

A. Dierkes-Globisch<sup>1</sup>  
A. Globisch<sup>2</sup>  
M. Schmitz<sup>3</sup>  
C. Schirpke<sup>4</sup>  
J. Schlegel<sup>5</sup>  
D. Koehler<sup>6</sup>  
K. Strassburger<sup>7</sup>  
R. Merget<sup>8</sup>

# Validität inhalativer Provokationstests zur Bestimmung der unspezifischen bronchialen Reaktivität – Vergleich verschiedener Methoden

*Validity of Nonspecific Bronchial Provocation Tests for the Diagnosis of Airway Hyperresponsiveness – A Comparison of Different Methods*

## Zusammenfassung

**Zielsetzung:** Ziel der Studie war, die Wertigkeit unterschiedlicher Methoden verschiedener Zentren für die Messung der bronchialen Hyperreaktivität zu vergleichen. **Methodik:** In die Studie eingeschlossen wurden 648 Patienten, die sich in sechs pneumologischen Zentren mit Kurzatmigkeit in der Anamnese vorstellten. Keiner nahm Medikamente ein, die die Testergebnisse hätte beeinflussen können bzw. hatte in den letzten zwei Wochen einen viralen Infekt durchgemacht. Alle Studienteilnehmer beantworteten einen Fragebogen nach früheren Symptomen und führten einen unspezifischen bronchialen Provokationstest entsprechend des jeweiligen Zentrumsprotokolls durch. Die Auswertung erfolgte durch ROC(receiver operating characteristic)-Kurven mit den Fragebogenantworten als Goldstandard. Verglichen wurden die diagnostischen Testsensitivitäten am Cut-off-Punkt der bronchialen Hyperreaktivität des jeweiligen Zentrums. **Ergebnisse:** Die ROC-Kurven zeigten eine schlechte Validität aller Tests, d. h., bei keinem der Tests wurden gleichzeitig eine akzeptable Sensitivität und Spezifität erreicht. Zwischen den einzelnen Zentren bestanden keine wesentlichen Unterschiede in den ROC-Kurvenverläufen. Allerdings unterschied sich die maximale Sensitivität am Cut-off-Punkt signifikant für die Frage „Pfeifen im Brustkorb in den letzten zwölf Monaten“ in den verschiedenen Zentren (0,35–0,70). Die Auswahl der einzelnen Frage hatte ansonsten keinen wesentlichen Einfluss auf die Testvalidität. **Schlussfolgerung:** Obwohl einige der Unterschiede zwischen den einzelnen Zentren möglicherweise durch Charakteristika der verschiedenen Kollektive bedingt sind, sind die großen Differenzen in den Testsensitivitäten inakzeptabel und unterstreichen die

## Abstract

**Objective:** To compare the validity of different methods for the assessment of bronchial hyperresponsiveness used by different centers. **Methods:** Case series of 648 subjects referred to six pulmonary centers, all with a history of shortness of breath without airway obstruction, without use of medication that might influence the tests and without viral infections during the previous two weeks. All subjects answered a questionnaire of recent symptoms and underwent bronchial challenge with a chemical stimulus according to each center's protocol. Analysis was performed by receiver operating characteristic (ROC) plots using the questionnaire's answers as the gold standard. Diagnostic test sensitivities at the cut-offs for bronchial hyperresponsiveness indicated by each center were compared. **Results:** ROC plots showed poor validity of all tests, i.e. both acceptable sensitivity and specificity were not observed with any test. There was no obvious difference of the slope of the ROC plots between the different centers. However, maximal sensitivity differed considerably: for "wheeze during the previous 12 months", sensitivity at each center's cut-off for the definition of bronchial hyperresponsiveness varied between 0.35 and 0.73. The choice of the question used as the standard had little influence on test validity. **Conclusion:** Although some of the differences between centers may be explained by subject characteristics, the large differences of the test sensitivities are unacceptable and underscore the need for standardization of these tests, primarily with respect to sufficient sensitivity.

## Institutsangaben

Bundeswehrzentral Krankenhaus Koblenz, Abteilung für Innere Medizin  
Bundeswehrzentral Krankenhaus Koblenz, Abteilung für Dermatologie  
Hochgebirgsklinik Davos, Schweiz  
Lungenklinik Lostau  
Universitätsklinik Mainz, Abteilung für Pneumologie  
Fachklinik Kloster Grafschaft, Schmallebenberg  
BASF AG, Ludwigshafen  
Berufsgenossenschaftliches Forschungsinstitut für Arbeitsmedizin (BGFA), Bochum

## Korrespondenzadresse

Prof. Dr. med. R. Merget · BGFA · Bürkle-de-la-Camp-Platz 1 · 44789 Bochum · E-mail: merget@bgfa.de

## Bibliografie

Pneumologie 2002; 56: 230–234 © Georg Thieme Verlag Stuttgart · New York · ISSN 0934-8387

Notwendigkeit einer Standardisierung für diese Testverfahren, und zwar primär hinsichtlich einer ausreichenden Sensitivität.

## Einleitung

Für die Diagnose von Erkrankungen mit variabler Atemwegsobstruktion werden unspezifische bronchiale Provokationstests eingesetzt [1]. Es besteht Übereinkunft, dass diese Tests für die Diagnose eines Asthma bronchiale eine schlechte Spezifität haben, jedoch aufgrund ihres hohen negativen prädiktiven Wertes in der klinischen Praxis zum Ausschluss von Asthma bronchiale genutzt werden [2]. Bezüglich des Testverfahrens wurde ein Vielzahl verschiedener Methoden veröffentlicht, die sich hinsichtlich Stimulus, Apparatur (z.B. Vernebler), Messprotokoll und Zielgrößen unterschieden [3–5]. Um die Testergebnisse vergleichen zu können, ist eine Standardisierung der Tests notwendig, zumal die Auswirkungen der verschiedenen Testprotokolle auf die diagnostische Wertigkeit nicht bekannt sind.

Unseres Wissens gibt es keine Untersuchung, die bronchiale Provokationstests verschiedener Methoden und verschiedener Zentren vergleicht. Ziel dieser Studie war, unspezifische bronchiale Provokationstests verschiedener Zentren unter Benutzung eines Fragebogens als Goldstandard zu vergleichen.

## Methodik

### Fragebogen

Alle Studienteilnehmer füllten nach Durchführung des unspezifischen bronchialen Provokationstests einen Fragebogen aus, der von den Mitgliedern der Arbeitsgruppe zuvor entwickelt worden war:

(1) Haben Sie in den letzten zwölf Monaten ein pfeifendes oder keuchendes Geräusch in Ihrem Brustkorb gehört? (2) Hat bei Ihnen der Arzt jemals die Diagnose Bronchialasthma gestellt? (3) Hatten Sie in den letzten zwölf Monaten nächtlichen Husten? (4) Hatten Sie in den letzten zwölf Monaten nächtliche Atemnot? (5) Hatten Sie in den letzten zwölf Monaten Atemnot oder Pfeifen im Brustkorb bei Nebel oder Kälte? (6) Hatten Sie bei dem eben durchgeführten Provokationstest Atemnot oder Pfeifen im Brustkorb? (7) Hatten Sie bei dem eben durchgeführten Provokationstest Husten? (8) Hatten Sie in den letzten zwölf Monaten ähnliche Symptome wie bei dem Provokationstest?

### Kollektiv

In die Studie eingeschlossen wurden 648 Patienten mit Atemnot aus sechs Zentren: n = 152 Universitätsklinik mit Poliklinik (Zentrum 1), n = 136 pneumologische Rehabilitationsklinik mit Poliklinik (Zentrum 2), n = 47 Haus der pneumologischen Maximalversorgung mit Poliklinik (Zentrum 3), n = 209 Haus der pneumologischen Maximalversorgung und Rehabilitationsklinik mit Poliklinik (Zentrum 4), n = 52 Universitätsklinik mit Poliklinik (Zentrum 5) und n = 52 Medizinischer Dienst einer chemischen Fabrik (Zentrum 6).

Ausschlusskriterien für die vorliegende Studie waren eine forcierte 1-Sekunden-Kapazität ( $FEV_1$ ) < 80% oder ein spezifischer Atemwegswiderstand ( $sRaw$ ) > 1,5 kPa · s. Keiner der Patienten nahm Medikamente, die die Testergebnisse hätten beeinflussen können bzw. berichtete über akute virale Infektionen in den letzten zwei Wochen.

### Bronchiale Provokationstests

Bronchiale Hyperreaktivität wurde in einem inhalativen Provokationstest mit einem unspezifischen Stimulus untersucht. Dabei benutzte jedes Zentrum einen anderen bronchialen Provokationstest (Tab. 1).

Tab. 1 Methoden der unspezifischen bronchialen Provokationstests der verschiedenen Zentren

Zentrum	Patienten (n)	Stimulus	(Kumulative) Dosis/ Konzentration	Vernebler	Messparameter/Gerät	Hyperreaktivitätskriterium	Referenz Autor/Jahr
1	152	MCH	0,035 – 13,3 mg*	Prallhelmvernebler (Heyer; Bad Ems)	spezifische Atemwegsleitfähigkeit (sGaw)/ Masterlab, Jaeger, Würzburg	$PD_{50} \leq 1$ mg	Gonsior et al. 1976 [6]
2	136	HIS	0,03 – 16 mg × ml <sup>-1</sup> (3,8 – 4180 µg)	Reservoir (Provotest II; Pari, Starnberg)	FEV <sub>1</sub> /MicroLab, Rochester, UK	$PC_{20} \leq 2$ mg × ml <sup>-1</sup>	Schmitz-Schumann et al. 1991 [7]
3	47	ACH	0,125 – 0,5 mg × ml <sup>-1</sup>	Düsenvernebler (Erka Masterjet, Kallmeyer, Bad Tölz)	FEV <sub>1</sub> /Flowscreen, Jaeger, Würzburg	$PC_{15} \leq 0,25$ mg × ml <sup>-1</sup>	keine Referenz
4	209	HIS	2 – 500 µg	Reservoir (Provotest II; Pari, Starnberg)	Peak Flow (PEF)/ Masterlab, Jaeger, Würzburg	$PD_{20} \leq 250$ µg	Köhler 1991 [8]
5	52	MCH	0,1 – 2,19 mg*	Düsenvernebler (Piccolo, Heyer, Bad Ems)	spezifische Atemwegsleitfähigkeit/Masterlab, Jaeger, Würzburg	$PD_{50} \leq 0,8$ mg	Schlegel 1991 [9]
6	52	CAR	0,28 – 4,73 mg	Reservoir (Provotest I; Pari, Starnberg)	spezifischer Atemwegswiderstand (sRaw)/Masterlab, Jaeger, Würzburg	$PD_{100} + sRaw \leq 1,93$ mg	Klein et al. 1991 [10]

\* Dosis durch Wiegen ermittelt. MCH, Metacholin. HIS, Histamin. ACH, Acetylcholin. CAR, Carbachol.  
 $PD_{100} + sRaw$  Provokationsdosis für eine Verdoppelung des spez. Atemwegswiderstands und Anstieg auf mindestens 2 kPa · s<sup>-1</sup>

## Statistik

Als Mittelwerte sind arithmetische Mittelwerte mit Standardabweichung angegeben, mit Ausnahme der Provokationskonzentration  $PC_xY$  bzw. der Provokationsdosis  $PD_xY$ , die einen  $\geq x\%$  Abfall des Lungenfunktionsparameters  $Y$  verursachte. Bei diesen Parametern wurde eine logarithmische Transformation durchgeführt und die geometrischen Mittelwerte mit 95%-Konfidenzintervall (KI) angegeben. Minima und Maxima sind in Klammern aufgeführt. Unterschiede zwischen den Gruppen wurden mit dem  $\chi^2$ -Test für Binominaldaten bzw. mit dem Wilcoxon-Mann-Whitney-Test oder dem Wilcoxon-Signed-Rank-Test für ungepaarte bzw. gepaarte Ordinaldaten bestimmt. Alle angegebenen  $p$ -Werte sind zweiseitig. Es wurde ein Signifikanzniveau von  $p < 0,05$  angenommen.

Um die diagnostische Validität der Tests zu bestimmen, benutzen wir die diagnostische Sensitivität und ROC-Kurven [11]. Bezüglich der Sensitivität wurden die Ergebnisse jedes Zentrums über das gesamte Spektrum der Provokationskonzentration bzw. -dosis sowie an der Maximalkonzentration bzw. -dosis und am von jedem Zentrum definierten Cut-off-Punkt der bronchialen Hyperreaktivität betrachtet.

## Ergebnisse

Die Gruppen unterschieden sich nicht bezüglich Alter und Geschlecht mit Ausnahme des Zentrums 2, die mehr weibliche Patienten als Zentrum 1, 3 und 4 ( $p < 0,05$ ) hatten (Tab. 2).

Mit keinem der Tests wurde gleichzeitig eine akzeptable Sensitivität und Spezifität bei derselben Provokationskonzentration oder -dosis erreicht (Abb. 1). Interessanterweise ergaben sich keine wesentlichen Unterschiede zwischen den ROC-Kurvenverläufen der verschiedenen Zentren.

Die maximalen diagnostischen Sensitivitäten differierten jedoch beträchtlich zwischen den einzelnen Zentren (Abb. 2). Beispielsweise betrug die Sensitivität bei maximaler Konzentration bzw. Dosis bezüglich Frage 1 („Haben Sie in den letzten zwölf Mona-

ten ein pfeifendes oder keuchendes Geräusch in Ihrem Brustkorb gehört?“) 0,99 für Zentrum 1 und 0,40 für Zentrum 3 (Mittel aller Zentren 0,81). Ähnlich variierte die Sensitivität am Cut-off-Punkt der bronchialen Hyperreaktivität zwischen 0,35 und 0,73 (Mittel aller Zentren 0,60). Die Wahl der als Standard definierten Frage hatte dabei insgesamt wenig Einfluss auf die Testvalidität, obwohl sich bei Zentrum 3 die Sensitivität durch Benutzung der Frage 2 („Hat bei Ihnen der Arzt jemals die Diagnose Bronchialasthma gestellt?“) an Stelle von Frage 1 deutlich von 0,40 auf 0,83 steigern ließ (Abb. 1). Frage 6 („Hatten Sie bei dem eben durchgeführten Provokationstest Atemnot oder Pfeifen im Brustkorb?“) hatte in allen Zentren die höchste Sensitivität (Mittel für die Maximalkonzentration/-dosis 0,91, Minimum 0,81, Maximum 0,98, Mittel für Cut-off-Punkt 0,70, Minimum 0,54, Maximum 0,79). Die anderen Fragen erbrachten keinen zusätzlichen Informationsgewinn.

## Diskussion

Der Nutzen einer Standardisierung unspezifischer bronchialer Provokationstests mit pharmakologischen Substanzen haben verschiedene Fachgesellschaften, wie die American Thoracic Society [12], die European Respiratory Society [13] sowie die Deutsche Gesellschaft für Pneumologie [14] zur Herausgabe entsprechender Leitlinien veranlasst. Trotz dieser Leitlinien sind die empfohlenen Testverfahren von einer Standardisierung weit entfernt und zumindest in Deutschland nicht ausreichend umgesetzt. Obwohl schon verschiedene Tests direkt miteinander verglichen wurden, ist uns keine Studie bekannt, die die Wertigkeit verschiedener Tests in unterschiedlichen Zentren untersucht hat. Als Goldstandard haben wir einen Fragebogen gewählt, der von der Arbeitsgruppe entwickelt wurde und von allen Untersuchern benutzt wurde. Die Arbeitsgruppe hatte sich zu der Entwicklung eines neuen, speziell auf die bronchiale Hyperreaktivität abgestellten Fragebogens entschlossen. Obwohl es sich um keinen validierten Fragebogen handelte, halten wir den Fragebogen für geeignet, da die Fragen im Einzelnen wörtlich den Fragen in international anerkannten Asthmafragebogen entsprechen

Tab. 2 Anamnestiche Daten und Ergebnisse der Lungenfunktions- und inhalativen Provokationstests

Zentrum	Patienten (n)	Geschlecht (m/f)	Alter (Jahre)	Lungenfunktion	bronchiale Hyperreaktivität	Prävalenz bronchialer Hyperreaktivität (%) <sup>*</sup>
1	152	95/57	39,5 ± 14,3 (18–72)	sGaw 3,62 ± 1,72 (0,22–7,2) kPa · s <sup>-1</sup>	PD <sub>50</sub> sGaw 0,49 [0,03; 5,63] mg MCH	68,4
2	136	55/81	40,3 ± 12,8 (18–75)	FEV <sub>1</sub> 97,3 ± 12,0 (80–140) % Soll	PC <sub>20</sub> FEV <sub>1</sub> 0,58 [0,01; 24,10] mg · ml <sup>-1</sup> HIS	64,0
3	47	34/13	41,7 ± 15,6 (15–69)	FEV <sub>1</sub> 105,3 ± 18,1 (80–154) % Soll	PC <sub>15</sub> FEV <sub>1</sub> 0,16 [0,05; 0,54] mg · ml <sup>-1</sup> ACH	31,9
4	209	110/99	43,7 ± 17,8 (7–97)	FEV <sub>1</sub> 98,3 ± 14,8 (80–165) % Soll	PD <sub>20</sub> PEF 0,10 [0,01; 0,76] mg HIS	34,4
5	52	27/25	41,3 ± 14,6 (14–67)	nicht durchgeführt	PD <sub>50</sub> sGaw 0,62 [0,09; 4,16] mg MCH	36,5
6	52	36/16	37,9 ± 12,8 (19–51)	sRaw 0,76 ± 0,22 (0,36–1,48) kPa · s <sup>-1</sup>	PD <sub>100+sRaw</sub> 0,74 [0,15; 3,56] mg CAR	55,8

\* Durch den jeweiligen Cut-off-Punkt jedes Zentrums ermittelt.

PD<sub>100</sub> + sRaw Provokationsdosis für eine Verdoppelung des spez. Atemwegwiderstands und Anstieg auf mindestens 2 kPa · s<sup>-1</sup>

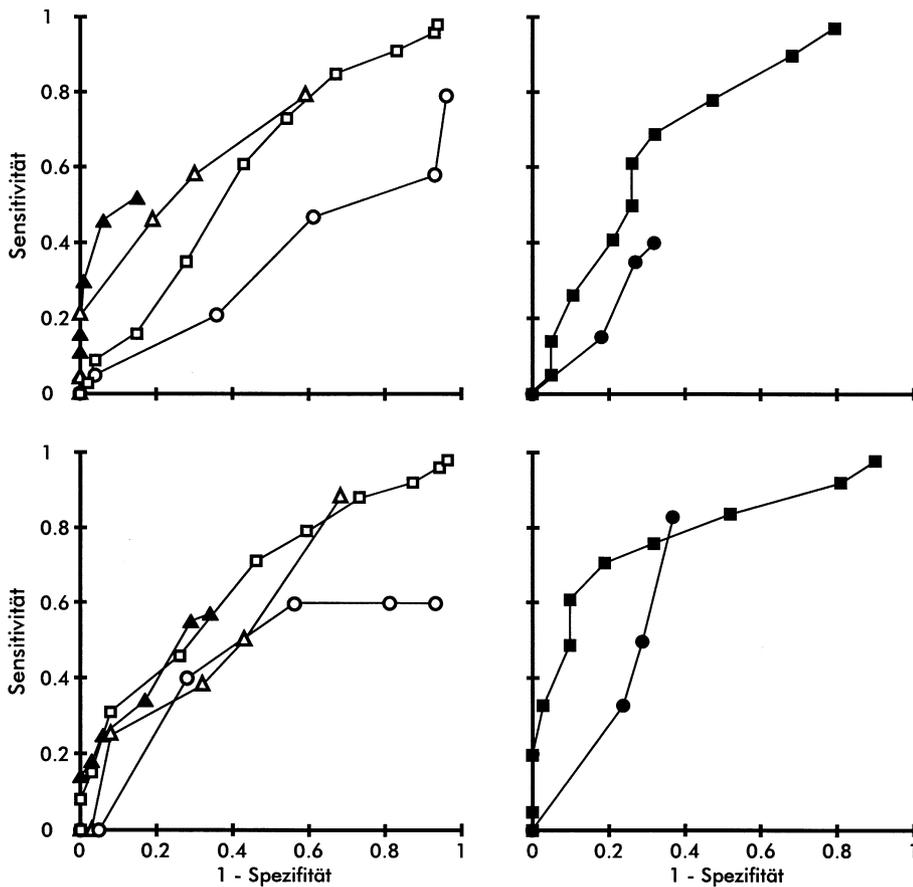


Abb. 1 ROC-Kurven für Frage 1 (Pfeifen in den letzten zwölf Monaten), oberer Teil und Frage 2 (Arzt Diagnose Asthma), unterer Teil. Die einzelnen Kurven stellen die verschiedenen Zentren dar (Zentrum 1: offenes Viereck, Zentrum 2: gefülltes Viereck, Zentrum 3: gefüllter Punkt, Zentrum 4: gefülltes Dreieck, Zentrum 5: offener Punkt, Zentrum 6: offenes Dreieck). Links sind die vier Zentren, die als Kriterium die PD benutzt haben, rechts die zwei, die die PC benutzt haben.

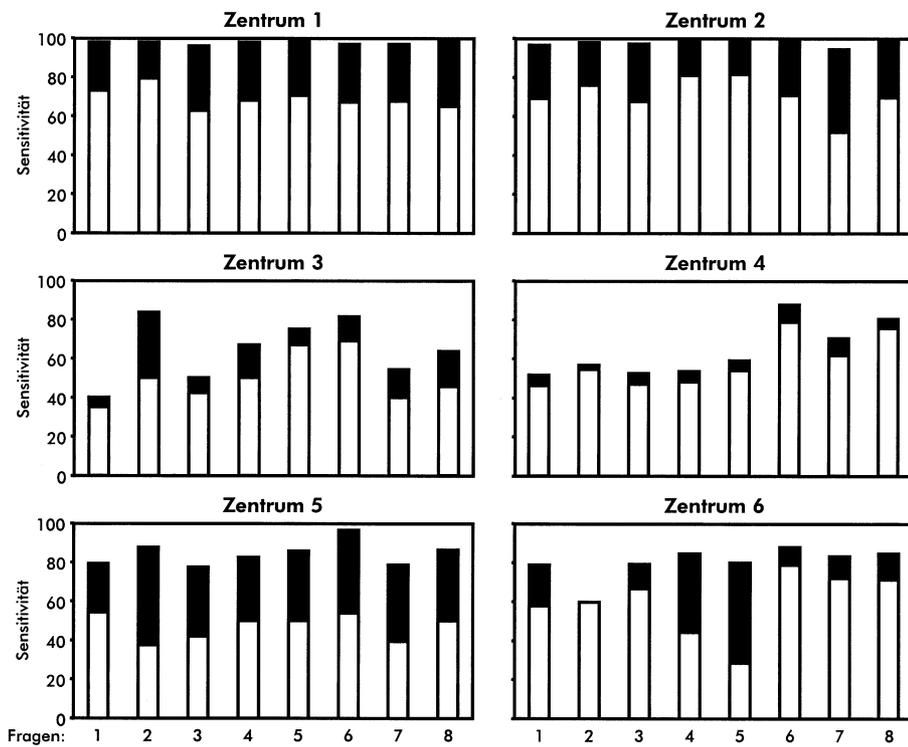


Abb. 2 Vergleich der diagnostischen Sensitivität am Cut-off-Punkt der bronchialen Hyperreaktivität (weiße Säulen) und an der maximalen Provokationskonzentration/-dosis (weiße und schwarze Säulen) für die Fragen 1 bis 8.

und zudem keine wesentlichen Unterschiede zwischen den einzelnen Fragen gefunden wurden.

Die Haupteinschränkung unserer Studie liegt in der möglichen Selektion der Patienten in den verschiedenen Zentren. Da wir

diese Selektion nicht ausschließen können, kann diese Studie nicht den besten oder schlechtesten Test bestimmen. Die vorliegende Studie sollte als Pilotstudie verstanden werden, die Ergebnisse wären ggf. an einem größeren und möglichst homogeneren (Asthma-)Kollektiv zu bestätigen.

Ziel dieser Untersuchung war, die Wertigkeit der Tests anhand der Sensitivität mit dem für jedes Zentrum verschiedenen Kriterium der bronchialen Hyperreaktivität grob abschätzen zu können. Wenn man nun eine Rangfolge der Zentren nach maximal vernebelter Dosis und Sensitivität am Cut-off-Punkt der bronchialen Hyperreaktivität erstellt, so stellt sich eine sehr gute Konkordanz dar, d. h., Zentren, die eine hohe Maximaldosis benutzen, verwendeten auch hohe Cut-off-Punkte und erreichten so höhere Testsensitivitäten.

Wenn man davon ausgeht, dass unspezifische Provokationen primär für den Ausschluss eines Bronchialasthmas geeignet sind und dementsprechend eine hohe Sensitivität (und damit geringe Spezifität) haben sollen, dann ist eines der wichtigsten Ergebnisse dieser Studie die Forderung nach einer Standardisierung der Tests bezüglich einer akzeptablen Sensitivität, d. h. ausreichend hohen Provokationsdosis. In unserer Testpopulation bedeutet dies eine „realistische“ Sensitivität zwischen 0,70 und 0,80. Wenn standardisierte Tests nicht erhältlich bzw. aus anderen Gründen nicht durchführbar sind, sollte diese Forderung verbindlich umgesetzt werden, zumal sie einfach zu erfüllen ist.

## Literatur

- <sup>1</sup> Scheffer AL. Guidelines for the diagnosis and the management of asthma. *J Allergy Clin Immunol* 1991; 88: 425–434
- <sup>2</sup> Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. Lung volumes and forced ventilatory flows. Report working party standardization of lung function tests European community for steel and coal. *Eur Respir J* 1993; 6 (Suppl 16): 5–40
- <sup>3</sup> Britton J, Mortagy A, Tattersfield A. Histamine challenge testing: comparison of three methods. *Thorax* 1986; 41: 128–132
- <sup>4</sup> Weersink EJ, v d Elshout FJ, van Herwaarden CV, Folgering H. Bronchial responsiveness to histamine and metacholine with forced expiration and with the forced oscillation technique. *Respir Med* 1995; 89 (5): 351–356
- <sup>5</sup> Hargreave FE, Ryan G, Thompson NC, O'Byrne PM, Latimer K, Juniper EF, Dolovich J. Bronchial responsiveness to histamine or methacholine in asthma: measurement and clinical significance. *J Allergy Clin Immunol* 1981; 68: 347–355
- <sup>6</sup> Gonsior E, Krüger M, Meier-Sydow J. Bronchial antigen provocation tests with the use of whole bodyplethysmography. *Acta allergol* 1976; 31 (4): 283–346
- <sup>7</sup> Schmitz-Schumann M, Markendorf A, Virchow jun C. Erfahrungen mit dem modifizierten inhalativen Provokationstest mit Histamin nach Cockcroft unter Verwendung des Pari Provocationstest I. *Pneumologie* 1991; 45: 670–673
- <sup>8</sup> Koehler D. Problematik der dosisgenauen Inhalation, dargestellt am Beispiel des Pari Provocationstestgerätes I. *Pneumologie* 1991; 45: 659–669
- <sup>9</sup> Schlegel J. Methodische Daten zum inhalativen Methacholin-Provokationstest (Dosis-Wirkungsbeziehung). *Pneumologie* 1991; 45: 684–686
- <sup>10</sup> Klein G, Wagner-Heck M, Zinn CG. Inhalativer Carbachol-Provokationstest mit einer Reservoirmethode. *Pneumologie* 1991; 45: 679–683
- <sup>11</sup> Zweig MH, Campbell G. Receiver operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem* 1993; 39 (4): 561–577
- <sup>12</sup> American Thoracic Society. Guidelines for methacholine and exercise challenge testing – 1999. *Am J Respir Crit Care Med* 2000; 161: 309–329
- <sup>13</sup> Sterk PJ, Fabbri LM, Quanjer PH, Cockcroft DW, O'Byrne PM, Anderson SD, Juniper EF, Malo JL. Airway responsiveness. Standardized challenge testing with pharmacological, physical and sensitizing stimuli in adults. *Eur Respir J* 1993; 6 (suppl 16): 53–83
- <sup>14</sup> Deutsche Gesellschaft für Pneumologie, Arbeitskreis „Bronchiale Provokationstests“. Leitlinien für die Durchführung bronchialer Provokationstests mit pharmakologischen Substanzen. *Pneumologie* 1998; 52: 214–220