

REVIEW ARTICLE

## 기능성 소화불량증 치료제의 임상시험 평가 가이드라인 권고안

이한희<sup>1,2</sup>, 정혜경<sup>3</sup>, 최명규<sup>1,2</sup>

가톨릭대학교 의과대학 소화기내과학교실<sup>1</sup>, 가톨릭 광의학연구소<sup>2</sup>, 이화여자대학교 의과대학 소화기내과학교실<sup>3</sup>

### Guideline Recommendation for Endpoints Used in Clinical Trials for Functional Dyspepsia

Han Hee Lee<sup>1,2</sup>, Hye-Kyung Jung<sup>3</sup> and Myung-Gyu Choi<sup>1,2</sup>

Division of Gastroenterology, Department of Internal Medicine, College of Medicine, The Catholic University of Korea<sup>1</sup>, Catholic Photomedicine Research Institute<sup>2</sup>, Division of Gastroenterology, Department of Internal Medicine, Ewha Womans University School of Medicine<sup>3</sup>, Seoul, Korea

Functional dyspepsia is a disease, in which there is no organic lesion but chronic and repetitive postprandial fullness, early satiation, epigastric pain, and epigastric burning. Functional dyspepsia is not life-threatening but its symptoms are relapsing and remitting and persist over a lifetime, limiting the social life and reducing the quality of life. Therefore, the treatment for acute relapsing period may help improve the short-term symptoms. Continuous medication may be needed to improve the long-term symptoms. Research designs to demonstrate the short-term efficacy of therapeutic agents may differ from clinical trials to demonstrate long-term efficacy. There are many difficulties in clinical trial design, implementation, and screening because there are no international standards of clinical trials for functional dyspepsia. The purpose of this guideline recommendation is to develop a standard for clinical trials, such as clinical trial subjects and evaluation methods, in the development of therapeutic agents for functional dyspepsia. The ultimate aim is to enhance the safety and efficacy of therapeutic agents for functional dyspepsia and promote the development of new therapeutic agents. (Korean J Gastroenterol 2018;72:170-178)

**Key Words:** Dyspepsia; Gastrointestinal diseases; Clinical trial; Endpoint determination; Patient reported outcome measures

## 서론

기능성 소화불량증은 기질적 병변은 없으나 상부위장관에 만성적이며 반복적인 위장 통증 혹은 불편함, 상복부 충만감, 팽만감, 조기 포만감, 더부룩함, 메스꺼움, 트림 등이 발생하는 질환이다.<sup>1</sup> 기능성 소화불량증은 생명을 위협하지는 않으나 일생동안 그 증상이 지속되면서 쉽게 완치되지 않아 사회 생활에 제약을 받고 삶의 질 저하를 가져올 수 있다.

기능성 소화불량증 임상시험에 대하여는 국제적 기준이 마련되지 않아 임상시험 계획, 수행, 심사 과정에서 어려운 점이 많았다. 본 연구는 국제 기준과 조화를 이룬 기능성 소화불량

증 임상시험 가이드라인(안)을 마련하기 위하여 다음의 과정을 거쳤다. 1) 새로이 개정된 로마기준 IV 문헌과 국내외 허가 관청에서 발간된 기능성 위장관 질환에 대한 지침을 검토하였다. 2) 현재까지 발간된 국내외 소화불량증의 임상시험 문헌을 검색하여 체계적 분석을 시행하였고 ClinicalTrials.gov에 등록된 계획서를 검색하였다. 임상시험에 사용된 유효성 평가 변수와 측정 방법에 대한 장단점을 조사하고 이를 평가, 분석하여 가이드라인 초안을 제정하였다. 3) 전문 학회, 제약회사, 식약청 기타 모든 이해관계자 및 별도로 초청된 외부 검토자로 전문가 그룹을 결성하여 지침 작성과정에서 자문을 얻었다. 4) 전문가 초청 미니 심포지엄에서 가이드라인 초안에 대

Received September 3, 2018. Revised October 12, 2018. Accepted October 16, 2018.

© This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2018. Korean Society of Gastroenterology.

교신저자: 최명규, 06591, 서울시 서초구 반포대로 222, 가톨릭대학교 의과대학 소화기내과학교실

Correspondence to: Myung-Gyu Choi, Division of Gastroenterology, Department of Internal Medicine, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul 06591, Korea. Tel: +82-2-2258-2083, Fax: +82-2-2258-2089, E-mail: choim@catholic.ac.kr,

ORCID: <https://orcid.org/0000-0003-4083-5187>

Financial support: None. Conflict of interest: None.

한 검토와 토론을 거쳐 제2안이 만들어졌다. 제2안은 18명의 외부 전문가들에게 보내져 10일간의 검토를 거친 후 10개의 항목에 대하여 델파이기법을 적용하여 온라인 투표를 실시하였으며 모든 항목에서 높은 동의수준(88.2-100%)을 얻었다. 5) 소화불량증 임상시험 가이드라인 최종안은 대한소화기 기능성질환·운동학회, 대한상부위장관·헬리코박터학회에 보내 검토를 요청하였으며 추인을 얻었다. 본 가이드라인 권고안은 기능성 소화불량증 치료제를 개발함에 있어 임상시험 대상자, 평가 방법 등 임상시험의 수행기준을 마련하여 지침서로 활용함으로써 기능성 소화불량증 치료제의 안전성과 유효성을 높이고 나아가 새로운 치료제의 개발을 촉진하는데 그 목적이 있다.

## 본 론

### 1. 기능성 소화불량증의 진단

기능성 소화불량증의 진단은 직접 진단에 이용할 수 있는 생물학, 생리학, 해부학적 특징이 없고 생물학적 표지자(biological marker)가 아직 밝혀지지 않아 환자의 증상을 기준으로 이루어진다. 로마기준 III에서 소화불량증으로 정의된 증상은 불편한 식후 충만감, 조기 포만감, 명치 부위 복통 혹은 쓰림(burning)의 네 가지 증상이다(Table 1). 이 네 가지 증상 중 하나 이상인 경우로 소화불량증은 정의되며, 명치동통증후군(epigastric pain syndrome, EPS)과 식후불편증후군(post-prandial distress syndrome, PDS)의 두 가지 아형으로 분류된다.<sup>1</sup> PDS는 식후 충만감(위 내에 음식이 계속 남아 있는 것 같은 불편한 증상)과 조기 포만감(식사를 시작하자마자 곧 배가 부르고 더 이상의 식사를 할 수 없는 느낌) 증상이 주 2회 이상 발생하는 경우이며, 상복부 팽만감, 식후 메스꺼움, 트림 등을 포함할 수 있다. EPS는 중등도의 명치 부위 통증이나 쓰림(명치 부근에서 발생하는 주관적이고 불편한 화끈거리는 느낌) 증상이 주 1회 이상 발생하며, 다른 부위의 통증이 아니어야 하고, 배변에 의하여 완화되지 않아야 하며, 담낭이나 오디 괄약근 질환이 없어야 한다. 이전 로마기준과

비교하면 PDS는 운동장애형과, EPS는 궤양형과 유사하다. 원인을 알 수 없는 메스꺼움, 구토, 가슴쓰림은 다른 범주로 분리되었다.

네 가지 주요 증상 이외의 증상들은 기능성 소화불량증 진단기준과는 무관하며 동반되어도 기능성 소화불량증으로 진단한다. 예를 들면, 복부팽만(bloating)은 과민성 장증후군이거나 기능성 복부팽만의 범주에 들어가나 소화불량증의 주요 증상을 함께 가지고 있으면 기능성 소화불량증으로 진단할 수 있다. 메스꺼움의 경우 중추신경계에서 시작된 경우가 더 많으며, 구토와 동반이 되어도 위에 국한되었다고 간주하지 않는다. 헛트림과 가슴쓰림은 위 질환이라기보다는 식도 질환으로 간주한다. 기능성 소화불량증의 정의는 개정된 로마기준 IV에서도 큰 차이가 없다. 로마기준 IV에서는 가슴쓰림이 동반되어도 전형적인 증상 기준에 합당한 경우 기능성 소화불량증으로 진단할 수 있다. 최근 한 연구 보고에 따르면 EPS의 통증과 가슴쓰림이 중복되는 경우가 많으나 가슴쓰림이 없는 EPS 환자에서도 다른 기능성 소화불량증에 비하여 24시간 보행성 식도산도 검사에서 비정상적인 위식도역류가 뚜렷하여 EPS 일부 환자는 위식도역류로 인하여 증상이 발생할 가능성을 제시하였다.

### 2. 기능성 소화불량증의 치료

기능성 소화불량증의 치료를 생활습관 조절과 약물 요법으로 시작한다. 기능성 소화불량증은 위약 효과가 높은 편이며 아직까지 약물 요법이 만족스러운 치료가 되지 못한다. 기능성 소화불량증의 치료에 사용되는 약제는 산분비 억제제, 위장관 운동 촉진제, 헬리코박터 파일로리 제균, 항우울제 등이 있으며 일부는 아직 식약처 허가를 받지 못하였다(Table 2).<sup>2</sup>

헬리코박터 파일로리 제균 요법은 22개의 무작위 대조 시험(randomized controlled trials, RCTs)의 4,896명 기능성 소화불량증 환자를 대상으로 평가되었다. 헬리코박터 파일로리 제균은 전반적인 소화불량 증상을 감소시키며 상대위험도 감소(relative risk reduction)는 0.91 (95% CI, 0.88-0.94), a number needed to treat는 12.5 (95% CI, 10-20)로 중증도의 효과로 보고되어 있다. 항생제의 기능성 소화불량증 증상 호전에 대한 효과는 헬리코박터 제균에 의할 수 있지만 상부위장관의 장내 세균총의 변화와 관련될 수도 있다.

산분비 억제제 중 프로톤펌프억제제(proton pump inhibitor, PPI)는 15개의 무작위 대조 시험(RCTs)의 5,389명의 기능성 소화불량증 환자에서 relative risk reduction이 0.83 (95% CI, 0.770-0.89), NNT 8 (95% CI, 614)로 위약 대비 우월한 효과를 보고하고 있다.<sup>3</sup> 우리나라에서 PPI는 소화불량증 환자의 치료약제로 허가되어 있지 않다.

위장관 운동 촉진제는 우리나라와 아시아에서 널리 사용되

**Table 1.** Diagnostic Criteria for Functional Dyspepsia in Rome III<sup>a</sup>

1. One or more of the following:
A. Bothersome postprandial fullness
B. Bothersome early satiation
C. Bothersome epigastric pain
D. Bothersome epigastric burning
2. No evidence of structural disease (including at upper endoscopy) that is likely to explain the symptoms

<sup>a</sup>Criteria fulfilled for the last 3 months with symptom onset at least 6 months prior to diagnosis.

고 있으나 서구에서 이루어진 위약 대비 대규모 임상시험에서는 효능이 입증되지 않았다. 일부 약제는 심장부작용이나 추체외로 부작용으로 인하여 사용이 제한되어 있다. 항정신성 약물도 위약 대비 우월성이 보고되고 있으나 그 효과는 3차병원으로 전원된 환자나 정신신체요소가 많은 환자에게서 있는 것으로 제한된다.<sup>4</sup>

### 3. 기능성 소화불량증의 치료에 대한 임상시험

#### 1) 임상 연구 설계

기능성 소화불량증 치료약제의 효능을 입증하기 위한 가장 이상적인 연구 설계는 위약 대조, 맹검, 병렬 그룹 디자인이다. 현재까지 전 세계적으로 위약 대비 우월성이 입증된 약제가 없기 때문에 기능성 소화불량증 신약에 대한 임상시험은 위약 대조군에 대하여 신약 효과의 우월성을 입증하는 것이 권고된다.<sup>5,6</sup>

환자를 일정한 임상시험 기간 동안 무작위로 위약군과 치료약제군으로 배정하여 임상시험 기간에 걸쳐 그 효과를 평가한다. 병행군 기본 구조로 치료를 받지 않는 baseline 기간과 치료 후 휴약기(washout) 기간을 포함한다. 일정한 전향적 기본 관찰기(prospective baseline observation period)를 권장하며 환자 선정에 적절한지 평가한다. 이 관찰 기간은 회상

이나 보고 오류를 피할 수 있고 환자가 증상을 현재 가지고 있는지를 확인할 수 있다. 이 기간 동안 기저 질병 상태를 문서화하고 시험 대상자가 데이터 수집 방법에 친숙해질 수 있게 한다. 증개의 시작은 사례 깊게 결정하며 무작위 배정 직전에 증상의 호전을 일으킬 수 있는 진단이나 검사를 피하는 것은 가치가 있다. 치료가 끝난 후에 일정 기간 추적 관찰은 치료약제의 견고성을 확인하기 위하여 필요하다. 추적 기간은 추정되는 약제의 치료 기전과 기능성 소화불량증 증상의 주기성과 관련하여 결정되는데 일반적으로 2주 이상의 추적 기간이 권고된다.

새로운 신약이 기존 약제에 비하여 편리하거나 경제적인 경우에 동등성 또는 비열등성 검증 시험을 추가로 시행할 수 있다. 표준 치료법과 대체 치료법 간의 차이에 대한 수용 범위에 대하여는 시험 전에 미리 결정되어야 하며, 이 정의에 대한 뚜렷한 근거가 제시되어야 한다. 비열등성이나 동등성을 입증하는 임상시험에서 대상 환자 수는 우월성 시험에 비하여 항상 크다.

#### 2) 임상 연구 기간

제3상 임상시험에서 활성약의 투여 기간은 약물의 약리학적 특성, 치료 기전과 기능성 소화불량증의 자연 경과를 고려하여 결정될 수 있다. 임상시험 설계의 기간은 반복 복용한

**Table 2.** Pharmacotherapy for Functional Dyspepsia

Type	Efficacy	KFDA permits	Usage status
<i>Helicobacter pylori</i> eradication	Superiority over placebo	No	
Acid-suppressive drugs			
Proton pump inhibitors	Superiority over placebo (in EPS)	No	
H <sub>2</sub> -receptor antagonists	Superiority over placebo (in EPS)	No	Indication for gastritis and GERD
Gastric acid neutralizing drugs	Not proven superiority over placebo	Yes	Can be prescribed for dyspepsia
Prokinetics			
Mosapride	Not proven superiority over placebo	Yes	
Itopride	Not proven superiority over placebo	Yes	
Domperidone	Some proofs of superiority over placebo	Yes	Usage restriction duo to QT prolongation
Metoclopramide, levopride	Not proven superiority over placebo	Yes	Risk of pseudo-parkinsonism
Acotiamide	Superiority over placebo (in PDS)	No	Permits in Japan/not yet introduced in Korea
Psychotropic drugs			
Buspirone	Superiority over placebo (in PDS)	No	patients with lots of psychosocial factors
Antidepressants-TCAs	Superiority over placebo (in EPS)	No	patients with lots of psychosocial factors
Antidepressants-SSRI	Not proven superiority over placebo	No	patients with lots of psychosocial factors
Herbal medicines			
Motilitone	Not proven superiority over placebo	Yes	Non-inferiority compared to Itopride
Iberogast	Superiority over placebo		
Rikkunshito	Not proven superiority over placebo		

KFDA, Korea Food & Drug Administration; EPS, epigastric pain syndrome; GERD, gastroesophageal reflux disease; PDS, postprandial distress syndrome; TCA, tricyclic antidepressant; SSRI, selective serotonin reuptake inhibitor.

약제의 효과에 대한 적절한 평가와 중단(withdrawal)의 영향을 평가할 수 있도록 충분히 길어야 한다. 제3상 임상시험에서 활성 치료의 단기 효과를 입증하기 위하여는 최소한 4주의 치료 기간이 추천되며 이보다 짧을 때에는 그것에 대한 정당화가 필요하다. 최근 여러 나라의 허가 기관에서는 기능성 위장장애의 임상시험 기간으로 12주를 표준으로 하고 있다. 만약 장기 사용용 염두에 두어 장기간 효과와 안정성을 획득하기 위하여는 최소한 6개월의 임상시험 기간을 권고한다.

### 3) 장기 연속 치료

장기간 지속적 사용을 목적으로 개발된 약물의 효능, 안전성 및 내약성을 확립하기 위하여는 잘 정의된 기능성 소화불량증 환자를 대상으로 대규모 이중 맹검, 위약 대조 임상시험을 실시해야 하고, 병렬 그룹 디자인의 임상시험이 채택되어야 한다. 기저 질병 활동을 문서화하고 환자가 데이터 수집 방법에 친숙해질 수 있게 하는 전향적인 기준 관찰(baseline observation)을 권장한다. 중개의 개입 시기는 주의 깊게 해야 한다.

장기간 연속 치료 임상시험은 어떤 반응이 지속되는지를 결정하고 치료 중단의 효과를 결정하기 위하여 충분하게 긴 시험 기간이 필요하다. 주기적이며 생명을 위협하지 않는 기능성 소화불량증의 자연 경과를 고려하면 최소한 6개월간의 활성 치료가 필요한 것으로 간주된다. 초기 이중 맹검 기간 이후에 치료 중단의 효과를 결정할 수 있도록 맹검 중단 연장 기간이 포함되어야 한다. 임상 설계는 장기 지속 효능, 중단(withdrawal), 반동효과(rebound effect) 및 안전성을 적절하게 평가할 수 있는가에 대한 정당성이 제시되어야 한다.

## 4. 임상시험 대상자

### 1) 선정기준

전향적 기능성 소화불량증 임상시험은 로마기준 IV의 기능성 소화불량증 진단기준을 만족시키는 대상자를 등록해야 한다. 임상시험에 참가하는 대상자는 임상시험에서 치료 반응을 평가할 수 있고 임상적으로 의미 있는 개선을 증명할 수 있도록 충분한 임상증상을 가지고 있어야 한다. 로마기준 IV에서는 괴로운(bothersome) 증상을 진단기준으로 하고 있으며 “괴로운”은 일상적인 활동을 방해하기 충분한 정도의 심한 증상으로 정의하고 있다. 증상들의 호전뿐만 아니라 증상의 “괴로운” 정도가 얼마나 완화되었는지 측정이 가능하여야 한다.

기능성 소화불량증 임상시험 대상 환자의 선정에서 고려해야 할 점은 기능성 소화불량증 환자는 다른 질환과의 중복이 흔하다는 점이다. 기능성 소화불량증은 위식도역류 질환, 변비 및 과민성 장증후군과 중복이 흔하다.<sup>7</sup> 임상에서 중복이 없는

순수한 기능성 소화불량증 환자는 많지 않다. 한 기능성 소화불량증 임상시험에서 과민성 장증후군이나 위식도역류 질환 환자를 배제하고 순수한 소화불량 환자만 선정하면 대상 환자의 규모가 11%로 매우 제한된다고 보고되어 있다.<sup>8</sup> 엄격하고 제한적인 선정기준은 시험약제의 효과를 입증하는데 도움이 되겠지만 이러한 임상시험의 결과를 한 가지 특정한 목표 증상만을 가진 환자가 드문 현실에서 임상 현장에 적용하기 어려울 것이다. 더구나 중복 환자는 한 가지 질환을 가진 환자보다 일상생활에 더 많은 지장을 가지고 있기 때문에 치료가 절실한 환자들이 많다. 그러므로 다른 기능성 위장 질환 중복 환자를 엄격하게 배제하는 것은 실효성에서 바람직하지 않다. 미국 소화기학회 백서에서도 중복 환자를 대상 환자에 포함하는 것에 동의하고 있다.<sup>6</sup> 과민성 장증후군이나 비미란성 위식도역류 질환과의 중복 환자도 기능성 소화불량증 증상이 더 현저하다면 포함이 가능하다. 중복 환자를 대상 환자에 포함하기 위하여는 그 선정과 제외기준을 임상시험 계획서에 명확하게 규정하여야 한다.

위마비증은 기능성 소화불량증과 증상만으로 구별할 수 없다. 따라서 위배출 검사를 시행하지 않는 한 위마비증 환자를 배제하기 어렵다. 그러나 현실적으로 위배출 검사를 대규모 임상 검사에서 시행하기 어렵고 위마비증의 유병률이 드문 것을 감안하면 기능성 소화불량증 임상시험의 배제기준으로 모든 연구에서 위배출 검사를 포함할 필요는 없다. 현재까지 컨센서스에 의하면 장시간 pH 검사나 위배출 검사는 기능성 소화불량증 증상의 병태생리를 이해하는데 중요하지만 기능성 소화불량증 임상시험의 선정기준이나 제외기준에 포함할 필요는 없다.

개발할 신약의 약리적 특성을 고려하여 대상 환자의 선정을 기능성 소화불량증 아형으로 제한할 수 있다. 증상 아형 또는 다른 기준(예: 증상의 중증도)에 따라 모집을 제한하는 임상시험은 왜 그러한 전략을 선정하였는지에 대한 명확한 근거를 제시하여야 하며, 이 임상시험의 결과는 넓은 정의의 기능성 소화불량증 환자에게 적용되지 않을 수도 있다는 것이 기록되어야 한다. 로마기준 IV에서는 PDS와 EPS의 두 가지 아형으로 분류하였지만 아형 분류가 병태생리를 명확하게 반영하지 못하고 이들 두 아형의 중복이 흔히 존재하므로, 기능성 소화불량증 임상시험에서 증상에 따른 아형 분류를 엄격하게 조정할 필요는 없다. 그러나 일부 임상시험에서 아형에 따른 치료약제 차이를 보인 경우도 있으므로 대안이 나오기까지는 로마 분류대로 PDS, EPS, EPS/PDS 중복군으로 분류하거나 병태생리에 보다 근접할 가능성이 있는 식사 관련 증상과 비식사 관련 증상으로 분류하는 것도 허용할 수 있다. 아형에 따라 가설을 생성하는 하위집단 분석이 수행될 수 있도록 설계되어야 한다.

대상 환자는 남성과 여성 모두 포함해야 하며 효과의 일관성을 보여줄 수 있는 의의 있는 하위 그룹 분석을 허용할 수 있을 만큼 많은 숫자를 모집해야 한다. 환자는 시험 기간에 평소 식이 요법과 생활 방식을 유지할 수 있어야 한다.

## 2) 제외기준

소화불량증의 원인이 규명된 환자는 배제한다. 예를 들어 서 내시경 검사에서 이상 소견의 진단, 규칙적인 NSAID의 복용, 가슴쓰림이 뚜렷한 환자는 배제되어야 한다.

헬리코박터 파일로리 세균의 감염이 확인된 환자는 최근 교토합의서나 로마위원회에서 주장하듯이 헬리코박터 파일로리 관련 소화불량증으로 기능성 위장 질환과 구별하고 있으므로 기능성 소화불량증 임상시험에서 헬리코박터 파일로리 감염 여부를 반드시 확인하고 감염된 환자를 임상시험 대상자에서 제외하는 것이 권고된다. 그러나 우리나라에서는 헬리코박터 파일로리 감염 유병률이 높고 제균 요법이 표준 치료법으로 허가되어 있지 않기 때문에 헬리코박터 파일로리 감염자를 기능성 소화불량증 임상시험에서 일률적으로 배제하는 것에 대한 반대 의견도 있다. 따라서 헬리코박터 파일로리 감염자를 대상 환자에 포함해야 할 경우는 그 이유를 명시하고 감염 유무에 따른 효과나 안전성에 대한 계층 분석을 할 것을 권고한다.

현재까지 이루어진 기능성 소화불량증 임상시험은 정상 내시경 검사 소견과 증상 진단으로 환자를 임상시험에 등록시켰는데, 이 경우 EPS로 선정된 위식도역류 질환 환자나 PDS와 유사한 위마비증 환자가 포함될 수 있다. 위식도역류 질환 환자를 배제하기 위하여 48시간 bravo pH 검사나 PPI run in을 하거나 위배출 검사로 위마비증 환자를 배제하는 엄격한 선정 기준을 택하는 유별난 임상시험도 생각할 수 있다. 현재까지의 컨센서스에 의하면 pH 검사나 위배출 검사는 기능성 소화불량증 증상의 병태생리를 이해하는데 중요하지만 기능성 소화불량증 임상시험의 선정이나 제외기준으로는 사용되지 않는다.<sup>9</sup>

## 5. 임상시험평가변수

기능성 소화불량증은 바이오마커가 없기 때문에 어떠한 평가변수(outcome variable)가 신약의 효과를 평가하기 좋은지는 아직 확정되지 않았다. 평가변수는 1차 평가변수와 2차 평가변수로 구분할 수 있다. 평가변수들은 연구가 개시되어 구체적인 자료를 수집하기 전에 미리 평가변수의 관측 방법, 측정기준, 평가기준에 대한 명확하고 구체적인 내용을 정해야 한다.

## 1) 환자 결과 보고(Patient Reported Outcomes, PROs)의 정의 및 사용

기능성 위장관 질환 임상시험에서 증상의 평가 및 정량화는 피험자 선정과 유효성 평가를 위한 주요 평가 방법이다. 기능성 위장관 질환 환자는 환자가 경험하는 증상 때문에 의료기관을 찾게 되므로 환자의 질병 경험을 측정하기 위한 구조화된 포맷의 환자 결과 보고(PRO)가 필요하다. PRO는 환자의 건강 상태에 대한 환자의 반응을 임상이나 다른 사람에 의하여 평가하지 않고 환자로써 스스로 얻어진 보고로 평가하는 것이다.<sup>10</sup> 위식도역류 질환이나 과민성 장증후군 임상시험에서 질병 특이적 가이드라인 등이 소개된 바 있다.<sup>11</sup> 그러나 아직까지 기능성 소화불량증에 대한 가이드라인이 정립되어 있지 않은데 이는 증상 복합체로 진단하는 기능성 소화불량증 진단과 치료 성과에 대한 확실한 자료가 선정되어 있지 않기 때문이다.

증상에 근거하여 피험자를 선정하거나 평가할 때마다 철저하고 충분히 검증된 증상 평가 방법을 사용해야 한다.<sup>10</sup> 문헌고찰 혹은 전문가 의견을 기반으로 유효성 평가지표(endpoint) 모델을 구축한다. 적절한 평가 대상을 선정한 후 개별 증상 아이템을 모으고 그룹화하여 기능성 소화불량증의 개념적인 구조를 세운 뒤, 환자의 질적 평가 등을 통하여 이를 확인한 후 초안을 완성한다. 내적일치도나 검사-재검사 등을 이용한 신뢰도 검증을 실시하고, 내용 타당도와 구조 타당도 검증을 시행한다. 유효성과 연관되어 “ability to detect change”를 추정하도록 권장하고 있는데 이는 임상적으로 유의한 그룹 간 차이나 중재에 따른 개인 간 변이 등 임상적으로 의미 있는 유효성 평가지표의 차이를 알아내는 능력을 의미한다. 환자의 상태를 평가하는 측정 도구를 개발하는데 있어서, 평가 도구의 타당성을 평가하기 위한 지표 중 하나로 minimally clinical important differences (MCIDs)가 있다. MCIDs는 진료 성과에 있어서 “임상적으로 의미 있는 최소한의 차이”를 의미한다. 이는 임상적보다는 환자 중심으로 진료 성과를 평가하는 것으로, “임상적으로 의미 있는 최소한의 차이”라는 표현이 객관적이지 못함에 따라, clinically important difference, minimal important difference 등 다양한 정의와 방법으로 활용되고 있다. PRO를 어떻게 개발하고 개발 단계를 기록하는지에 대한 안내는 Table 3에 요약하였다.<sup>12</sup>

## 2) 1차 평가변수

지금까지 이루어진 기능성 소화불량증 임상시험에서 사용된 1차 유효성 평가변수는 다양하고 아직까지 표준화되어 있지 않다.<sup>13</sup> 임상시험의 평가지표(endpoint)로 적합한 바이오마커나 진단 검사는 아직 없다. 평가지표로서 현재까지 검토

된 객관적 검사는 위배출 검사와 nutrient/saline load test이다. 두 검사 모두 소화불량 증상의 병태생리를 이해하는데 유용하지만 기능성 소화불량증 시험의 평가지표로 충분히 검증되지 않았다. 기능성 소화불량증은 증상들의 수집으로 정의되므로 가장 적절한 평가지표는 증상 호전에 근거한다. 기능성 소화불량증 임상시험에 사용할 수 있는 검증된 PRO 측정도구는 아직 개발되어 있지 않다. 이에 대하여 미국과 유럽의 허가 관청에서도 1차 유효성변수에 대한 논의를 계속하고 있어 국제적인 표준안이 마련되면 제시된 안에 따라 국내 지침의 개정이 필요할 것이다. 국제적인 표준안이 마련되어 있지 않은 현실에서 최종안이 나오기 전까지 잠정안이 필요한 실정이다.

기능성 소화불량증 증상 평가를 위하여 여러 평가도구들이 개발되어 있다. 가장 많이 사용된 평가지표는 전반적 평가(overall treatment evaluation, OTE), 이분법적 답변 방식(binary outcome measures), 종합증상 설문지이다.

#### (1) 전반적 치료 평가(OTE)

기능성 소화불량증 치료제의 임상시험에서 전반적인 상태의 변화를 하나의 질문에 환자가 응답하는 방법이 많이 사용되었다. 전반적 치료 평가는 실제 임상에서 의사가 환자의 치료 효과를 평가하는 상황과 가장 유사하다는 장점이 있다. 그러나 개별 증상에 대한 평가가 아니기 때문에 호전이나 악화의 이유를 알 수가 없고, 개선 정도를 정량화할 수 없어, 회상오류(recall bias) 가능성으로 인하여 Food and Drug Administration (FDA)에서는 전반적 상태에 대한 단일 질문을 1차 평가변수로 더 이상 추천하지 않고 있다. 국내에서 motilitone 임상시험에서 OTE를 7점 리커드 척도로 사용되었으며, 최근에 이루어진 acotinamide 임상시험에서도 OTE를 7점 리커드 척도로 사용되었다.<sup>14,15</sup>

#### (2) 이분법적 답변 방식(binary outcome measures)

이분법적 답변 방식은 지금까지 여러 높은 수준의 대규모 functional gastrointestinal disease (FGID) RCTs에서 1차

유효성 평가변수로 사용되었다. 로마 III 안내서에서는 응답하기 쉽고 해석이 명확한 장점 때문에 이분법적 답변을 1차 유효성 평가변수로 추천하였지만 최근의 등록 임상시험에서 이분법적 답변의 사용을 더 이상 추천하지 않는다. 문제점으로는 ① 전부는 아니지만 어떤 환경에서는 이분법적 답변은 일부 baseline severity에 의하여 좌우된다. ② 이분법적 답변은 증상에서 MCID나 health-related quality of life를 검출하지 못할 수 있다. ③ 이분법적 답변은 주요한 질병 영역을 추적하는 능력이 없고 임상 질환 아형 간 감별할 수 없다. ④ 이분법적 답변은 환자 포커스 그룹에서 유래된 것이 아니다.<sup>16,17</sup> 이분법적 답변은 등록 임상에서 1차 유효성 평가변수로 더 이상 인정하지 않지만 2차 평가변수로 허용하고 있다. 비등록 임상에서는 1차 평가변수로 아직도 인정되고 있다.

#### (3) 종합증상 설문지

다문항증상 설문지의 사용이 가능한 대체 평가지표로 제안되어 있다. 현재까지 사용 가능한 소화불량증 종합증상 설문지들은 장점이 있고 여러 단계의 검증이 이루어져 있지만 평가지표 기준을 만족시키지 못한다. 일부는 focus group에서 개발되었지만 검증 과정에서 cognitive debriefing을 시행한 것은 드물다. 대부분은 소화불량증과 위식도역류 질환과 같이 혼재된 대상 집단에서 개발되었으므로 fit-for-purpose에 맞지 않는다. 일부 평가 도구는 구조 타당도와 반응도에 대한 검증이 이루어지지 않았고, 대부분 회상 기간이 길고 MCID 기준을 제시하지 않았다. 빈도와 강도만 측정되어 있고 다른 attribute of potential relevant가 없다. 최근 개발된 the Leuven Postprandial Distress Scale (LPDS)은 FDA와 European Medicines Agency 권고안에 맞추어 개발과 검증이 완료되었다.<sup>18</sup> LPDS는 증상 패턴과 인자 분석을 통하여 추출한 3가지 증상(식후 충만감, 조기 포만감, 상복부 팽만감)으로 구성되었으며 리커트 척도(0-4)로 평가한다. LPDS score는 민감도와 신뢰도가 높으며 European Medicines Agency의 지지를 받아 앞으로 PDS의 임상시험에 사용될 수 있을 것으로 기대된다. 그러나 EPS 아형이나 EPS-PDS over-

**Table 3.** Developing PROs for FGIDs

1. For FGID registration trials, the FDA and EMA require fit-for-purpose PROs that are specifically designed for the precise target population being studied.
2. As FGIDs are an international phenomenon, it is important to create PROs that can be accurately translated into different languages.
3. All items in a PRO should be carefully tested in cognitive interviews with FGID patients who are representative of the target population.
4. For clinical trials, it is generally recommended to use endpoints with a 1 day recall or less.
5. For FGID registration trials, PROs must have empirical evidence of validity and reliability and must have a priori MCID benchmarks and responder definitions.

PROs, patient reported outcomes; FGID, functional gastrointestinal disease; FDA, Food and Drug Administration; EMA, European Medicines Agency; MCID, minimal clinically important difference.

lap 환자군에서 LPDS는 검증되어 있지 않다.

현재 사용 가능한 소화불량증 종합증상 설문지들은 등록 임상시험의 1차 평가변수로 적합하지 않으며 2차 평가변수로 사용될 수 있다. 더구나 국내에서 검증이 완전히 이루어진 평가 도구들이 없는 실정이다. 다국가 임상시험에서는 서로 다른 언어로 표현되어 검증이 이루어진(cognitive debriefing) 증상 설문지가 사용되어야 한다. 현재까지 외국에서 개발된 소화불량증 종합증상 설문지들이 임상시험에서 사용되었지만 제대로 검증이 이루어지지 않았다는 문제점이 항상 제기되어 왔다.

### 3) 2차(지원) 유효성 평가변수

2차 유효성 평가변수는 1차 평가변수 측정과 개인의 증상간의 일치를 보여줌으로써 결과를 강화시키기 위하여, 중재의 기전을 강조하기 위하여, 안정성이나 치료의 비용 효과의 평가, 환자가 가장 또는 최소한의 이점을 갖는 것을 예측하는 변수를 알기 위하여 선택할 수 있다. 임상시험의 계획서에 2차 유효성 평가변수의 선택에 대한 정당화가 제시되어야 한다.

건강 관련 삶의 질은 가장 중요한 2차 유효성 평가변수의 하나로 간주된다. 건강 관련 삶의 질 평가 도구는 FDA PROs 개발 기준을 만족해야 한다.<sup>10</sup> 현재까지 국내에서 개발된 건강 관련 삶의 질 평가 도구는 없지만 외국에서 개발되어 국내 연구자들에 의하여 번역, 재현성, 구조, 타당도, 반응도에 대한 검증이 이루어진 Nepean Dyspepsia Index-Korean version (NDI-K)의 사용을 고려할 수 있다.

### 4) 한국형 소화불량증 종합증상 설문지의 개발과 잠정안

기능성 소화불량증의 임상시험에서 환자의 증상 평가 도구는 환자가 쉽게 설명하고 이해할 수 있어야 하며 환자의 상태 변화에 민감해야 한다. 변화 측정에는 호전뿐만 아니라 악화도 포함되어야 한다. 증상 설문지의 각 증상의 중증도는 시각적 통증 강도 비율 등급(visual analogue scales, VAS) 또는 여러 포인트의 리커트 척도를 사용해야 한다. 리커트 척도에

서 평가 포인트 수는 사용 방법마다 다양할 수 있지만 최소한 5개 포인트는 포함되어야 한다. 소화불량증의 증상은 최소한 다음 7가지 증상이 포함되어야 한다; 상복부 통증(stomach pain), 상복부 쓰림(upper abdominal burning), 메스꺼움(nausea), 팽만감(bloating), 식후 충만감(postprandial fullness), 조기 포만감(early satiety)과 트림(burping/belching). 적합한 기능성 소화불량증 증상 평가 도구는 소화불량증의 7가지 증상을 포함하여 강도와 빈도 외에도 괴로운 정도(bothersomeness)의 척도를 포함한 증상 점수의 개발을 권고하고 있다.<sup>6</sup>

소화불량증 증상은 단어로 표현하기 힘들므로 그림(pictogram)을 통하여 더욱 쉽고 효과적으로 느낌을 기술할 수 있다. 그림을 추가하면 단순 설명 문항보다 환자가 증상 개념을 이해하는 데 도움을 준다.<sup>19</sup> 그림은 환자와 소통을 원활하게 해 줄 가능성을 가지고 있으며 일부 장벽을 극복할 수 있다. 모든 그림은 이해를 확인하기 위하여 여러 환자들을 통하여 검증되어야 한다.

국내 기능성 소화불량증 임상시험의 1차 유효성 변수에 대한 잠정안으로 환자의 치료에 대한 전반적인 평가인 OTE (7단계 리커트 척도)와 증상 종합점수의 변화 두 가지를 1차 평가변수로 사용할 것을 권유한다. 통계적으로 유효한 변화가 양변수에 모두 확인되어야 한다. 통계적으로 의미 있는 결과는 임상적 정당성이 확인되어야 한다. 아직까지 기능성 소화불량증의 임상적 평가지표를 평가하는 검증된 관측도구(outcome measure)가 인정되어 있지 않기 때문에 임상시험 담당자는 관측도구 선택의 정당한 이유를 제출해야 한다.

우리나라에서는 아직까지 FDA guidance에서 권고하는 검증이 완료된 PROs가 개발되어 있지 않지만 NDI-K의 증상 평가지는 재현성에 대한 검증이 이루어졌으며 7가지 소화불량증 증상이 포함되어 있다. 최근 국내 임상시험(Motilitone)에서 사용되면서 일부 구조 타당도 검정이 이루어졌으므로 NDI-K의 증상 평가지의 사용을 고려해 볼 수 있다.<sup>20</sup> 최근 정혜경/최명규가 개발한 상부위장관 증상 설문지의 사용도 대

**Table 4.** Defining MCID and Responder Status

1. The main result of the study must be based on the evaluation of the primary outcome measure, as stated in the protocol before the study begins. The primary outcome should be stated in absolute numbers and should include a 95% CI.
2. When trial results are expressed in terms of the proportion of responders in each arm of the study, a responder should be defined as someone showing an improvement equal to or greater than the MCID.
3. When continuous measures are used to assess trial results, the differences between treatment arms should be compared with the MCID.
4. When reporting p-values, actual values and not thresholds should be provided. Investigators are encouraged to report the NNT in clinical trials, and the NNH can be calculated based on the risk of adverse effects and can be weighed against the NNT.
5. The primary analysis should be an ITT analysis and must include all patients randomized.

MCID, minimal clinically important difference; CI, confidence interval; NNT, number needed to treat; NNH, number needed to harm; ITT, intention-to-treat.

안으로 제시될 수 있지만 위식도역류 질환 환자만을 대상으로 검증이 이루어졌기 때문에 소화불량증 환자를 대상으로 하는 추가 검증이 필요하다.

#### 5) 자료수집 전략

기능성 소화불량증 임상시험 가이드라인이 확정되지 않은 상태에서 일차 증상 관측도구 잠정안(interim primary symptom outcome)을 평가하기 위하여 매일 일기 사용을 권고한다. 정확한 증상 패턴과 강도를 평가하기 위하여 매일 기록을 권고하며, 일일 평가 일기 사용으로 회상 오류를 최소화할 수 있다. 종이 일기는 방문 직전에 한꺼번에 적을 우려가 있기 때문에 종이 일기보다는 전자 일기의 사용을 추천한다. 전자 일기는 알람을 제공하기 때문에 일지의 완성도를 높일 수 있다.

#### 6. 반응자의 정의

임상시험 계획서는 임상적으로 관련 있다고 간주될 수 있는 만큼 뚜렷한 1차 평가변수의 개선으로 반응자를 사전에 정의해야 한다(a priori). 임상시험은 점수의 평균 변화보다는 반응자로 정의될 수 있는 만큼 규정된 개선량을 달성한 환자의 비율을 비교해야 한다(Table 4).<sup>12</sup>

반응자 또는 반응을 달성하는 시간 간격에 대해서도 사전적으로 지정되어야 한다. 만약 대상자가 치료 기간의 최소 50%의 주 또는 일(예를 들어, 6/12주 또는 42/84일) 동안에 미리 정의한 개선에 도달하였다면, 대상자는 전반적인 반응자로 분류한다. 임상시험은 임상시험 등록시 선정기준에 속하는 개별 증상의 변화에 대한 측정법이 포함되어야 한다. 종합적인 증상 척도를 사용할 경우 이러한 척도에서 최소한의 임상적 변화에 대한 검증이 필요하고, 반응자 비율에서 최소한의 임상적 관련이 있는 변화는 사전에 정의되어야 한다.

#### 7. 부작용과 안정성 모니터링

모든 부작용을 보고하는 것이 중요하다. 신약을 테스트할 때에는 새로운 기대하지 않았던 부작용이 일어날 수 있기 때문이다. 기능성 소화불량증은 생명을 위협하지 않는 질환으로써 약물 후보 물질에 대한 안전성이 가장 중요하게 고려되어야 한다. 기능성 소화불량증은 만성 질환으로 여겨지므로 이를 대상으로 하는 약물의 적절한 장기 안전성 평가가 필요하다.

## 결 론

본 연구에서 수행한 기능성 소화불량증 치료제 임상시험 가이드라인은 임상시험 연구자 및 평가자에게 최신 임상시험과 관련된 규정을 국문으로, 효율적으로 정리하여 제공하므로

활용가치가 높다. 국내에서 소화불량증 치료제의 시장 규모는 2천억 원 이상으로 매우 높아 국내의 여러 제약회사들이 치료제를 개발하고 있거나 제형 개선, 복합제 개발이 활발하게 이루어지고 있으나 임상시험 가이드라인이 없어 혼란을 주고 있는 현실에서 큰 도움이 될 것이다. 아직까지 세계적 표준안이 마련되지 않은 현실에서 우리나라의 독자적인 가이드라인이 개발됨에 따라 제약회사는 훨씬 쉽게 임상시험을 설계하고 임상시험 결과의 신뢰도가 향상되어 국제 경쟁력이 강화될 수 있다. 식약처는 임상시험 승인 과정이 표준화됨으로써 임상시험 심사 업무의 투명성과 일관성이 보장될 수 있다. 환자 단체들도 가이드라인을 활용하여 임상시험의 승인과 수행을 감시할 수 있어 환자 권익 보호에 도움이 된다. 임상시험에 대한 가이드라인은 국내 임상시험의 활성화, 세계적 수준의 임상시험 수행, 양질의 임상시험 결과의 보고 등으로 이어지며, 이는 결과적으로 한국 임상시험의 국제적 수준으로의 향상을 이룰 수 있을 것으로 기대된다.

## REFERENCES

1. Stanghellini V, Chan FK, Hasler WL, et al. Gastrointestinal disorders. *Gastroenterology* 2016;150:1380-1392.
2. Moayyedi PM, Lacy BE, Andrews CN, Enns RA, Howden CW, Vakil N. ACG and CAG clinical guideline: management of dyspepsia. *Am J Gastroenterol* 2017;112:988-1013.
3. Pinto-Sanchez MI, Yuan Y, Bercik P, Moayyedi P. Proton pump inhibitors for functional dyspepsia. *Cochrane Database Syst Rev* 2017;3:CD011194.
4. Ford AC, Luthra P, Tack J, Boeckstaens GE, Moayyedi P, Talley NJ. Efficacy of psychotropic drugs in functional dyspepsia: systematic review and meta-analysis. *Gut* 2017;66:411-420.
5. Irvine EJ, Tack J, Crowell MD, et al. Design of treatment trials for functional gastrointestinal disorders. *Gastroenterology* 2016;150:1469-1480.e1.
6. Vakil NB, Howden CW, Moayyedi P, Tack J. White paper AGA: functional dyspepsia. *Clin Gastroenterol Hepatol* 2017;15:1191-1194.
7. Park JM, Choi MG, Cho YK, et al. Functional gastrointestinal disorders diagnosed by Rome III questionnaire in Korea. *J Neurogastroenterol Motil* 2011;17:279-286.
8. Vakil N, van Zanten SV, Chang L, et al. Comprehension and awareness of symptoms in women with dyspepsia. *Aliment Pharmacol Ther* 2005;22:1147-1155.
9. Howden CW, Vakil NB. American gastroenterological association drug development conferences: a model for future collaborations. *Clin Gastroenterol Hepatol* 2017;15:1195-1196.
10. Guidance for Industry. Patient-Reported Outcomes Measures: Use in Medical Product Development to Support Labeling Claims. [Internet]. Rockville: US Department of Health and Human Services Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Devices and Radiological Health (CDRH); 2009 [updated 2009 Dec; cited 2018 Sep 3].



Available from: <http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf>

11. Guidance for Industry: Irritable Bowel Syndrome - Clinical Evaluation of Drugs for Treatment. [Internet]. Rockville: US Department of Health and Human Services Food and Drug Administration, Center for Drug Evaluation and Research (CDER); 2012 [updated 2012 May; cited 2018 Sep 3]. Available from: <http://www.fda.gov/downloads/Drugs/Guidances/UCM205269.pdf>
12. Drossman DA, Chang L, Chey WD, Kellow J, Tack J, Whitehead WE. Rome IV functional gastrointestinal disorders: disorders of gut-brain interaction. 4th ed. Raleigh: Rome Foundation, Inc., 2016.
13. Ang D, Talley NJ, Simren M, Janssen P, Boeckstaens G, Tack J. Review article: endpoints used in functional dyspepsia drug therapy trials. *Aliment Pharmacol Ther* 2011;33:634-649.
14. Choi MG, Rhee PL, Park H, et al. Randomized, controlled, multi-center trial: comparing the safety and efficacy of DA-9701 and itopride hydrochloride in patients with functional dyspepsia. *J Neurogastroenterol Motil* 2015;21:414-422.
15. Matsueda K, Hongo M, Tack J, Saito Y, Kato H. A placebo-controlled trial of acotiamide for meal-related symptoms of functional dyspepsia. *Gut* 2012;61:821-828.
16. Corsetti M, Tack J. FDA and EMA end points: which outcome end points should we use in clinical trials in patients with irritable bowel syndrome? *Neurogastroenterol Motil* 2013;25:453-457.
17. Guideline on the evaluation of medicinal products for the treatment of irritable bowel syndrome. [Internet]. London: European Medicines Agency; 2013 [updated 2014 Sep 25; cited 2018 Sep 3]. Available from: [https://www.ema.europa.eu/documents/scientific-guideline/guideline-evaluation-medicinal-products-treatment-irritable-bowel-syndrome-revision-1\\_en.pdf](https://www.ema.europa.eu/documents/scientific-guideline/guideline-evaluation-medicinal-products-treatment-irritable-bowel-syndrome-revision-1_en.pdf)
18. Carbone F, Vandenberghe A, Holvoet L, et al. Validation of the leuven postprandial distress scale, a questionnaire for symptom assessment in the functional dyspepsia/postprandial distress syndrome. *Aliment Pharmacol Ther* 2016;44:989-1001.
19. Tack J, Carbone F, Holvoet L, Vanheel H, Vanuytsel T, Vandenberghe A. The use of pictograms improves symptom evaluation by patients with functional dyspepsia. *Aliment Pharmacol Ther* 2014;40:523-530.
20. Cho YK, Choi MG, Kim SH, et al. The effect of mosapride on quality of life in functional dyspepsia. *Korean J Gastroenterol* 2004;43:160-167.