European Journal of Integrative Medicine 15 (2017) 47–63

Contents lists available at ScienceDirect

European Journal of Integrative Medicine

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

Review article

The reliability and validity of instruments measuring pattern identification in Korean medicine: A systematic review

Mi Mi Ko[a], Myeong Soo Lee[b], Stephen Birch[c], Ju Ah Lee[d,*]

[a] KM Fundamental Research Division, Korea Institute of Oriental Medicine, Daejeon, Republic of Korea
[b] Clinical Research Division, Korea Institute of Oriental Medicine, Daejeon, Republic of Korea
[c] Kristiania University College, Institute of Health Sciences, Oslo 0855, Norway
[d] Department of Korean Internal Medicine, College of Korean Medicine, Gachon University, Republic of Korea

ARTICLE INFO

Keywords:
Pattern identification
Korean medicine
Questionnaires
Instruments
Reliability
Validity

ABSTRACT

Introduction: Pattern identification (PI) is a unique concept in traditional East Asian medicine that refers to diagnostic process. This study aims to summarize and critically evaluate the reliability and validity of all Korean designed PI instruments.

Methods: A systematic literature search was conducted in MEDLINE, EMBASE, and eight Korean medical databases from the study's inception to August 2017 to identify all studies that described Korean PI instruments. We included PI instruments without restriction on the types of diseases or, conditions of participants. General characteristics of the included instruments, reliability verification, internal consistency, and the types of validity, including construct validity, content validity, and criterion validity, were reviewed.

Results: Thirty-six PI instruments were identified and analyzed. Ten of them were PI diagnostic tools for specific diseases. Fifteen were related diagnoses of PI for non-specific diseases. Four instruments involved the diagnosis of PI for specific decoction treatment and seven were related to the diagnosis of constitution, including Sasang constitution. The most commonly used statistical test was Cronbach's alpha, an estimator of internal consistency. A total of ten studies conducted test-retest reliability with varying time intervals, and the test-retest coefficient was moderate to good.

Conclusions: This article examined the reliability and validity of the PI instruments used in Korea and the need to improve the standardisation of PI diagnosis. Almost studies reported the value of internal consistency only. Clinical studies on the application of future questionnaires will be needed in the future.

1. Introduction

Pattern identification (PI) is a unique concept in traditional East Asian medicine that refers to diagnostic process. Although there are many PI instruments, a systematic review of the literature on the reliability of these diagnostic tool has not previously been carried out. Because evidence-based medicine (EBM) is becoming increasingly common in medicine, including complementary and alternative medicine (CAM), the number of clinical trials has steadily increased for at least a half-century [1]. The number and quality of measurement tools needed to conduct clinical trials continues to increase at a rapid pace.

Clinical trials invariably demand measurement of the effect of a medication or procedure in a disease process. For example, the measurement of the effect of treatment on a neurologic disease requires appropriate tools. Understanding the qualities of a scale is important not only for those planning clinical trials, but also for clinicians who wish to interpret trials reported in the literature and who want to compare reported treatments using different scales [2].

A useful measure or scale should have the following characteristics: (1) it should be appropriate to the task; (2) it should be valid, that is, it should measure what it purports to measure; (3) it must be accurate and should accurately measure what it purports to measure; (4) it must be reliably reproducible (precise); (5) it should be efficient and easy to use, with little special training; (6) it should be sensitive to change in the underlying condition yet relatively insensitive to symptom fluctuation; and (7) it should be consistent over time, that is, not subject to so-called frame-of-reference shifts [2].

As a type of CAM, traditional East Asian medicine (TEAM) has a specific diagnosis system called Pattern identification (PI). PI is the diagnosis system of medicine characterized by its own theoretical basis and practical experience in TEAM. It is a unique diagnostic system that uses a comprehensive analysis of symptoms and signs to assess the

* Corresponding author.
E-mail address: motoong@gmail.com (J.A. Lee).

http://dx.doi.org/10.1016/j.eujim.2017.09.005
Received 19 May 2017; Received in revised form 18 September 2017; Accepted 18 September 2017

cause, nature, and location of the illness, the patient's physical condition, and the patient's treatment [3,4]. Although PI is widely accepted in individual diagnosis and treatment worldwide, there remains insufficient clinical evidence regarding the therapeutic effects of PI as well as the absence of objective diagnostic tools.

Recently, several studies about clinical evidence regarding the therapeutic effects of PI and the absence of an objective diagnostic tool were conducted. A few diagnostic tools that are mainly questionnaires have been developed.

However, these diagnostic tools have several limitations. In particular, the reliability and validity of questionnaires were not established. In this review, we aimed to identify, describe and evaluate all available questionnaires designed to measure PI in Korean Medicine (KM) and to identify major trends in the use/development of the questionnaires of PI in research and practice. Finally, we suggest areas for future research.

2. Methods

2.1. Data sources and search terms

A systematic review was conducted to collect studies about PI instruments in KM. The following electronic Korean medical databases were searched without restriction of language from their respective inceptions up to September 2016: the Korean Studies Information Service System (KISS), DBPIA, Korea Institute of Science and Technology Information, Research Information Service System (RISS), KoreaMed, the Korean Medical Database (KMbase), the Oriental Medicine Advanced Searching Integrated System (OASIS) and the National Assembly Library. Also, we updated the search results August 2017. We also searched EMBASE, and MEDLINE (last searched August, 2017). The search terms used were "pattern identification", "Syndrome differentiation", "pattern diagnosis", "patternization", "diagnosis", "questionnaire", "instrument", "reliability", "validity" and Korean language terms or English terms related to pattern identification and clinical trials in traditional KM (S2 File). The different search terms were used on these Korean databases due to the difficulty of implementing identical search strategy such as that found in international databases. The references in all located articles were also searched. Hard copies of all articles were obtained and read in full.

2.2. Study selection and eligibility

All included studies contained Korean PI instruments. We included studies without restrictions to the types of diseases or conditions of participants. We excluded PI instruments which were not included information about the reliability and validity. We excluded studies that provided only PI results without PI instruments.

2.3. Data extraction

Two reviewers (MMK, JAL) independently read the titles and abstracts of the references found in the search and excluded irrelevant studies. All studies retrieved from the screening were read in full and the eligibility was independently evaluated by two researchers. Disagreements were resolved via a discussion by authors and an arbiter (MSL). A qualitative summary of results in the included studies were performed. We extracted the study ID (first author, year of publication) and characteristics of the participants.

2.4. Description of the instruments

Descriptive data extracted from the publications included the target/test population, type of PI, number of items, item generation, number of scales, and domains to which the items could be classified. Results related to reliability and validity were also included.

2.5. Reliability

Reliability is the degree to which an assessment tool produces stable and consistent results. Internal consistency and the test-retest reliability are defined as two of the main types of reliability evidence and include parallel forms reliability and, inter-rater reliability. Internal consistency reliability is a measure of reliability used to evaluate the degree to which different test items that probe the same construct produce similar results and is usually measured with Cronbach's alpha [5,6]. Test-retest reliability is a measure of reliability obtained by administering the same test twice over a period of time to a group of individuals and measured with Pearson's (or Spearman's) correlation coefficient [6].

2.6. Validity

A variety of types of validity exist, each designed to ensure that specific aspects of measurement tools are accurately measuring what they are intended to measure and that the results can be applied to real-world settings. The validity of the instruments was assessed by means of the contents validity, construct validity and criterion validity [5,7]. Content validity refers to how accurately a measurement tool taps into the various aspects of the specific construct in question, and this validity is most often measured by relying on the knowledge of people who are familiar with the construct being measured [8]. Criterion-related validity usually includes any validity strategies that focus on the correlation of the test being validated with some well-respected outside measure(s) of the same objectives or specifications. Construct validity has traditionally been defined as the experimental demonstration that a test is measuring the construct it claims to be measuring. To take a unified definition of construct validity, we could demonstrate it using content analysis, correlation coefficients, factor analysis, ANOVA studies demonstrating differences between differential groups or pretest-posttest intervention studies, multi-trait/multi-method studies, etc. [5,7,9]. For diagnostic accuracy tests, ROC curve analysis is used to interpret validity indexes. ROC curve analysis is used in clinical epidemiology to quantify how accurately medical diagnostic tests can discriminate between two patient states [10].

3. Results

3.1. Description of the chosen PI instruments by the studies

The literature search revealed 535 studies, of which 483 studies were excluded after screening the abstracts and titles. A total of 52 studies were read in full and evaluated, of which 16 articles were excluded because they did not conduct reliability or validity research (S1 File). The study process is summarized in a flow diagram in Fig. 1. Finally, 36 studies met our inclusion criteria. In total, the 36 included studies were all conducted in Korea and published between 1993 and 2016.

All PI instruments were questionnaires (Table 1). Ten of them were PI diagnostic tools for a specific disease. There were several diseases involved within the included studies: Oriental Obstetrics & Gynecology [11], obesity [12], stroke [13,14], depression [15], alcoholic hepatitis [16], neck pain [17], knee pain [18], Tic disorder in children [19], and anxiety disorder [49]. Fifteen of them were related to the diagnosis of PI for non-specific disease. There were various PI diagnostic tools for non-specific diseases involved within the included studies: Cold-Heat [20–22,48], Joseup (dry-dampness) [23], Yin-Yang [24], Yol (Heat) [25], Blood stasis [26,27], Damum (phlegm-retained fluid) [28], Yin [29,30], Cold-Heat & Deficiency-Excess [22], Deficiency-Excess [31], and Phlegm [32] patterns.

Four instruments considered the diagnosis of PI for decoction treatment: Pyungweesan (Pingwei San) [33], Yukmijiwhang-tang (Liu Wei Di Huang Tang) [36], Guibi-tang (Gui Pi Tang) [35], and Buzhongyiqi-tang [34]. Seven were related to the diagnosis of constitution: