Validation of the Asthma Predictive Index and comparison with simpler clinical prediction rules

Nora A. Leonardi, MSc,[a] Ben D. Spycher, PhD,[a] Marie-Pierre F. Strippoli, MSc,[a] Urs Frey, MD, PhD,[b] Michael Silverman, MD, FRCPCH,[c] and Claudia E. Kuehni, MD, MSc[a] *Bern and Basel, Switzerland, and Leicester, United Kingdom*

Background: The loose and stringent Asthma Predictive Indices (API), developed in Tucson, are popular rules to predict asthma in preschool children. To be clinically useful, they require validation in different settings.

Objective: To assess the predictive performance of the API in an independent population and compare it with simpler rules based only on preschool wheeze.

Methods: We studied 1954 children of the population-based Leicester Respiratory Cohort, followed up from age 1 to 10 years. The API and frequency of wheeze were assessed at age 3 years, and we determined their association with asthma at ages 7 and 10 years by using logistic regression. We computed test characteristics and measures of predictive performance to validate the API and compare it with simpler rules.

Results: The ability of the API to predict asthma in Leicester was comparable to Tucson: for the loose API, odds ratios for asthma at age 7 years were 5.2 in Leicester (5.5 in Tucson), and positive predictive values were 26% (26%). For the stringent API, these values were 8.2 (9.8) and 40% (48%). For the simpler rule early wheeze, corresponding values were 5.4 and 21%; for early frequent wheeze, 6.7 and 36%. The discriminative ability of all prediction rules was moderate (*c* statistic ≤ 0.7) and overall predictive performance low (scaled Brier score < 20%).

Conclusion: Predictive performance of the API in Leicester, although comparable to the original study, was modest and similar to prediction based only on preschool wheeze. This highlights the need for better prediction rules. (J Allergy Clin Immunol 2011;127:1466-72.)

Key words: *Asthma, wheeze, children, prognosis, prediction, longitudinal, cohort study, validation*

Abbreviations used
API: Asthma Predictive Index
EFW: Early frequent wheeze
EW: Early wheeze
LRC: Leicester 1998b Respiratory Cohort
NPV: Negative predictive value
OR: Odds ratio
PPV: Positive predictive value
TCRS: Tucson Children's Respiratory Study

The Asthma Predictive Index (API), developed in the Tucson Children's Respiratory Study (TCRS), is a clinical tool to predict the risk of asthma at school age in preschool children.[1] The authors defined 2 prediction rules: the loose API for a medium and the stringent API for a high risk of later asthma. The API is a popular clinical prediction rule and is commonly considered the standard to which new rules are compared.[2-5] It has been used for various purposes, such as recruiting high-risk children for clinical trials[6,7] or testing whether physiological measurements distinguish between medium and high-risk groups.[8,9] It has also been widely advertised for use in clinical management of preschool wheeze.[10-14]

But how well does the API predict development of asthma in different settings? Testing the accuracy of prediction rules in different populations (external validation) is crucial for application in clinical practice.[15-19] Two previous studies have compared the performance of the API with that of novel risk scores, but without comparing its performance with the original study and discussing its applicability in the new setting.[4,5] Besides its performance, another important criterion for acceptance of prediction rules in clinical practice is ease of implementation.[15,18,20] The API is a rather complex prediction rule based on frequency of early wheeze (none, infrequent, frequent) combined with information on family and personal history (parental asthma, eczema, allergic rhinitis, wheeze without colds) and eosinophil counts. Clinical patterns of early wheeze are known to be strong predictors of later respiratory disease on their own.[2-5,21,22] This raises the question of how much prognostic information is gained by the additional complexity of the API.

The aims of this study were, first, to validate the API in an independent, longitudinal, population-based cohort and, second, to compare its performance with simpler prediction rules based only on the frequency of preschool wheezing.

METHODS

Study populations

Development cohort. The TCRS, a birth cohort study in Arizona, contacted 1596 healthy infants between 1980 and 1984.[23] Baseline clinical information for the API was assessed from questionnaires at enrollment and at ages 2 and 3 years and from a blood sample at age 1 year; the outcome asthma was assessed at ages 6, 8, 11, and 13 years (Table I).[1,24]

From [a]the Institute of Social and Preventive Medicine, University of Bern; [b]the University Children's Hospital Basel, University of Basel; and [c]the Division of Child Health, Department of Infection, Immunity and Inflammation, University of Leicester.

Supported by the Swiss National Science Foundation (PDFMP3-123162, 823B-046481, 3200B0-122341) and Asthma UK (07/048).

Disclosure of potential conflict of interest: The authors have declared that they have no conflict of interest.

Received for publication October 11, 2010; revised February 25, 2011; accepted for publication March 1, 2011.

Available online March 31, 2011.

Reprint requests: Claudia E. Kuehni, MD, MSc, Institute of Social and Preventive Medicine, University of Bern, Finkenhubelweg 11, CH-3012 Bern, Switzerland. E-mail: kuehni@ispm.unibe.ch.

0091-6749/$36.00

doi:10.1016/j.jaci.2011.03.001

TABLE I. Characteristics of the development and validation cohorts

Characteristics	Development cohort: TCRS*		Validation cohort: LRC†	
Location	Tucson, Ariz		Leicester, UK	
Climate	Desert		Temperate maritime	
Recruitment year	1980-1984		1998	
Study design	Prospective cohort		Prospective cohort	
Recruitment	Healthy infants seeing pediatricians of a large HMO in Tucson		Random sample from the general population	
Age at recruitment (y)	0		1	
Mother's ethnicity (%)	White (80)		White (81)	
	Mexican American (20)		South Asian (19)‡	
Sex (% males)	49		52	
Children contacted (N)	1596		4300	
Questionnaire survey§	0 y	1246 (100%)		
	2 y (1.6 ± 0.4)	1055 (85%)	2 y (1.5 ± 0.3)	3392 (100%)
	3 y (2.9 ± 0.5)	940 (75%)	3 y (2.5 ± 0.3)	2405 (71%)
	6 y (6.3 ± 0.9)	1025 (82%)	7 y (6.5 ± 0.3)	2092 (62%)
	11 y (10.9 ± 0.6)	955 (77%)	10 y (9.8 ± 0.3)	1521 (45%)
Blood sample	1 y (0.9 ± 0.1)	912 (73%)		

HMO, Health maintenance organization; *UK*, United Kingdom.
*Data as reported by Castro-Rodriguez et al,[1] Taussig et al.[23,24]
†Data as reported by Kuehni et al.[25]
‡Indian, Pakistani, or Bangladeshi origin.
§Includes only surveys used for validation of the API in LRC. Data are mean age at survey (mean years ± SD), number responded (response rate calculated on the basis of children who replied to first survey).

Validation cohort. The Leicester 1998b Respiratory Cohort (LRC, www.leicestercohorts.org) is a population-based stratified random sample of children recruited in 1998 at the age of 1 year in Leicestershire, United Kingdom, that includes 3500 white and 800 South Asian children.[25] Parents received standardized questionnaires at the child's age of 2, 3, 5, 7, and 10 years (Table I), with questions on respiratory symptoms, diagnoses and treatments, environmental exposures, parental history of atopic diseases, ethnicity, and socioeconomic situation.[26,27] The Leicestershire Health Authority Research Ethics Committee approved the study.

Outcome and prediction rules

In both cohorts, clinical information assessed at the ages of 2 and 3 years was used to define prediction rules for active asthma at school age (6 and 11 years in TCRS; 7 and 10 years in LRC). Table II shows definitions of predictor and outcome variables and prediction rules in both cohorts. In TCRS, asthma was predicted by using the API only.[1] In LRC, we predicted asthma by using (1) the API, (2) simpler prediction rules based only on frequency of wheeze, and (3) a random rule to illustrate the predictive performance of chance alone. While the performance of random rules is predictable, they are useful to put the performance of the other rules into perspective.

Loose and stringent API. As originally proposed, the API was based on 7 clinical features (predictor variables) assessed in the first 3 years of life: early wheeze (EW), early frequent wheeze (EFW), wheeze without colds, parental asthma, personal history of eczema, allergic rhinitis, and blood eosinophils.[1] We approximated the API by using similar or comparable information from LRC. The 3 wheeze variables could be matched closely, whereas some adaptations were made for parental asthma, eczema, and allergic rhinitis; eosinophilia was replaced by a surrogate (Table II).

As in the original study, a child with EW and 1 of 2 major risk factors (parental asthma, personal eczema) or 2 of 3 minor risk factors (allergic rhinitis, wheeze without colds, eosinophilia) was assigned a medium risk of later asthma (positive loose API). A child with EFW and the same combination of risk factors was considered at high risk of later asthma (positive stringent API).

Frequency of wheeze. A child with EW was assigned a medium risk, a child with persistent EW or EFW a high risk, and a child with persistent EFW a very high risk.

Random rules. Preschool children were randomly assigned positive random 1 and random 2 rules with probabilities equal to the prevalence of the loose and stringent indices, respectively.

Statistical analyses

As in the original study, only children with complete information on all predictor variables or with sufficient information to confirm a positive loose or stringent API were included in the analysis.[1]

First, we compared the prevalence of asthma, the predictor variables, and prediction rules between the 2 cohorts. Second, we compared the performance of the API in the development (TCRS) and validation (LRC) cohorts. Third, we compared the performance of the API with that of simpler and random prediction rules. This comparison was restricted to data from LRC only.

We used a variety of measures to compare predictive performance across cohorts and prediction rules, including the standard test characteristics odds ratio (OR), sensitivity, specificity, and positive (PPV) and negative predictive values (NPV). Sensitivity measures the fraction of children with later asthma who were identified as at risk at preschool age; specificity measures the fraction of children without later asthma who were not identified as at risk. Clinically more relevant are PPV, the probability of later asthma in a child identified as at risk, and NPV, the probability of no later asthma in a child not identified as at risk.

We also assessed measures of discrimination, calibration, and overall predictive performance of the API. *Discrimination* is the ability of a rule to correctly distinguish between children with and without asthma. We assessed discrimination by using the concordance (c) statistic. Discrimination is not better than chance if $c = 0.5$, moderate if $c > 0.6$, good if $c > 0.8$, and perfect if $c = 1$.[16] *Calibration* is the extent of agreement between predicted risk and the frequency of observed outcomes; for example, if a 20% probability of asthma is predicted, the observed frequency of asthma will be close to 20% for a well calibrated rule. As a measure of calibration, we compared PPV (the predicted probability of asthma for those with a positive API) and 1 – NPV (the predicted probability for those with a negative API) between the 2 cohorts. As a summary measure of *overall predictive performance*, combining discrimination and calibration, we used the scaled Brier score, which measures the average difference between predicted and actual outcomes.[17,28] The scaled Brier score ($\text{Brier}_{\text{scaled}} = 1 - \text{Brier}/\text{Brier}_{\text{max}}$) ranges from 0% for a noninformative model to 100% for a perfect prediction of outcomes. Overall performance could be assessed in LRC only.