Clinical Application of the UK Working Party’s Criteria for the Diagnosis of Atopic Dermatitis in the Chinese Population by Age Group

Li Wang1, Lin-Feng Li1; on behalf of Professional Committee of Dermatology, Chinese Association of Integrative Medicine
1Department of Dermatology, Allergy and Clinical Immunology Centre, Beijing Friendship Hospital, Capital Medical University, Beijing 100050, China

Abstract

Background: Atopic dermatitis (AD) is a common inflammatory skin disease with an increasingly significant prevalence. The prevalence of AD depends greatly on how its diagnosis is done. The UK Working Party’s diagnostic criteria for AD are simple and easy to apply without invasive laboratory tests. This study assessed the clinical utility of these criteria in China.

Methods: Data were collected from 6208 patients at 31 tertiary hospitals in 13 Chinese provinces/municipalities from March 2014 to May 2014. The agreement between the UK diagnostic criteria and the clinical records for AD was assessed by Cohen’s kappa.

Results: The overall agreement between the UK diagnostic criteria and clinical diagnosis was fair (kappa = 0.40). A slightly better agreement was found in patients aged between 4 and 9 years (kappa = 0.48), while fair agreement was found in the group <4 years and the group ≥10 years (kappa = 0.27 and 0.39, respectively). Using the UK party’s criteria as the standard, the sensitivity, specificity, positive predictive value, and negative predictive value of the clinical diagnosis of AD were 62.3%, 89.2%, 38.0%, and 95.7%, respectively.

Conclusions: Our study indicates a modest ability among Chinese dermatologists to apply the UK Working Party’s diagnostic criteria for AD, especially in patients aged <4 years and ≥10 years. Since there is no gold standard for AD diagnosis, it is important to determine how AD is identified when evaluating a diagnostic tool.

Key words: Accuracy; Atopic Dermatitis; China; Criteria; Diagnosis

INTRODUCTION

Atopic dermatitis (AD) is a common inflammatory skin disease characterized by chronic and relapsing pruritic eczematosus lesions. The condition is the result of multiple factors, including a hyperstimulated cutaneous immune system, a genetically determined compromised skin barrier, and exposure to triggering environmental stimuli.[1] Recently, there has been a significant increase in the prevalence of AD; the disease affects 15–30% of children and 2–10% of adults.[2] In 1980, Hanifin[3] established diagnostic criteria for AD, which have been validated in both hospital and community settings.[4] These criteria have been used worldwide for the last 30 years in diagnosing AD.[5] Hanifin and Rajka require the assessment of 27 separate criteria (4 major and 23 minor). This results in a complex daily practice. In 1994, the UK Working Party proposed minimal diagnostic criteria for AD; these require that an individual must have an itchy skin condition plus three or more of the following: a history of flexural involvement, a history of asthma/hay fever, a history of a generalized dry skin, the onset of rash under the age of 2 years, or visible flexural dermatitis.[6–8] These simplified criteria are easy to operate with, since they can take under 2 min/patient to ascertain and do not require invasive or laboratory tests. In addition, the sensitivity and specificity of the UK Working Party’s criteria are in good agreement with those of Hanifin and Rajka.[5]
The UK Working Party’s diagnostic criteria were validated in several independent studies from different countries, including China. Although the UK Working Party’s criteria have been used for AD diagnosis in China for many years, the ability of dermatologists in China to apply these criteria in diagnosing AD has not been evaluated. Therefore, this study evaluated the accuracy of clinical diagnoses of AD by comparing clinical diagnostic parameters to the UK diagnostic criteria.

**Methods**

**Study design and participants**

A total of 6265 outpatients were initially screened. Ultimately, 6208 cases completed the study, while 57 cases were eliminated due to a lack of clinical records. This study was designed as a cross-sectional survey using the collection of primary data on individuals diagnosed with eczema or dermatitis. This survey was carried out at 31 tertiary hospitals from 13 provinces/municipalities [Table 1]. The survey was conducted from March 2014 to May 2014. Consent was obtained from each participant.

The clinical diagnosis of AD was made by two dermatologists based on physical examinations and responses to a questionnaire completed by the dermatologist and his patients. The diagnosis was considered final when the two dermatologists were in agreement; otherwise, a final decision was made by a senior dermatologist. The questionnaire contains information on gender, age, age at appearance of lesion(s), duration of lesion(s), itching degree score, distribution and clinical features of skin lesions, history of asthma or hay fever, history of atopic disease in relatives, history of general dry skin, history of eczema in childhood, and history of flexural eczema.

**Data analysis**

The agreement between the clinical diagnosis and the UK Working Party’s diagnostic criteria for AD was evaluated by Cohen’s kappa. The interpretation of the kappa coefficient was based on Landis and Koch’s classifications: slight <0.20, fair 0.21–0.40, moderate 0.41–0.60, substantial 0.61–0.80, and almost perfect; 0.81–1.00. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the clinical diagnosis were determined and the UK Working Party’s diagnostic criteria were considered to be the gold standard.

Sensitivity was defined as the percentage of patients who met the UK criteria and were diagnosed with AD by clinicians. Specificity was defined as the percentage of patients who did not meet the UK criteria and were not diagnosed with AD. PPV was defined as the percentage of patients who did not meet the UK criteria and were diagnosed with AD whereas NPV was defined as the percentage of patients who did not meet the UK criteria among the total patients who were not diagnosed with AD. All statistical analyses were performed using SAS version 9.3 (SAS Institute Inc., Cary, NC, USA). A $P < 0.05$ was considered to be statistically significant.

**Results**

**Demographic characteristics**

Among the 6208 patients who completed the study, there were 3183 men (51.3%) and 3025 women (48.7%). The mean age was 37.8 ± 18.2 years. Overall, AD was clinically diagnosed in 975 cases. Among the AD patients, 370/975 (37.9%) met the UK criteria as well while 605 did not. On the other hand, among the 5233 patients who were not clinically diagnosed with AD, 224 cases met the UK criteria. These 224 “misdiagnosed” patients had the following skin diseases: unclassified eczema ($n = 132$), irritant contact dermatitis ($n = 11$), allergic contact dermatitis ($n = 42$), photosensitivity dermatitis ($n = 4$), autosensitization dermatitis ($n = 13$), seborrhoeic dermatitis ($n = 24$), lichen simplex chronicus ($n = 20$), prurigo ($n = 2$),...
corticosteroid-dependent dermatitis ($n = 2$), unclassified eczema and irritant contact dermatitis ($n = 22$), and unclassified eczema, allergic contact dermatitis, and seborrhoeic dermatitis ($n = 2$).

Since the prevalence of AD varies among different ages and the UK party’s diagnostic criteria for AD differ with age as well,[26] we categorized our study participants into three groups based on age: <4 years old, 4–9 years old, and ≥10 years old. The prevalence of AD and the proportion of patients who met the UK criteria were determined among the three categories [Table 2]. AD was clinically diagnosed for 36.3%, 35.3%, and 13.9% of cases aged <4 years, 4–9 years, and ≥10 years, respectively, while the UK criteria were identified for 19.6%, 22.8%, and 8.5% of cases aged <4 years, 4–9 years, and ≥10 years, respectively.

**Agreement between the clinical diagnosis of atopic dermatitis and the UK Working Party’s diagnostic criteria**

The agreement between the clinical diagnosis of AD and the UK Working Party’s diagnostic criteria was determined by computing Cohen’s kappa. Using this, the agreement scored “fair” (kappa = 0.40) among the entire study population (6208). Among different age categories, the agreement between the clinical diagnosis of AD and UK party’s diagnostic criteria was scored “fair” for cases aged <4 years and ≥10 years (kappa = 0.27 and kappa = 0.39, respectively), while agreement was scored “moderate” (kappa = 0.48) for cases aged 4–9 years [Table 3].

**Accuracy of the clinical diagnosis**

The accuracy of the clinical diagnosis was evaluated by determining the sensitivity, specificity, PPV, and NPV by comparison to the UK diagnostic criteria of AD in different age groups [Table 4].

The sensitivity and specificity of clinical diagnosis of AD were 62.3% and 89.2%, respectively, with a low PPV (38.0%) and high NPV (95.7%). When the sensitivity was stratified by age category, it ranged from 60.1% to 79.3%, while the specificity ranged from 70.8% to 90.4%.

**Discussion**

The UK Working Party’s AD diagnostic criteria were created in 1994. This diagnostic tool relies on clinical features and medical history. The UK criteria are easy to apply, especially for children, which make them appropriate for large population-based surveys and hospital studies.

The accuracy of the clinical diagnosis of AD compared to the UK diagnostic criteria has not been well studied in China. Therefore, in this study, we investigated this issue among patients with different ages. The results of our study

---

### Table 2: Prevalence and clinical features of AD by the clinical diagnoses and the UK Working Party’s diagnostic criteria

| Variables                              | <4 years | 4–9 years | ≥10 years |
|----------------------------------------|----------|-----------|-----------|
|                                        | Total    | UK criteria | Clinical diagnosis | Total    | UK criteria | Clinical diagnosis | Total    | UK criteria | Clinical diagnosis |
| Total, n (%)                           | 281 (100)| 55 (19.6)  | 102 (36.3)   | 232 (100)| 53 (22.8)  | 82 (35.3)          | 5695 (100)| 486 (8.5)   | 791 (13.9)          |
| Gender, n (%)                          |          |            |             |          |            |                   |          |             |                   |
| Male                                   | 127 (45.2)| 33 (26.0)  | 50 (39.4)    | 111 (47.8)| 30 (26.3)  | 41 (32.3)          | 2945 (51.7)| 208 (7.1)   | 379 (12.9)          |
| Female                                 | 154 (54.8)| 22 (14.3)  | 52 (33.8)    | 121 (52.2)| 23 (14.9)  | 41 (26.6)          | 2750 (48.3)| 278 (10.1)  | 412 (15.0)          |
| Average age (years), mean ± SD         | 1.8 ± 0.9| 1.9 ± 0.9  | 1.8 ± 0.9    | 7.0 ± 1.9| 6.9 ± 2.2  | 6.7 ± 2.0          | 40.8 ± 15.6| 36.0 ± 13.7 | 36.9 ± 14.7          |
| Age of onset (years), mean ± SD        | 1.3 ± 0.7| 1.1 ± 0.7  | 1.0 ± 0.7    | 5.1 ± 3.6| 2.8 ± 2.8  | 3.9 ± 2.7          | 37.7 ± 15.7| 30.7 ± 14.5 | 32.9 ± 15.3          |
| Duration (years), mean ± SD            | 0.6 ± 0.7| 0.9 ± 0.9  | 0.7 ± 0.7    | 2.0 ± 2.1| 3.8 ± 2.7  | 2.9 ± 2.5          | 3.2 ± 5.0  | 5.5 ± 6.8   | 4.2 ± 4.4           |
| Atopic skin condition, n (%)           | 275 (97.9)| 55 (19.6)  | 101 (35.9)   | 213 (91.8)| 53 (22.8)  | 79 (34.1)          | 5590 (98.2)| 486 (8.5)   | 788 (13.8)          |
| History of involvement of the skin creases, n (%) | 36 (12.8)| 27 (9.6)   | 28 (10.0)    | 43 (18.5)| 39 (16.8)  | 39 (16.8)          | 528 (9.3)  | 369 (6.5)   | 296 (5.2)           |
| History of involvement of the cheeks, n (%) | 70 (24.9)| 49 (17.4)  | 21 (7.5)     | 106 (45.7)| 29 (12.5)  | 44 (19.0)          | N/A       | N/A         | N/A                 |
| Personal history of asthma or hay fever, n (%) | 50 (17.8)| 27 (9.6)   | 22 (7.8)     | 50 (21.6)| 36 (15.5)  | 38 (16.4)          | 802 (14.1)| 225 (4.0)   | 154 (2.7)           |
| History of atopic disease in the first degree relative, n (%) | 14 (5.0) | 8 (2.8)    | 8 (2.8)      | N/A       | N/A         | N/A                | N/A       | N/A         | N/A                 |
| History of general dry skin in the last year, n (%) | 45 (16.0)| 21 (7.5)   | 18 (6.4)     | 51 (22.0)| 26 (11.2)  | 33 (14.2)          | 1461 (25.7)| 192 (3.4)   | 271 (4.8)           |
| Visible flexural eczema, n (%)         | 260 (92.5)| 54 (19.2)  | 97 (34.5)    | 201 (86.6)| 53 (22.8)  | 81 (34.9)          | 4691 (82.4)| 470 (8.3)   | 750 (13.2)          |
| Eczema involving the cheeks/forehead and outer limbs, n (%) | 184 (65.5)| 47 (16.7)  | 81 (28.8)    | N/A       | N/A         | N/A                | N/A       | N/A         | N/A                 |
| Onset under the age of 2 years, n (%)  | N/A      | N/A        | N/A          | 32 (13.8)| 24 (10.3)  | 21 (9.1)           | 26 (0.5)  | 14 (0.2)    | 15 (0.3)            |

SD: Standard deviation; N/A: Not available; AD: Atopic dermatitis.
showed that agreement between the clinical diagnosis and the UK diagnostic criteria varied with the patient's age. Against our expectations, the overall agreement between the clinical diagnosis and the UK diagnostic criteria was only "fair," which may indicate a modest ability among Chinese dermatologists to apply the UK criteria practically for patients with skin lesions, especially for those aged <4 years or ≥10 years. The overall sensitivity of the clinical diagnosis compared to the UK criteria was 62%, while the overall specificity of the clinical diagnosis was 89%; PPV ranged from 35.3% to 51.2%, and NPV was significantly higher, 95.7%. These values are in relatively good agreement with what has been reported about the validation of UK diagnostic criteria in Japanese children and adults since the sensitivity, specificity, and PPV were about 70%, 90%, and 45%, respectively, for a similar study conducted in Japan. Interestingly, the accuracy of the UK diagnostic criteria for AD in non-Asian populations was comparable to that observed in the current study and the Japanese study; the sensitivity and specificity among a South African population were 44% and 98%, respectively. The sensitivity of different criteria for the diagnosis of AD among adults/adolescents is different from that among infants or children. Notably, the sensitivity of UK criteria is relatively low for diagnosis of AD among Chinese adults/adolescents. Furthermore, in some patients, AD is not present during childhood while it starts later in life (late-onset AD). This might partly contribute to the low concordance rate between the clinical diagnoses and the UK criteria in patients aged ≥10 years in present study. In a meta-analysis including 19 validation studies, the sensitivity of the UK diagnostic criteria had a wide range, from 10% to 100%. This, on the other hand, indicates that the accuracy of the UK criteria is variable. In the same context, it is worth noting that there is no gold standard for a definite diagnosis of AD. A study by Schram et al. showed that refined Millennium criteria (typical morphology, early age of onset, Dennie-Morgan fold, and historical and actual flexural involvement) had superior sensitivity and specificity to that of the UK Working Party criteria and the Hanifin and Rajka criteria for the diagnosis of AD. The variability in practical approaches used in AD diagnosis might partially explain the only fair agreement between the UK diagnostic criteria and physical diagnosis seen in the current study.

Notably, more than one-third of patients who met the UK criteria (224/594) were misdiagnosed either primarily with unclassified eczema or secondarily with allergic contact dermatitis. This may be attributable to the inability of dermatologists to clinically differentiate between AD, eczema, and other dermatitis.

All hospitals involved in this study are tertiary hospitals located in urban areas; usually, the level of expertise for dermatologists at these hospitals is higher than that demanded of dermatologists in secondary- or primary-care settings. Thus, it is surprising that the current survey shows only "fair" agreement between the clinical diagnosis and the UK diagnostic criteria of AD. This shows an opportunity for the further improvement of the clinical training of physicians at tertiary hospitals and more urgently in secondary- or primary-care settings in rural regions of China.

**Limitations**

The main limitation of this study is its relatively small sample size when considered in the context of the total 1.3 billion inhabitants of China. This factor may have led to a selection bias. In addition, this is a hospital-based survey; thus, it cannot avoid the presence of some selection bias. Patients were recruited from multiple centers, but all centers were

**Table 3: Concordance between the clinical diagnoses and the UK Working Party's diagnostic criteria for AD**

| Age | UK criteria | |
|-----|-------------|--|
|     | Positive    | Negative | Total |
| age ≤9 years | 36 | 66 | 102 | 0.27 |
| Clinical diagnosis positive | 19 | 160 | 179 |
| Total | 55 | 226 | 281 |
| 4≤ age <9 years | 42 | 40 | 82 | 0.48 |
| Clinical diagnosis positive | 11 | 139 | 150 |
| Total | 53 | 179 | 232 |
| 10≤ age ≥10 years | 292 | 499 | 791 | 0.39 |
| Clinical diagnosis positive | 194 | 4710 | 5040 |
| Total | 486 | 5209 | 5695 |
| Total | 370 | 605 | 975 | 0.40 |
| Clinical diagnosis negative | 224 | 5009 | 5233 |
| Total | 594 | 5614 | 6208 |

KV: Kappa value.

**Table 4: Information on sensitivity, specificity, and positive and negative predictive values of AD in different age groups**

| Age (years) | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) |
|-------------|-----------------|-----------------|---------|---------|
| Value       | 95% CI          | Value           | 95% CI  | Value   | 95% CI  |
| <4          | 65.5 51.4–77.8  | 70.8 64.4–76.6  | 35.3 26.1–45.4  | 89.4 83.9–93.5  |
| 4–9         | 79.3 65.9–89.2  | 77.7 70.8–83.5  | 51.2 39.9–62.4  | 92.7 87.3–96.3  |
| ≥10         | 60.1 55.6–64.5  | 90.4 89.6–91.2  | 36.9 33.5–40.4  | 96.0 95.5–96.6  |
| Total       | 62.3 58.3–66.2  | 89.2 88.4–90.0  | 38.0 34.9–41.1  | 95.7 95.1–96.3  |

PPV: Positive predictive value; NPV: Negative predictive value; CI: Confidence interval; AD: Atopic dermatitis.
located in provincial capitals or central cities; none were located in the countryside. Therefore, it is impossible to have avoided some selection bias due to the nonhomogeneous population and the spatial distribution of patients.

In conclusion, our study shows a fair agreement between the clinical diagnoses of AD and the UK criteria. The dermatologists’ ability to recognize AD in patients aged 4–9 years was better than in those aged <4 years or ≥10 years. It is critical to know how the diagnosis was made when evaluating the epidemiological data of AD. The current results need to be further validated in a larger and broader study including more participants of different ages recruited from different medical institutions located in different regions of China.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

REFERENCES
1. Chalmers DA, Todd G, Saxe N, Milne JT, Tolosana S, Ngcelwane PN, et al. Validation of the U.K. Working Party diagnostic criteria for atopic eczema in a Xhosa-speaking African population. Br J Dermatol 2007;156:111-6. doi: 10.1111/j.1365-2133.2006.07606.x.
2. Samochcki Z, Dejewska J. A comparison of criteria for diagnosis of atopic dermatitis in children. World J Pediatr 2012;8:355-8. doi: 10.1007/s12519-012-0381-1.
3. Hanifin J. Diagnostic features of atopic dermatitis. Acta Derm Venereol (Stockh) 1980;92:44-7.
4. Brenninkmeijer EE, Schram ME, Leeflang MM, Bos JD, Spuls PI. Diagnostic criteria for atopic dermatitis: A systematic review. Br J Dermatol 2008;158:754-65. doi: 10.1111/j.1365-2133.2007.08412.x.
5. Gu H, Chen XS, Chen K, Yan Y, Jing H, Chen XQ, et al. Evaluation of diagnostic criteria for atopic dermatitis: Validity of the criteria of Williams et al. in a hospital-based setting. Br J Dermatol 2001;145:428-33. doi:10.1111/j.1365-2133.2001.04379.x.
6. Williams HC, Burney PG, Hay RJ, Archer CB, Shipley MJ, Hunter JJ, et al. The U.K. Working Party’s Diagnostic Criteria for Atopic Dermatitis. I. Derivation of a minimum set of discriminators for atopic dermatitis. Br J Dermatol 1994;131:383-96.
7. Williams HC, Burney PG, Strachan D, Hay RJ. The U.K. Working Party’s Diagnostic Criteria for Atopic Dermatitis. II. Observer variation of clinical diagnosis and signs of atopic dermatitis. Br J Dermatol 1994;131:397-405.
8. Williams HC, Burney PG, Pembroke AC, Hay RJ. The U.K. Working Party’s Diagnostic Criteria for Atopic Dermatitis. III. Independent hospital validation. Br J Dermatol 1994;131:406-16.
9. de Menezes TN, Oliveira EC, de Sousa Fischer MA. Validity and concordance between self-reported and clinical diagnosis of hypertension among elderly residents in northeastern Brazil. Am J Hypertens 2014;27:215-21. doi: 10.1093/ajh/hpt181.
10. Wu CS, Lai MS, Gau SS, Wang SC, Tsai HJ. Concordance between patient self-reports and claims data on clinical diagnoses, medication use, and health system utilization in Taiwan. PLoS One 2014;9:e112257. doi: 10.1371/journal.pone.0112257.
11. Sacki H, Oiso N, Homma M, Iizuka H, Kawada A, Tamaki K. Prevalence of atopic dermatitis in Japanese adults and community validation of the U.K. diagnostic criteria. J Dermatol Sci 2009;55:140-1. doi: 10.1016/j.jdermsci.2009.04.004.
12. Sacki H, Iizuka H, Mori Y, Akasaka T, Takagi H, Kitajima Y, et al. Community validation of the U.K. diagnostic criteria for atopic dermatitis in Japanese elementary schoolchildren. J Dermatol Sci 2007;47:227-31. doi: 10.1016/j.jdermsci.2007.04.006.
13. Liu P, Zhao Y, Mu ZL, Lu QJ, Zhang L, Yao X, et al. Clinical features of adult/adolescent atopic dermatitis and Chinese criteria for atopic dermatitis. Chin Med J 2016;129:757-62. doi: 10.4103/0366-6999.178960.
14. Schram ME, Leeflang MM, DEN Ottolander JP, Spuls PI, Bos JD. Validation and refinement of the Millennium Criteria for atopic dermatitis. J Dermatol 2011;38:850-8. doi: 10.1111/j.1346-8138.2011.01202.x.