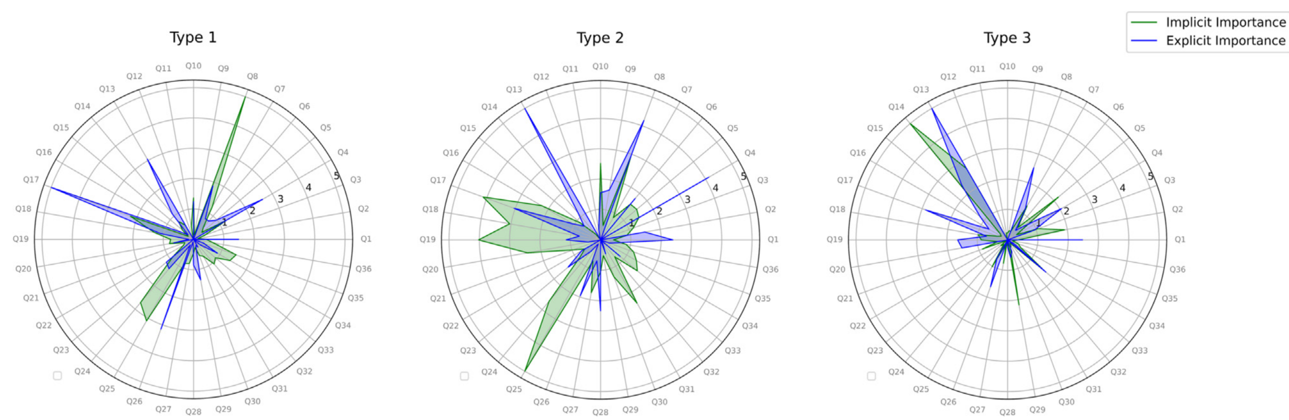

Fig. 3. Radar plots of implicit and explicit importance scores. Distribution of importance scores in explicit and implicit approaches is presented in a radar plot for the three patterns. The green line represents implicit importance score and the blue line represents explicit importance score.

in individual cases for the identification. For instance, Q18 was the most important symptom for identifying patterns of FD generally, but not for any specific patterns. Furthermore, Q17 was not notable for identifying patterns of FD generally, but stood out as the most important symptom for identifying patterns of FD type 1. In contrast, Q13 was the second important symptom for identifying patterns of FD generally, and was also the most important symptom for identifying patterns of FD type 2 and type 3 specifically (Fig. 2A). Based on machine learning, key symptoms were similar when identifying patterns of FD both generally and specifically. For instance, Q25 and Q8 were the most important symptoms for identifying patterns of FD generally, as well as for identifying certain patterns of FD type 1 and FD type 2, respectively. Furthermore, Q14 was the most important symptom for identifying patterns of FD type 3 and also an important symptom for identifying patterns of FD generally (Fig. 2B). Identifying important symptoms of particular patterns of FD can be useful for clinicians in diagnosing each pattern of FD.

Generally speaking, the significance of information can be decided when symptom manifestation is salient compared to others. It does not always mean that the symptom is more severe in the specific pattern. Instead, the significance of information can be found even when the symptom is less severe in the pattern. High implicit importance scores can be derived from both low symptom levels and high symptom levels. For example, Q8 had high implicit importance scores for type 1, especially with low symptom scores in type 1 compared to other patterns. Similarly, Q25 had a high implicit importance score, with low symptom scores especially in type 2. Conversely, Q14 indicated a high implicit importance score with a high symptom score in type 3 (Fig. 2C). These findings suggest that the absence of symptoms in Q8 and Q25 might help distinguish types 1 and 2, while severe symptom of Q14 might help distinguish type 3. Conventional instruments for pattern identification usually follow scoring methods where each item has a given point for its presence in a particular pattern, and the summation of points determines the pattern. This method adds points only with the presence of the symptom but not for the absence. However, the absence of a symptom can be important to rule out certain patterns and may have already been processed intuitively by doctors. This was also revealed as mutual exclusion of clinical features in a data-driven method using a latent class analysis model.[21]

In the current study, we found that the important symptoms generally overlapped between explicit and implicit approaches (Fig. 3). There were positive correlations of importance between the implicit and explicit scores except for type 2. Some information consciously recognized as important by the doctor may differ from what becomes an important feature when predicted by machine learning. This difference might be derived from the methodological attribute. Implicit scores obtained from machine learning consider both significantly low and high symptom scores to distinguish the patterns, while for explicit scores, doctors are likely to rate the importance based on the severity of the symptom. In type 1, the severities of important items were rather high (Q24) or low (Q8). On the other hand, the severities of important items were mostly low for type 2 (Q17, Q19, Q25), while high for type 3 (Q13, Q14). These variations might affect the results of correlation between explicit and implicit importance scores in three pattern types. Especially, no significant correlation in importance scores between explicit and implicit approaches was observed for type 2 since the symptom scores of important items were mostly low.

Pattern identification is a doctor's conceptual process of diagnosing and prescribing, traditionally characterized by its veiled nature. Several efforts have been made to standardize and visualize this process. Instruments have been developed by experts to guide pattern identification for various diseases or symptoms, and recently, machine-learning-based methods have been utilized to unveil this process.[22,23] These developments can suggest the importance level of clinical knowledge used in the pattern identification process, aiding doctors in making medical decisions. The concept of implicit and explicit knowledge from Park's study is relevant in representing these two approaches.[15] The study showed an incongruent relationship between implicit and explicit knowledge in differentiating patterns in patients with allergic rhinitis, but a congruent relationship for prescription purposes. In this study, we further investigate specific knowledge for specific patterns within these two approaches. To capture more diagnosis-like information for explicit knowledge, we asked doctors to provide explicit information during the diagnosis for each case. This allowed us to suggest the critical information needed to distinguish patterns specifically, with concordance between the implicit and explicit approaches.

In actual clinical practice, clinicians select and pose a few key questions to patients to assess them more effectively, rather than asking every question on a questionnaire. A few questions may suffice to distinguish between patterns. This procedure could be converted into a hierarchical model, similar to a decision tree model that determines the most effective strategy to reach a diagnosis. In this model, by choosing "yes" or "no" for several significant features, we can distinguish the pattern of a patient. Developing diagnostic tools based on the key features identified by machine-learning models and validated by expertise could be a practical option to support clinicians more reliably.

Given the significance of pattern identification for deciding a treatment plan or outcomes,[24,25] the low inter-rater reliability of this process has been continuously highlighted.[26] Diverse efforts have been made to standardize this process by developing questionnaires or tools with data-driven approaches.[14] However, there is also a question of the necessity for high inter-rater reliability as it may not assure better clinical outcomes.[9] In this study, when each doctor explicitly answered the important symptoms, the important information was not consistent—that is, the information which doctors paid more attention to differed among doctors. Treatment regimens may also differ as clinicians adapt their clinical strategies based on their perspectives, which can be a characteristic of personalized medicine.

This study has limitations. First, the number of cases was not sufficient to accurately extract important information for each pattern. There were fewer cases diagnosed with type 3 pattern than types 1 or 2, which may have affected the overall importance. Second, we did not evaluate how the diagnosed pattern type is related to the prescription or treatment plans. TEAM's treatment principles involve concepts of "root treatment" and "branch treatment"; where treatment can be adjusted based on both the chief symptoms and the patterns. Our study only focused on the patterns. The treatment strategy might vary depending on the patient's unique symptoms. Third, this study focused on identifying important information related to diverse symptoms closely linked to the disease. However, other factors beyond disease-related symptoms, such as obesity, medical history, and holistic features like pulse or tongue, also contribute to identifying the patterns.[27] Additionally, incorporating diverse aspects of clinical information into modeling has led to novel subtyping in FD patients.[28] Future studies should further explore how medical decisions, including pattern diagnosis and prescriptions, are made based on both disease-related symptom and non-disease-related symptom factors.

In conclusion, we identified critical symptoms for distinguishing patterns of FD using both explicit and implicit approaches. Additionally, we identified important symptoms for pattern identification in general, as well as symptoms specific to individual patterns. The explicit approach, based on doctors' clinical judgments, and the implicit approach, using the random forest model, both provided insights into symptom significance. Our findings suggest that these approaches can contribute to developing tools for pattern identification with enhanced accuracy and reliability. Future research should explore the incorporation of additional patient features and treatments to further refine diagnostic tools and support personalized medicine in TEAM.

Author contribution

Da-Eun Yoon (Conceptualization, Formal analysis, Visualization, Writing - original draft), Heeyoung Moon (Data curation, Resources), In-Seon Lee (Writing - review and editing), Younbyoung Chae (Conceptualization, Supervision, Writing - review and editing). All authors read and approved the final manuscript.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Funding

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (Grant number: HF22C0023) and by the Institute of Information and Communications Technology Planning and Evaluation (IITP) grant funded by the Korea government (MSIT) [No. RS-2022-00155911, Artificial Intelligence Convergence Innovation Human Resources Development (Kyung Hee University)].

Ethical statement

All data were collected from experiments conducted in accordance with guidelines issued by the human subjects committee and approved by the institutional review board of Kyung Hee University, Seoul, Republic of Korea (approval number: KHSIRB-22-074RA).

Data availability

The authors can provide upon reasonable request.

Acknowledgments

We sincerely appreciate the contributions of the Korean medical doctors who participated in this study: Seunguk Kweon (Kyung Hee GJ Korean Medical Clinic), Tae Hyeong Kim (Kyung Hee GJ Korean Medical Clinic), and Min-Ji Park.

References

1. Lee YS, Ryu Y, Chae Y. Acupoint selection based on pattern identification results or disease state. *Integr Med Res*. 2020;9(2):100405.
2. Ferreira AS, Lopes AJ. Chinese medicine pattern differentiation and its implications for clinical practice. *Chin J Integr Med*. 2011;17(11):818–823.
3. Zhang J, Chen T, Wen Y, Siah KTH, Tang X. Insights and future prospects of traditional Chinese medicine in the treatment of functional dyspepsia. *Phytomedicine*. 2024;127:155481.

4. Zhou W, Su J, Zhang H. Efficacy and safety of acupuncture for the treatment of functional dyspepsia: meta-analysis. *J Altern Complement Med*. 2016;22(5):380–389.
5. Ha NY, Jeong H, Lee H, Ko SJ, Park JW, Kim J. Safety and effectiveness of traditional herbal medicine Siho-sogan-san in functional dyspepsia: a systematic review and meta-analysis. *J Ethnopharmacol*. 2023;313:116518.
6. Kim SY, Hong SH, Park JW, et al. Analysis of diagnostic decision in acupuncture from the actual functional dyspepsia patient's clinical information. *Integr Med Res*. 2020;9(4):100419.
7. Kang BK, Park TY, Lee JA, et al. Reliability and validity of the Korean standard pattern identification for stroke (K-SPI-Stroke) questionnaire. *BMC Complement Altern Med*. 2012;12:55.
8. Chen XL, Wen Y, Wu ZC, et al. Development of a traditional Chinese medicine syndrome-specific scale for ulcerative colitis: the large intestine dampness-heat syndrome questionnaire. *Evid Based Complement Alternat Med*. 2018;2018:4039019.
9. Jacobson E, Conboy L, Tsering D, et al. Experimental studies of inter-rater agreement in traditional Chinese medicine: a systematic review. *J Altern Complement Med*. 2019;25(11):1085–1096.
10. Popplewell M, Reizes J, Zaslawski C. Consensus in traditional Chinese medical diagnosis in open populations. *J Altern Complement Med*. 2019;25(11):1109–1114.
11. Ho LT, Chung VC, Wong CH, et al. Evaluating traditional Chinese medicine diagnostic instruments for functional dyspepsia: systematic review on measurement properties. *Integr Med Res*. 2021;10(3):100713.
12. Won J, Lee BH, Jung WM, Chae Y, Lee H. Herbal medicine for inflammatory bowel diseases: development of pattern identification algorithms by retrospective analysis of case series data. *Eur J Integr Med*. 2020;36:101114.
13. Huang WT, Hung HH, Kao YW, et al. Application of neural network and cluster analyses to differentiate TCM patterns in patients with breast cancer. *Front Pharmacol*. 2020;11:670.
14. Ho L, Xu Y, Zhang NL, et al. Quantification of prevalence, clinical characteristics, co-existence, and geographic variations of traditional Chinese medicine diagnostic patterns via latent tree analysis-based differentiation rules among functional dyspepsia patients. *Chin Med*. 2022;17(1):101.
15. Park M, Kim MH, Park SY, Choi I, Kim CE. Individualized diagnosis and prescription in traditional medicine: decision-making process analysis and machine learning-based analysis tool development. *Am J Chin Med*. 2022;50(7):1827–1844.
16. Moon H, Yoon DE, Seo Y, Lee IS, Chae Y. A prospective observational study of optimal acupoint selection on patients with functional gastrointestinal disorders. *Medicine (Baltimore)*. 2023;102(28):e34316.
17. Ha NY, Ko SJ, Park JW, Kim J. Development of a standard tool of pattern identification for functional dyspepsia: a cross-sectional study from Korea. *Healthcare (Basel)*. 2024;12(23):2331.
18. Fawagreh K, Gaber MM, Elyan E. Random forests: from early developments to recent advancements. *Syst Sci Control Eng*. 2014;2(1):602–609.
19. Zhang H, Ni W, Li J, Zhang J. Artificial intelligence-based traditional Chinese medicine assistive diagnostic system: validation study. *JMIR Med Inform*. 2020;8(6):e17608.
20. Wang H, Liu X, Lv B, Yang F, Hong Y. Reliable multi-label learning via conformal predictor and random forest for syndrome differentiation of chronic fatigue in traditional Chinese medicine. *PLoS One*. 2014;9(6):e99565.
21. Zhang NL, Fu C, Liu TF, et al. A data-driven method for syndrome type identification and classification in traditional Chinese medicine. *J Integr Med*. 2017;15(2):110–123.
22. Jung WM, Park IS, Lee YS, et al. Characterization of hidden rules linking symptoms and selection of acupoint using an artificial neural network model. *Front Med*. 2019;13(1):112–120.
23. Bae H, Lee S, Lee CY, Kim CE. A novel framework for understanding the pattern identification of traditional asian medicine from the machine learning perspective. *Front Med (Lausanne)*. 2021;8:763533.
24. Kim SH, Jeong JH, Lim JH, Kim BK. Acupuncture using pattern-identification for the treatment of insomnia disorder: a systematic review and meta-analysis of randomized controlled trials. *Integr Med Res*. 2019;8(3):216–226.
25. Kim CH, Yoon DE, Lee YS, Jung WM, Kim JH, Chae Y. Revealing associations between diagnosis patterns and acupoint prescriptions using medical data extracted from case reports. *J Clin Med*. 2019;8(10).
26. Schnyer RN, McKnight P, Conboy LA, et al. Can reliability of the Chinese medicine diagnostic process be improved? Results of a prospective randomized controlled trial. *J Altern Complement Med*. 2019;25(11):1103–1108.
27. Birch S, Alraek T, Lee MS, Kim TH. Thinking about traditional medicine diagnostic patterns and instruments. *Integr Med Res*. 2021;10(4):100745.
28. Park SY, Bae H, Jeong HY, Lee JY, Kwon YK, Kim CE. Identifying novel subtypes of functional gastrointestinal disorder by analyzing nonlinear structure in integrative biopsychosocial questionnaire data. *J Clin Med*. 2024;13(10):2821.