

Reliability and Validity of Patient Oriented Eczema Measure (POEM) Korean Version

Yoon-Ha Kim¹, Jong-Min Han¹, In-Chul Jung², Weechang Kang³, Yang-Chun Park¹

¹Department of Internal Medicine, College of Korean Medicine, Daejeon University, Daejeon, Republic of Korea

²Department of Neuropsychiatry, College of Korean Medicine, Daejeon University, Daejeon, Republic of Korea

³Department of Statistics, College of Business Administration, Daejeon University, Daejeon, Republic of Korea

Objectives: We developed a Korean translation of this questionnaire by sequential forward-and-backward translation. The purpose of this study is to validate the Korean version of the POEM, the “POEM-K” by Korean patients with atopic dermatitis (AD).

Methods: A single-center prospective study was conducted with 50 participants diagnosed with AD. The POEM was translated to Korean version by an expert panel. Scoring Atopic Dermatitis (SCORAD) and Short Form 36 Health Survey (SF-36) were used as external comparators.

Results: Twenty men and thirty women between the ages of 18 and 63 participated in the study. The test-retest reliability of the total POEM-K was estimated using the intra-class correlation coefficient (ICC) that showed a strong agreement (ICC = 0.72). By using Pearson correlation coefficients, we compared the POEM-K to SCORAD and yielded a concurrent validity that showed a significant result. The responsiveness of the POEM-K was represented by the effect size (ES) of 1.41 and statistically significant ($p = 0.004$).

Conclusions: The Korean version of the POEM is a reliable, valid, and responsive disease-specific questionnaire for assessing the symptoms and quality of life of Korean patients with AD.

Key Words : Atopic dermatitis, Patient Oriented Eczema Measure, Quality of life, Reliability, Validity

Introduction

Atopic dermatitis (AD) is a chronic eczematoid dermatosis caused by an abnormal immune reaction.^{1,2} Prevalence rates of AD vary from 8 % to 20 %.³ According to the analysis of National Health Insurance (NHI) health-insurance payment data, the average number of annual patients is 1.04 million in Korea.⁴ The number of total patients decreased slightly from 2008 to 2012. However, the number of inpatient admissions increased by 1.5-fold from 894

in 2008 to 1,367 in 2012, which indicated the increasing trend of severe cases.⁴ Especially, School-age children are seriously affected with respect to the growth, body development, family relations, and social life, with emotional disturbances and insomnia the foremost issues.⁵

Regarding AD, reliable assessment tools are necessary to evaluate the effects of therapies and monitor the progress of disease in clinical practice or research. Schmitt et al. reviewed the validity of 20 measures of AD in 45 eligible articles. Only Scoring

• Received : 24 October 2016 • Revised : 27 December 2016 • Accepted : 27 December 2016
• Correspondence to : Yang-Chun Park
Department of Internal Medicine, Dunsan Korean Medicine Hospital,
College of Korean Medicine, Daejeon University
75, Daedeok-daero 176beon-gil, Seo-gu, Daejeon 35235, Republic of Korea
Tel : +82-42-470-9126, Fax : +82-42-470-9486, E-mail : omdpyc@dju.kr

of Atopic Dermatitis (SCORAD), Eczema Area and Severity Index (EASI), and Patient Oriented Eczema Measure (POEM) showed validity and reliability.⁶⁾ While SCORAD and EASI are primarily designed for clinical trial work and too time-consuming, POEM simply consists of 7 questions, which allow for simple application in clinical research as well as in general practice. The SCORAD and EASI are evaluated through an interview between a doctor and patient regarding the extent of lesion dispersal and severity, where the language barrier and cultural difference are minor concerns. The POEM, however, requires patients to self-report their symptoms, and therefore needs to be properly translated and verified. With this background, we report the development of the Korean version of the POEM (“POEM-K”) to prove the reliability and validity.

Methods

1. Patient Oriented Eczema Measure (POEM) and elaboration of Korean version

The POEM includes 7 questions regarding the number of days each symptom appeared. All questions are based on 5-point severity scales (Table 1).¹⁰⁾ The POEM-K was translated by an American physician, who can speak both English and Korean. A review committee comprising a Korean linguist reviewed the translation and then produced a Korean

draft, which was retranslated by the third bilingual translator. The review committee then reviewed the retranslated document and finalized the text through a cognitive debriefing process.

2. Study design

This study was approved by the institutional review board (IRB) of the Korean Medicine Hospital of Daejeon University, and the approval number is DJOMC-106. All participants understood the purpose and method of this study and agreed to participate with written informed consents. Participants were recruited from outpatients for 184 days from August 1, 2013, to January 1, 2014.

Patients diagnosed with AD according to the criteria of Hanifin and Rajka⁷⁾ by research physicians were included in the study. Exclusion criteria included age younger than 18 years or older than 65 years and diseases that seriously affected participants' quality of life (QoL), such as liver cirrhosis, chronic renal failure, and chronic heart failure.

The sample size was calculated so that the number of participants, who have a score greater than 1 in every question, could be at least 20 in the study. The information needed to estimate the sample size was based on the study of Charman et al., where the lowest occurrence rate was reported to be 40% for the item “weeping”.⁸⁾ Therefore, at least 50 ($=20/0.4$) participants were required.

Participants were asked to complete the POEM-K,

Table 1. Content of the Patient Oriented Eczema Measure (POEM)¹⁰⁾

Questions
1. Over the last week, on how many days has your/your child's skin been itchy because of the eczema?
2. Over the last week, on how many nights has your/your child's sleep been disturbed because of the eczema?
3. Over the last week, on how many days has your/your child's skin been bleeding because of the eczema?
4. Over the last week, on how many days has your/your child's skin been weeping or oozing clear fluid because of the eczema?
5. Over the last week, on how many days has your/your child's skin been cracked because of the eczema?
6. Over the last week, on how many days has your/your child's skin been flaking off because of the eczema?
7. Over the last week, on how many days has your/your child's skin been felt dry or rough because of the eczema?

Responses are scored as follows: no days, 0; 1–2 days, 1; 3–4 days, 2; 5–6 days, 3, every day, 4.

SCORAD and Short Form-36 (SF-36) at the first visit, and the POEM-K and the global rating of change (GRC) 2 and 6 weeks later (Fig. 1). The GRC scores were determined to evaluate self-perceived changes in symptoms since the first visit. Responses were scored from +7 (a very great deal better) to -7 (a very great deal worse), where 0 indicate no change.⁹⁾ The participants were allowed to take any medicines during this study.

3. Statistical analysis

Data were reported as means \pm standard deviation (SD) unless specified otherwise. The level of significance for all statistical tests was set at $p < 0.05$. The test-retest reliability was estimated by the Bland and Altman Plot¹⁰⁾ to calculate the intra-class correlation coefficient (ICC) in 32 people with no change or a slight change ($|GRC| \leq 1$) of symptoms at the second or third visit. The concurrent validity was evaluated by estimating the Pearson correlation coefficient between POEM-K and other instruments (SCORAD and SF-36). For responsiveness, we analyzed the differences of the POEM-K scores between the visits with GRC scores greater than 3 and their former visits. Here, the GRC score of minimum 4 was considered a significant improvement.¹¹⁾ The responsiveness was evaluated with the effective size (ES), estimated as the mean difference of POEM-K scores divided by the SD of

the POEM-K scores of the former visits. Statistical analyses were performed by the SAS software (SAS Institute, Inc., Cary, NC, USA).

Results

1. Demographics of participants

Fifty participants were initially enrolled in the study. One participant did not attend the center at the third visit. All participants who replied once or more were included in the analysis. The age of participants ranged from 18 to 63, with an average age of 29.2 ± 12.9 . The numbers of female and male participants were 30 and 20, respectively. The duration of AD ranged from 2 years to 44 years, with a mean period of 15.7 ± 8.0 years. Regarding severity of eczema, most of participants (96%) conformed to mild-moderate, while the remainder (4%) had severe eczema (Table 2).

2. Test-retest reliability of POEM-K

The average GRC values at the second and third visits were -0.6 ± 2.2 and -0.5 ± 3.1 , respectively. At each visit, patients answered that their symptoms, on average, slightly worsened. The subjects of the test-retest reliability analysis of the POEM-K were those with no change or a slight change ($|GRC| \leq 1$) in symptoms at the second or third visit. This group was composed of 7 patients with a GRC score

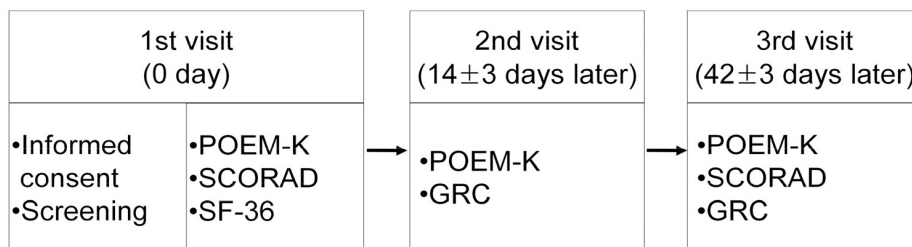


Fig. 1. Flow of study.

POEM-K: Korean version of Patient Oriented Eczema Measure; SCORAD: Scoring of Atopic Dermatitis index; SF-36: Short Form 36 Health Survey; ADPIQ: Atopic Dermatitis Pattern Identification Questionnaire; GRC: global rating of change.

Table 2. Participant Characteristics at First Visit

Variables	Value
Gender (number/total)	
Male	20/50
Female	30/50
Age, years	
Mean	29.2 ± 12.9 ^{a)}
Range	18-63
Blood pressure (mmHg)	
Systolic blood pressure	120.7 ± 14.7
Diastolic blood pressure	73.2 ± 10.2
Pulse(frequency/min)	74.6 ± 10.3
Temperature (°C)	36.3 ± 0.3
Breath (frequency/min)	19.5 ± 0.9
Duration of AD (year)	15.7 ± 8.0
Severity of Eczemab) (number/total)	
Mild-moderate	48/50
Severe	2/50

a): mean ± standard deviation.

b): mild-moderate ≤ 50 of SCORAD, severe > 50 of SCORAD

of -1 (very slightly worse), 16 patients with a GRC score of 0 (approximately the same), and 9 patients

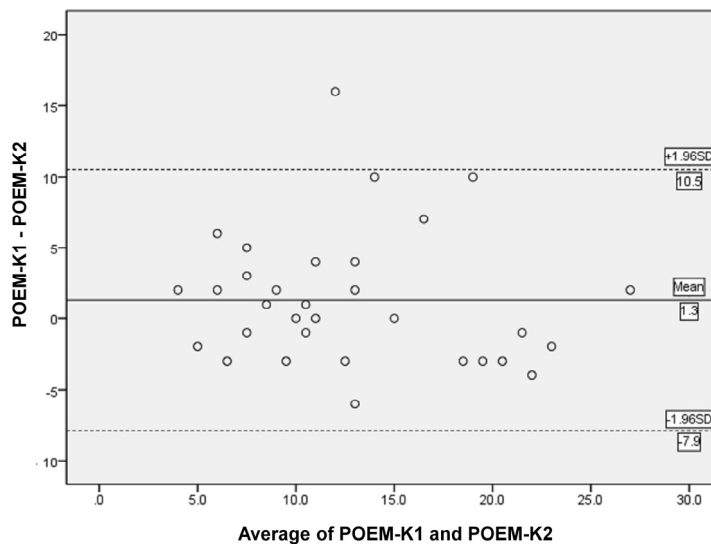
with a GRC score of 1 (very slightly better): the total number of participants was 32.

The POEM-K score of the patients with a slight change of symptoms decreased by an approximate average of 1.3 ± 4.7 points; however, this change was found statistically insignificant (p -value: 0.126 via the paired t -test). All points but one were contributed in the mean $\pm 1.96 \times$ SD of the Bland-Altman Plot; and therefore, we can confirm repeatability¹⁰⁾ (Fig. 2).

In the ICC analysis, the test-retest consistency of the total POEM-K score was 0.72, representing a strong agreement. The consistency of each question is the minimum value of 0.53 and the maximum of 0.76 (Table 3); both values were more than the moderate agreement.¹²⁾

3. Concurrent validity of POEM-K

The total score of POEM-K and individual questions showed a significantly positive linear

**Fig. 2.** Bland-Altman Plot of POEM-K.

POEM-K1: The first POEM-K of participants whose first GRC scores were from -1 to 1 and the second POEM-K of participants whose second GRC scores were from -1 to 1; POEM-K2: The second POEM-K of participants whose first GRC scores were from -1 to 1, and the third POEM-K of participants whose second GRC scores were from -1 to 1.

Table 3. Intraclass Correlation Coefficients of POEM-K and Each Item

Each item of POEM-K	Intraclass Correlation Coefficients
Itching	0.68
Sleep disturbance	0.76
Bleeding	0.53
Weeping or oozing	0.62
Crack	0.54
Flaking	0.58
Dry or rough	0.63
Total	0.72

relationship with the total score of SCORAD and the “intensity” and “pruritus” domains. The total score of POEM-K and all of the questions did not show a significant correlation with the “extent” domain of SCORAD. The total score of POEM-K and the question regarding “sleep disturbance” showed a significant correlation with the “sleep” domain of SCORAD (Table 4).

All questions of the POEM-K had a positive linear relationship with the “general health” domain of the SF-36. In particular, the question regarding “sleep disturbance” of the POEM-K had a significant relationship. The question regarding “itching” of the POEM-K and the “physical function” domain of the SF-36 significantly had a negative linear relationship (Table 5).

4. Responsiveness of POEM-K

The responsiveness of the POEM-K was evaluated by analyzing the variations in the POEM-K scores of 8 participants (5 patients with a GRC score of 4 and 3 patients with a GRC score of 5), who answered that their symptoms had improved at the second or third visit. The variation in the POEM-K scores of these participants was an average of 9.3 ± 6.2 points (p -value = 0.004). Effective size was estimated to be 1.41 (Table 6).

Discussion

Various clinical outcome measures are used in clinical studies of AD because there is no reliable laboratory test to assess and monitor the severity of the disease. Among these measures, the SCORAD, EASI, and POEM have shown adequate validity and reliability.⁹ The POEM is composed of 7 questions regarding objective symptoms (bleeding, weeping, skin cracking, and flaking) and subjective symptoms (itching, sleep loss, and skin dryness). Each question concerns the length of time that patient experienced symptoms during the past week prior to the test¹⁴. In the case of the SCORAD and EASI, expert techniques are required to apply the tests to AD patients. The POEM is more approachable: patients

Table 4. Correlation Coefficient between SCORAD and POEM-K

r	POEM-K	POEM-K	POEM-K	POEM-K	POEM-K	POEM-K	POEM-K	POEM-K
<i>p</i> -value	Total	1	2	3	4	5	6	7
SCORAD	0.54	0.33	0.33	0.39	0.32	0.34	0.33	0.39
Total	<0.001	0.020	0.021	0.005	0.022	0.015	0.020	0.006
SCORAD	0.14	0.14	0.01	0.06	-0.05	0.14	0.12	0.22
Ex	0.325	0.339	0.969	0.688	0.752	0.325	0.400	0.121
SCORAD	0.46	0.29	0.08	0.35	0.33	0.31	0.37	0.33
Int	0.001	0.040	0.591	0.013	0.019	0.027	0.008	0.021
SCORAD	0.59	0.43	0.40	0.50	0.40	0.35	0.27	0.27
Pru	<0.001	0.002	0.004	<0.001	0.005	0.012	0.055	0.055
SCORAD	0.32	-0.03	0.78	0.15	0.15	0.12	0.01	0.16
Sleep	0.026	0.819	<0.001	0.284	0.283	0.395	0.938	0.267

Ex: Extent; Int: Intensity; Pru: Pruritus; Sleep: Sleep Loss, POEM-K 1: Itching, POEM-K2: Sleep disturbance, POEM-K3: Bleeding, POEM-K4: Weeping or oozing, POEM-K5: Crack, POEM-K6: Flake off, POEM-K7: Dry or rough.

Table 5. Correlation Coefficient between SF-36 and POEM-K

r	POEM-K	POEM-K	POEM-K	POEM-K	POEM-K	POEM-K	POEM-K	POEM-K
<i>p</i> -value	Total	1	2	3	4	5	6	7
SF-36	0.22	0.07	0.36	0.18	0.06	0.12	0.01	0.15
GH	0.131	0.645	0.010	0.209	0.655	0.415	0.938	0.304
SF-36	-0.19	-0.32	-0.07	-0.14	-0.09	-0.04	-0.15	-0.04
PF	0.198	0.026	0.626	0.320	0.531	0.772	0.291	0.775
SF-36	-0.09	-0.27	0.08	-0.18	-0.08	0.03	-0.10	0.06
RP	0.522	0.059	0.575	0.205	0.574	0.833	0.503	0.676
SF-36	-0.01	-0.11	-0.06	-0.01	-0.07	0.18	0.08	-0.14
RE	0.935	0.452	0.681	0.948	0.628	0.223	0.580	0.347
SF-36	-0.13	-0.23	-0.04	-0.14	-0.04	-0.02	-0.02	-0.15
SF	0.369	0.112	0.797	0.352	0.788	0.901	0.878	0.295
SF-36	0.02	-0.15	0.05	-0.04	0.00	0.17	0.06	-0.11
BP	0.905	0.290	0.717	0.807	0.999	0.225	0.664	0.458
SF-36	-0.07	-0.06	-0.04	-0.11	-0.15	-0.02	0.01	0.08
VT	0.644	0.655	0.794	0.427	0.315	0.876	0.982	0.592
SF-36	-0.01	-0.08	-0.09	-0.18	-0.11	0.12	0.10	0.15
MH	0.944	0.577	0.555	0.222	0.451	0.408	0.473	0.294

GH: General Health; PF: Physical Function; RP: Role limitation Physical; RE: Role limitation-Emotion; SF: Social Function; BP: Body Pain; VT: Vitality; MH: Mental Health.
 POEM-K 1: Itching, POEM-K2: Sleep disturbance, POEM-K3: Bleeding, POEM-K4: Weeping or oozing, POEM-K5: Crack, POEM-K6: Flake off, POEM-K7: Dry or rough.

Table 6. Responsiveness of Effective Size of POEM-K

Variables	N	Mean	SD	<i>p</i> -value
Total POEM-K1	8	14.3	6.6	
Total POEM-K2	8	5.0	2.7	
Total POEM-K1 - Total POEM-K2	8	9.3	6.2	0.004 ^{a)}
Effective Size of POEM-K ^{b)}			1.41	

POEM-K1: Former POEM-K score of participants whose GRC scores were more than 4.

POEM-K2: Latter POEM-K score of participants whose GRC scores were more than 4

a): Paired t-test for changes of the POEM-K scores between the former and the latter visits.

b): Effective size = E(POEM-K1 - POEM-K2)/SD of POEM-K1

can be more easily engaged as they evaluate and make decisions regarding the relief of their symptoms by treatments.

For the test-retest reliability, the 32 participants with no change or a slight change of symptoms were analyzed. In the Bland-Altman Plot, score variations of 31 participants (96.9%) contributed in the mean \pm 1.96 SD were between +10.5 and -7.9 points (mean 1.3), and that of 1 participant (3.1%) was out of the range. This result confirmed an acceptable repeatability, though the range of variations was larger than that of Charman et al.'s study.⁸⁾ The ICC of the total POEM-K had a degree of strong

agreement, whereby the ICC of each question is the minimum value of 0.53, and the maximum of 0.76; all of the values are more than a certain level of moderate agreement and can prove the reliability of the POEM-K.

In the criterion-related validity analysis of the POEM-K and SCORAD, a meaningful positive linear relationship was found in all the domains except for the extent domain of the SCORAD. These results suggest that POEM-K have the validity for using as convenient tool for assessing AD. The absence of relationship between POEM-K and extent domain of SCORAD may be explained by the fact that

POEM-K did not have questions about the extent of disease. As a tool for measuring QoL in AD patients, all questions of POEM-K showed positive linear relationships with “general health”, “physical function”, and “social function” domains of SF-36. Especially “sleep disturbance” of POEM-K and “general health” domain of SF-36 had significant correlation coefficients, so did “itching” of POEM-K and “physical function” domain of SF-36. However, correlations between POEM-K and other domains of SF-36 were neither consistent nor significant. These results were similar with the correlation between SCORAD and SF-36. Only “extent” domain of SCORAD and “body pain” domain of SF-36 as well as “pruritus” domain of POEM-K and “social function” domain of SF-36 (data not shown) had a significant correlation coefficient. Therefore, other tools complementing POEM-K seem to be necessary for evaluation of QoL in AD patients.

Responsiveness of POEM-K was tested using the ES estimated from 8 participants with the GRC score of 4 or 5, because the responsiveness could be overrated if the GRC score was greater than 6. The ES of the POEM-K was 1.41. The variation of the POEM-K scores significantly increased for those patients who reported that their symptoms had improved (mean \pm SD: 9.3 ± 6.2 ; p -value: 0.004). We can verify and therefore prove the responsibility of the POEM-K.

POEM-K is a well-validated, reliable, and responsive scale. POEM-K could not sufficiently reflect the QoL of AD patients. Despite this limitation, POEM-K was concise and easily interpreted. Therefore, POEM-K could be used broadly as a tool for evaluating AD in clinical practice and research.

Acknowledgements

The authors thank Hywel C. Williams, Professor of Dermato-epidemiology, University of Nottingham,

for allowing us to use the POEM. 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: HI12C1954, HI15C0006). This work was also supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2015M3A9E3052912). The authors thank Dr. Yoon Yeo at Purdue University for critical reading of this manuscript.

References

1. Leung DY, Bieber T. Atopic dermatitis. *Lancet*. 2003;361(9352):151-60.
2. Boguniewicz M, Leung DY. 10. Atopic dermatitis. *The Journal of allergy and clinical immunology*. 2006;117(2 Suppl Mini-Primer): S475-80.
3. Williams H, Stewart A, von Mutius E, Cookson W, Anderson HR, International Study of A, et al. Is eczema really on the increase worldwide? *The Journal of allergy and clinical immunology*. 2008;121(4):947-54 e15.
4. Institute HI. Every year 1 million people were treated for severe itching due to "atopic dermatitis". 2014. Available at: URL: <http://www.nhic.or.kr/cms/board/board/Board.jsp?act=VIEW&communityKey=B0039&boardId=22846>.
5. Absolon CM, Cottrell D, Eldridge SM, Glover MT. Psychological disturbance in atopic eczema: the extent of the problem in school-aged children. *The British journal of dermatology*. 1997;137(2):241-5.
6. Schmitt J, Langan S, Williams HC. European Dermato-Epidemiology N, What are the best outcome measurements for atopic eczema? A systematic review. *The Journal of allergy and clinical immunology*. 2007;120(6):1389-98.

7. Hanifin JM, Rajka G. Diagnostic features of atopic dermatitis. *Acta Derm Venereol Suppl* (Stockh). 1980;59:44-7.
8. Charman CR, Venn AJ, Williams HC. The patient-oriented eczema measure: development and initial validation of a new tool for measuring atopic eczema severity from the patients' perspective. *Archives of dermatology*. 2004;140(12):1513-9.
9. Kocks JW, Tuinenga MG, Uil SM, van den Berg JW, Stahl E, van der Molen T. Health status measurement in COPD: the minimal clinically important difference of the clinical COPD questionnaire. *Respir Res*. 2006;7:62.
10. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-10.
11. Huisman AN, Wu MZ, Uil SM, van den Berg JW. Reliability and validity of a Dutch version of the Leicester Cough Questionnaire. *Cough*. 2007;3:3.
12. Portney LG, Watkins MO. Foundations of clinical research applications to practice. Prentice Hall Inc. 2000:5660-567.