EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Original article

Diagnosis of asthma in children: findings from the Swiss Paediatric Airway Cohort

Carmen CM. de Jong, Eva SL. Pedersen, Rebeca Mozun, Dominik Müller-Suter, Anja Jochmann, Florian Singer, Carmen Casaulta, Nicolas Regamey, Alexander Moeller, Cristina Ardura-Garcia, Claudia E Kuehni

Please cite this article as: de Jong CC, Pedersen ES, Mozun R, *et al*. Diagnosis of asthma in children: findings from the Swiss Paediatric Airway Cohort. *Eur Respir J* 2020; in press (https://doi.org/10.1183/13993003.00132-2020).

Diagnosis of asthma in children: findings from the Swiss Paediatric Airway Cohort

Carmen CM. de Jong[1], Eva SL. Pedersen[1], Rebeca Mozun[1], Dominik Müller-Suter[2], Anja Jochmann[3], Florian Singer[4,5], Carmen Casaulta[4,6], Nicolas Regamey[7], Alexander Moeller[8], Cristina Ardura-Garcia[1], Claudia E Kuehni[1,4]

Affiliation:
1. Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland
2. Paediatric Respiratory Medicine, Kantonsspital Aarau, Aarau, Switzerland
3. Paediatric Respiratory Medicine, University Children's Hospital (UKBB), University of Basel, Basel, Switzerland
4. Paediatric Respiratory Medicine, Children's University Hospital of Bern, University of Bern, Bern, Switzerland
5. PedNet Bern, Children's University Hospital of Bern, University of Bern, Bern, Switzerland
6. Paediatric Respiratory Medicine, Kantonsspital Graubunden, Chur, Switzerland
7. Division of Paediatric Pulmonology, Children's Hospital, Lucerne, Switzerland
8. Division of Paediatric Pulmonology, University Children's Hospital of Zurich, University of Zurich, Zurich, Switzerland

Corresponding author

Claudia E. Kuehni, University of Bern, Institute of Social and Preventive Medicine, Mittelstrasse 43, CH-3012 Bern, Switzerland, claudia.kuehni@ispm.unibe.ch

Take Home Message (142/256 characters): Asthma diagnosis does not seem straightforward, even for experienced pulmonologists; this highlights the need for new evidence-based guidance.

Conflict of interest: None

Funding: Swiss National Science Foundation: 32003B_162820, the Lung League St. Gallen and the Allergiestiftung Ulrich Müller-Gierok.

Author Contributions: C. Kuehni and C. de Jong conceptualised and designed the study. D. Müller-Suter, A. Jochmann, F. Singer, C. Casaulta, N. Regamey, and A. Moeller supervised data collection. C. de Jong analysed the data and drafted the manuscript. E. Pedersen and C. Ardura supported the statistical analysis and gave input for interpretation of the data. All authors critically revised the manuscript and approved the final manuscript as submitted.

Keywords: asthma, wheeze, diagnostic testing, cohort, children

Word count: 2996/3000

Abstract (248/250)

Introduction: Diagnosing asthma in children remains a challenge because respiratory symptoms are not specific and vary over time.

Aim: In a real-life observational study, we assessed the diagnostic accuracy of respiratory symptoms, objective tests, and two paediatric diagnostic algorithms proposed by GINA and NICE to diagnose asthma in school-aged children.

Methods: We studied children aged 5-17 years referred consecutively for evaluation of suspected asthma to pulmonary outpatient clinics. Symptoms were assessed by parental questionnaire. The investigations included specific IgE measurement or skin prick tests, measurement of fractional exhaled nitric oxide, spirometry, body plethysmography, and bronchodilator reversibility. Asthma was diagnosed by paediatric pulmonologists based on all available data. We assessed diagnostic accuracy of symptoms, tests, and diagnostic algorithms by calculating sensitivity, specificity, positive and negative predictive values, and area under the curve (AUC).

Results: Among 514 participants, 357(70%) were diagnosed with asthma. The combined sensitivity and specificity (sensitivity/specificity) was highest for any wheeze (0.75/0.65), dyspnoea (0.56/0.76), and wheeze triggered by colds (0.58/0.78) or by exercise (0.55/0.74). Of the diagnostic tests, the AUC was highest for specific total resistance (sRtot) (0.73) and lowest for the residual volume (RV) total lung capacity (TLC) ratio (0.56). The NICE algorithm had a sensitivity of 0.69 and specificity of 0.67, whereas the GINA algorithm had a sensitivity of 0.42 and specificity of 0.90.

Conclusion: This study confirms the limited usefulness of single tests as well as existing algorithms for the diagnosis of asthma. It highlights the need for new and more appropriate evidence-based guidance.

Introduction

Diagnosing asthma in children remains a challenge because respiratory symptoms such as wheeze and cough are not specific and vary over time, and asthma lacks an effective stand-alone diagnostic test. Clinically, physicians diagnose asthma based on a combination of symptoms, physical examination, and diagnostic tests [1-3]. Among available tests, spirometry and body plethysmography in combination with bronchodilator reversibility testing demonstrate reversible airway obstruction. Bronchial provocation tests measure bronchial hyper-responsiveness (BHR), fractional exhaled nitric oxide (FeNO) indicates airway inflammation, and allergy tests show underlying atopy. Diagnostic algorithms that combine these tests have recently been proposed by the National Institute for Health and Care Excellence (NICE) and the Global Initiative for Asthma (GINA) [1-3].

However, the accuracy of these diagnostic algorithms in school-aged children suspected for asthma is uncertain, which can lead to both under- and over treatment [3-5]. Systematic literature reviews done as part of the ongoing task force of the European Respiratory Society (ERS) on asthma diagnosis in children find few relevant studies [1-3]. A population-based assessment of the diagnostic accuracy of the algorithm proposed in the NICE guideline recommends against guideline implementation in the absence of better evidence [6]. In a recent assessment of diagnostic tests in 111 school-aged outpatients referred for suspected asthma, we found that accuracy was highest for reported triggers and severity of wheeze, FeNO, and BHR testing by methacholine or exercise [7]. But because the study of the NICE algorithm was based on the general population, and our study took place in a research setting, we lack data on the usefulness of these tests in everyday clinical situations. Body plethysmography in particular has also not been previously assessed. We therefore set out to assess the diagnostic accuracy of respiratory symptoms, diagnostic tests, and two paediatric diagnostic algorithms by GINA and NICE in school-aged children who were referred to respiratory outpatient clinics with suspected asthma.

Methods

Study population and study design

We relied upon baseline measurements of the Swiss Paediatric Airway Cohort (SPAC), a prospective clinical cohort embedded in routine care [8]. This analysis used data from children aged 5-17 years invited to participate between July 2017 and June 2019. The primary care physicians wrote down the reason for referral in the referral letter. The 514 participants were consecutive referrals for symptoms suggestive of asthma: wheeze, cough (except wet cough without wheeze), exercise induced breathing problems, or dyspnoea. The children referred by their primary care physician for other reasons such as investigations of allergic rhinoconjunctivitis alone, or of sleep-disordered breathing, were not included. The study inclusion process is shown in Figure 1.

Study procedures

Participant families received a parental questionnaire either accompanying an invitation letter to attend the clinic or upon arrival at the outpatient clinic. At the visit, children underwent clinical evaluation, allergy testing (unless allergy test results were reported in the referral letter), measurement of FeNO, spirometry, body plethysmography, and bronchodilator reversibility testing. When indicated, children also underwent bronchial hyperresponsiveness testing by methacholine, exercise, or mannitol either during the same visit or at a follow-up visit within three months. Ethical approval was obtained from the Bernese ethics committee (KEB 2016-02176) and all participating parents and adolescents aged 14 years or older gave written informed consent.

Clinical asthma diagnosis

The clinical diagnosis of definite asthma, probable asthma, or other disease was the one noted by the experienced paediatric pulmonologist in the letter sent back to the referring primary care physician. The diagnosis was based on medical history, clinical examination, and all test results, and was regarded as the reference standard. The results of the parental questionnaire were not available to the paediatric pulmonologists when making the diagnosis. When the diagnosis was unclear (that is, was described as

probable asthma) and a follow-up visit took place within 3 months, we used the clinical diagnosis from the follow-up visit.

Respiratory symptoms and diagnostic tests

The parental questionnaire included key questions about lower respiratory symptoms from the International Study of Asthma and Allergies in Childhood (ISAAC) and further detailed questions from the Leicester respiratory cohort questionnaires [9, 10]. Dyspnoea was assessed using the question: “Hatte Ihr Kind in den letzten 12 Monaten Atemnot?” (German) meaning “In the past 12 months did your child have episodes of breathlessness?” in English. All diagnostic tests were performed according to published guidelines [11, 12]. Children were asked to withhold short-acting beta2-agonists for 8 hours, inhaled corticosteroids, leukotriene antagonists, and long acting beta2-agonists for 24 hours, and antihistamines and sodium cromoglycate for >72 hours before the visit.

Skin prick tests and specific IgE measurements were used to measure atopy. Skin prick tests were done using histamine, saline and birch, grass, mugwort, alternaria, cat, and house dust mites (D. pteronyssinus and farinae) [13]. A wheal size of ≥3 mm was considered positive in case the positive control (histamine) had a wheal size of ≥3 mm and the negative control (saline) had a wheal size of <3mm. Specific IgE levels for birch, grass, mugwort, alternaria, cat, house dust mites were measured in serum samples using the fluorescence enzyme immunoassay/immunocap (Thermo Fisher Scientific, Uppsala, Sweden). IgE levels were considered positive at the detection threshold (≥0.35 kU/l).

Fractional exhaled nitric oxide (FeNO) was measured in doublets before spirometry using the portable multi-gas analyser (NIOX MINO®, Aerocrine, Sweden) in St. Gallen and the CLD 88sp (Ecomedics) in Bern, Basel, Aarau, Zurich, and Luzern in accordance with published guidelines [12]. These devices show good agreement [14].

Spirometry and body plethysmography were performed using American Thoracic Society (ATS) criteria for paediatric lung function testing and a Jaeger masterscope (Erich Jaeger GmbH, Würzburg, Germany) using JLAB software (version 4.34) (Basel and St. Gallen) or MasterScreen Pneumo spirometer using Sentrysuite

software (Bern, Zurich, Aarau, and Lucerne). Spirometry was done in triplicate by experienced lung function technicians who performed quality control during the measurement and recorded the best measurement. The flow-volume curve was checked by the responsible paediatric pulmonologist. Results of the spirometry (forced expiratory volume in 1 second (FEV_1) and forced expiratory flow (FEF) at 50% of the forced vital capacity (FVC)) are expressed as z-scores based on GLI-2012 reference standards [15] and as proportion (FEV_1/FVC). The results of body plethysmography are expressed as kPa·s for the specific effective airway resistance (sReff) and specific total airway resistance (sRtot) and as proportion (residual volume/total lung capacity).

Bronchodilator reversibility tests were performed if FEV1 was ≤90%, FEF75 ≤67%, FEF50 ≤67%, or FEF25 ≤67% (Lucerne, Zurich and Aarau), if SReff >180% or FEF50 <80% (Bern), or in all patients (St. Gallen and Basel). All centres gave salbutamol 400 μg (Ventolin® pMDI via spacer) to assess bronchodilator reversibility. Spirometry was repeated in duplicate after 10 (Lucerne and Basel), 15 (St. Gallen, Aarau, and Bern) or 20 minutes (Zurich). Bronchodilator reversibility was calculated by the following equation: (post-bronchodilator FEV1–pre-bronchodilator FEV1)x100/pre-bronchodilator FEV1.

Statistical analysis

We calculated sensitivity, specificity, positive predictive value and negative predictive value, Youden's Index (sensitivity + specificity – 1), area under the curve (AUC) and 95% confidence intervals (CI) for the reported symptoms and diagnostic tests to diagnose asthma. The cut-off with the best diagnostic accuracy was the value with the highest Youden's index. We assessed the diagnostic accuracy of tests if they were performed in at least 70% of the children. To assess the diagnostic accuracy of bronchodilator reversibility we did a subanalysis in children with obstructive lung function (FEV1/FVC <80%) [1, 16]. We performed an initial sensitivity analysis in which we classified children with "probable asthma" as having "no asthma". We did a second sensitivity analysis using only steroid naïve children. We applied the asthma diagnosis algorithms by GINA and NICE to assess how they would have performed in a clinical setting and calculated sensitivity,

specificity, positive predictive value, and negative predictive value. We used STATA software (version 15; College Station, Texas) for statistical analysis.

Results

Characteristics of the study population

Of the 514 children fulfilling the inclusion criteria, 294 (57%) were male and the median age was 9 years. Nearly two-thirds of the referred children reported wheeze, and over half wheeze and/or cough. Asthma had been diagnosed in 356 (69%) of the participants. Table 1 presents the full characteristics of the participants. Exercise limitation (56 children, 11%) and cough (71 children, 14%), both not due to asthma, were frequent other diagnoses (table S1).

Table 1. Characteristics of the study participants (N=514)

	Total n (%)	
Age, median [IQR]	9	[7-12]
Sex, male	294	(57)
BMI, median z-score [IQR]	0.3	[-0.4-1.1]
Respiratory symptoms in the last 12 months		
Any wheeze	317	(62)
More than 3 attacks of wheeze	170	(33)
Wheeze with colds	230	(45)
Exercise-induced wheeze	232	(45)
Wheeze triggered by pollen	127	(25)
Wheeze triggered by house dust	81	(16)
Wheeze triggered by pets	64	(12)
Awakening due to wheeze	182	(35)
Cough longer than 4 weeks	214	(42)
Night cough	271	(53)
Cough more than others	281	(55)
Dyspnoea	230	(45)
Inhalation medication in the last 12 months		
Any	395	(77)
Short-acting B2-agonist, alone	152	(30)
ICS +/- Short-acting B2-agonist	114	(22)
ICS + Long-acting B2-agonist	129	(25)

IQR: inter quartile range, BMI: body mass index, ICS: inhaled corticosteroids

Diagnostic accuracy of respiratory symptoms to diagnose asthma

Any reported wheeze in the past 12 months had the highest sensitivity (74%) and Youden's Index (0.39) for asthma (table 2). Specificity was highest for frequent attacks (>3/year) (84%), awakening due to wheeze (82%), and wheeze triggered by pollen (92%), house dust (97%), or pets (97%). Youden's Index was also relatively high for wheeze triggered by colds (0.36), exercise (0.30), and dyspnoea (0.31).

Table 2. Diagnostic accuracy of respiratory symptoms to diagnose asthma (N=514)

	A+S+ n	**A-S+** n	**A+S-** n	**A-S-** n	**Sens** % (95%CI)	**Spec** % (95%CI)	**PPV** % (95%CI)	**NPV** % (95%CI)	**YI**
Respiratory symptoms in the past 12 months									
Any wheeze	262	55	90	100	74 (70-79)	65 (56-72)	83 (78-87)	53 (45-60)	0.39
> 3 attacks of wheeze	145	25	200	129	42 (37-47)	84 (77-89)	85 (79-90)	39 (34-45)	0.26
Wheeze with colds	196	34	141	120	58 (53-63)	78 (71-84)	85 (80-90)	46 (40-52)	0.36
Exercise-induced wheeze	192	40	154	115	55 (50-61)	74 (67-81)	83 (77-87)	43 (37-49)	0.30
Wheeze triggered by									
Pollen	115	12	227	145	34 (29-39)	92 (87-96)	91 (84-95)	39 (34-44)	0.26
House dust	76	5	259	150	23 (18-28)	97 (93-99)	94 (86-98)	37 (32-42)	0.19
Pets	59	5	274	152	18 (14-22)	97 (93-99)	92 (83-97)	36 (31-40)	0.15
Awakening due to wheeze	155	27	191	127	45 (39-50)	82 (76-88)	85 (79-90)	40 (35-46)	0.27
Cough > 4 weeks	140	74	209	81	40 (35-45)	52 (44-60)	65 (59-72)	28 (23-33)	-0.08
Night cough	189	82	153	74	55 (50-61)	47 (39-56)	70 (64-75)	33 (27-39)	0.03
Cough more than others	200	81	146	70	58 (52-63)	46 (38-55)	71 (65-76)	32 (26-39)	0.04
Dyspnoea	192	38	154	116	55 (50-61)	75 (68-82)	83 (78-88)	43 (37-49)	0.31

A+S+: children with asthma diagnosis and reported symptom, A-S+: children without asthma diagnosis but with symptom, A+S-: children with asthma diagnosis but without symptom, A-S-: children without asthma and without symptom, Sens: sensitivity, Spec: specificity, PPV: positive predictive value, NPV: negative predictive value YI: Youden's-Index: Sensitivity + Specificity -1

Diagnostic accuracy of tests to diagnose asthma

The tests done in each centre and their results are shown in tables S2 and S3. Allergy tests were performed in 467 (91%) of the 514 children. FeNO was performed in 501 (97%), spirometry in all 514, body plethysmography in 432 (84%), and bronchodilator reversibility in 381 children (74%). Of these performed measurements, we excluded 63, 19, 45, and 15, respectively, because of poor quality. The accuracy of bronchial hyperresponsiveness testing was not assessed, because it was only performed in 210 (41%) children.

The cut-off values with the best diagnostic accuracy were ≥1 positive test for the allergy test, ≥23ppb for FeNO, ≤-0.7 *z* score for FEV1, <84% for FEV1/FVC, ≤-0.3 *z* score for FEF50, ≥0.9 kPa·s for sReff, ≥1.1 kPa·s for sRtot, ≥25% for RV/TLC, and ≥7% increase in FEV1 for the bronchodilator reversibility test (table 3). The diagnostic accuracy (area under the curve) was highest for sRtot (0.74), allergy tests (0.70), FEF50 (0.69), and FeNO (0.68). The accuracy was lowest for RV/TLC (0.56) and the bronchodilator reversibility test (0.60). However, bronchodilator reversibility had the highest accuracy (0.75) when we analysed only data from children with FEV1/FVC <80% (figure 2).

Table 3. Diagnostic accuracy of diagnostic tests to diagnose asthma N=514

	A+T+ n	A-T+ n	A+T- n	A-T- n	Sens % (95%CI)	Spec % (95%CI)	PPV % (95%CI)	NPV % (95%CI)	YI	AUC
Clinical tests										
Allergy test[1]										0.70
≥1 positive test*	260	52	81	74	76 (71-81)	59 (50-67)	83 (79-87)	48 (40-56)	0.35	
≥2 positive tests	199	34	142	92	58 (53-64)	73 (64-81)	85 (80-90)	39 (33-46)	0.31	
FeNO										0.68
≥20ppb	157	31	147	103	52 (46-57)	77 (69-84)	84 (77-89)	41 (35-48)	0.29	
≥21ppb	153	27	151	107	50 (45-56)	80 (72-86)	85 (79-90)	41 (35-48)	0.30	
≥23ppb*	145	18	159	116	48 (42-53)	87 (80-92)	89 (83-93)	42 (36-48)	0.34	
≥25ppb	139	16	165	118	46 (40-52)	88 (81-93)	90 (84-94)	42 (36-48)	0.34	
Spirometry										
FEV1										0.66
≤-0.7 z-score*	148	30	195	121	43 (38-49)	80 (73-86)	83 (77-88)	38 (33-44)	0.23	
≤-1.0 z-score	109	23	234	128	32 (27-37)	85 (78-90)	83 (75-89)	35 (30-41)	0.17	
FEV1/FVC										0.65
<80%	120	15	216	128	36 (31-41)	90 (83-94)	89 (82-94)	37 (32-43)	0.25	
<84%*	174	33	162	110	52 (46-57)	77 (69-84)	84 (78-89)	40 (35-47)	0.29	
<90%	245	84	91	59	73 (68-78)	41 (33-50)	74 (69-79)	39 (31-48)	0.14	
FEF50										0.69
≤-0.3 z-score*	171	31	122	93	58 (52-64)	75 (66-82)	85 (79-89)	43 (37-50)	0.33	
≤-1.0 z-score	96	13	197	111	33 (27-38)	90 (83-94)	88 (80-93)	36 (31-42)	0.22	
Bodyplethysm.										
sReff[2]										0.66
≥0.9 kPa·s/l*	118	25	114	76	51 (44-57)	75 (66-83)	83 (75-88)	40 (33-47)	0.26	
≥1.0 kPa·s/l	96	18	136	83	41 (35-48)	82 (73-89)	84 (76-90)	38 (31-45)	0.24	
sRtot[3]										0.74
≥1.0 kPa·s/l	35	11	4	6	90 (76-97)	35 (14-62)	76 (61-87)	60 (26-88)	0.25	
≥1.1 kPa·s/l *	32	7	7	10	82 (66-92)	59 (33-82)	82 (66-92)	59 (33-82)	0.41	
RV/TLC										0.56
≥25%*	204	80	61	36	77 (71-82)	31 (23-40)	72 (66-77)	37 (28-48)	0.08	
Bronchodilator rev.										0.60
FEV_1										
≥7% increase*	188	43	86	49	69 (63-74)	53 (43-64)	81 (76-86)	36 (28-45)	0.22	
≥10% increase	160	37	114	55	58 (52-64)	60 (49-70)	81 (75-86)	33 (26-40)	0.18	
≥12% increase	145	34	129	58	53 (47-59)	63 (52-73)	81 (74-86)	31 (24-38)	0.16	
Bronchodilator rev.										0.75
if FEV1/FVC <80%[4]										
≥7% increase*	89	4	23	9	79 (71-87)	69 (39-91)	96 (89-99)	28 (14-47)	0.49	
≥10% increase	73	3	39	10	65 (56-74)	77 (46-95)	96 (89-99)	20 (10-34)	0.42	
≥12% increase	65	2	47	11	58 (48-67)	85 (55-98)	97 (90-99)	19 (10-31)	0.43	

A+T+: children with asthma diagnosis and positive test result, A-T+: children without asthma diagnosis but positive test result, A+T-: children with asthma diagnosis but negative test result, A-T-: children without asthma and negative test result, Sens: sensitivity, Spec: specificity, PPV: positive predictive value, NPV: negative predictive value, YI: Youden's-Index: Sensitivity + Specificity -1, AUC: area under the curve, FeNO: fractional exhaled nitric oxide, ppb: parts per billion, FEV_1: forced expiratory volume in 1 second, FVC: forced vital capacity, FEF50: forced expiratory flow at 50% of FVC, sReff: specific effective airway resistance, sRtot: specific total airway resistance, RV: residual volume, TLC: total lung capacity, Bronchodilator rev.: bronchodilator reversibility

Displayed cut-offs chosen based on proposed cut-offs from previous publications

*Cut-off with maximum combined sensitivity and specificity (highest Youden's-Index)

[1] Number allergens for which the skin prick test is positive: wheal size ≥3 or the specific IgE test was positive: ≥0.35 kU/l.

[2] Reported by 4 centres

[3] Reported by 2 centres

[4] N= 126, cut-off chosen based on proposed cut-off from previous publications and guidelines

Sensitivity analysis

In the sensitivity analysis, where we classified children diagnosed with "probable asthma" as having "no asthma", the cut-offs with highest combined sensitivity and specificity changed only slightly for FEV1 (from ≤-0.7 *z* score to ≤-0.2 *z* score) and FEF50 (from ≤-0.3 *z* score to ≤-0.5 *z* score). The diagnostic accuracy (area under the curve) remained highest for FEF50 (0.73) and FeNO (0.70) and increased for FEV1/FVC (from 0.65 to 0.72), and sReff (from 0.66 to 0.68). The accuracy remained lowest for RV/TLC (0.53) and the bronchodilator reversibility test (0.60) (table s4).

In the second sensitivity analysis, which included only ICS naïve children, asthma was diagnosed in 156 (58%) of the 271 children (table s5). The cut-offs with highest combined sensitivity and specificity changed for FeNO from ≤23 ppb to ≤28 ppb, FEV1/FVC from 84% to 86%, sRtot from 1.1 kPa·s to 1.5 kPa·s, and RV/TLC from ≥25% to ≥35%. The diagnostic accuracy remained highest for sRtot (0.77), the allergy test (0.71), FEF50 (0.70), and FeNO (0.71) as in the main analysis. The accuracy was still lowest for RV/TLC (0.51) and bronchodilator reversibility (0.56) (table S6).

Diagnostic accuracy of algorithms to diagnose asthma

We applied the GINA diagnostic algorithm to the 514 children suspected for asthma. We were able to pass 91 children through the proposed pathway to the step "treat for asthma" (figure 3). Of these, 81 (positive predictive value [PPV] 89%) were diagnosed with asthma by our paediatric pulmonologist. Of the 210 children we could pass through the algorithm to the step "consider alternative diagnosis," 111 were diagnosed with asthma, and 99 (negative predictive value [NPV] 47%) were not. The sensitivity of the algorithm was 42% and the specificity was 90%. In 168 children, the GINA algorithm would have been inconclusive because they ended at the step "repeat on another occasion or arrange other tests"; the paediatric pulmonologists in our study diagnosed 132 of these children with asthma.

We also applied the NICE diagnostic algorithm to the 514 children suspected for asthma. We were only able to pass 38 children through to the step "diagnose asthma" (figure S2). Of these 38 children, 35 (PPV 92%)

were diagnosed with asthma by the paediatric pulmonologists. Of the 22 children whom we could pass through to the step "refer for specialist assessment," 18 were diagnosed with asthma and 6 (NPV 27%) were not. The sensitivity was thus 69% and the specificity 67%. However, 362 (83%) children were stuck at "2 weeks of PEF monitoring." From this step onwards we could not apply the NICE diagnostic algorithm.

Discussion

This study found that the commonly used tests, alone or in combination as suggested by GINA or NICE, are not very helpful in diagnosing asthma in routine, clinical care. We found that the combined sensitivity and specificity to diagnose asthma in our study was highest for any wheeze (sensitivity/specificity, 74/65) and the diagnostic accuracy (area under the curve) was highest for sRtot (0.74), a positive allergy test (0.70), FEF50 (0.69), FeNO (0.68), and the bronchodilator reversibility test in children with FEV1/FVC <80% (0.75), and lowest for RV/TLC (0.56). The NICE algorithm relied too much on 2 weeks' PEF monitoring, which should have been done in 83% of children according to the algorithm. This is not practical for an outpatient setting. The GINA paediatric diagnostic algorithm was specific (90%), but not sensitive (42%).

This is the first study of the diagnostic accuracy of body plethysmography and the largest study to investigate the usefulness of respiratory symptoms, diagnostic tests, and algorithms to diagnose asthma in routine care. A few studies have assessed the accuracy of symptoms and tests in school-aged children consecutively referred for evaluation of possible asthma [7, 17-19]. They all found that reported wheeze was sensitive (ranging 75-86%) but not specific (64-73%), and that frequent wheeze and awakening due to dyspnoea were specific (84-90%) but not sensitive (33-54%), which is in line with our findings. Another way to phrase this would be that, if for example frequent wheeze is mentioned in the history, a physician is highly likely to make an asthma diagnosis. In our previous study in a different clinical population, the combined sensitivities and specificities were highest for the same symptoms. In that study, wheeze with colds and dyspnoea also scored high [7]. Also in that study and as reported by Woo et al., we confirmed that a positive skin prick test was sensitive (68-90%) but not specific (32-40%) [20]. The area under the

curve (AUC) for FeNO in our study (0.68) was lower than in our previous study (0.79), and in a Korean study by Eom et al. (0.80) [7, 21]. In steroid naïve children, we expected the diagnostic accuracy for FeNO to be higher, because ICS treatment reduces FeNO levels. This was not the case, which could have two reasons: (1) the ICS naïve children were less severely affected and episodes were mostly triggered by colds. FeNO levels are lower in mildly affected children without allergic triggers, reducing the diagnostic accuracy in treated children. (2) The question on ICS use was imprecise. We asked about use of ICS in the last 12 months. This includes current use of ICS but did not allow for an analysis of current use of ICS separately. FEV1/FVC had low diagnostic accuracy in all studies. FEF25-75 (0.81) and FEF50 (0.69) seem to perform better. Differences in the AUC between the Korean and our study could be due to their exclusion of the children with unclear asthma. A study by Korten et al. found that adding sRtot and RV/TLC measured during bodyplethysmography to spirometry measures improved the agreement with the asthma control test in a cohort of 145 asthmatic children from Germany [22]. In our study, sRtot had the highest AUC to diagnose asthma. However, if we want to assess its value in addition to spirometry, we would need to perform a prospective study.

The NICE algorithm has previously been tested using data from the Manchester Asthma and Allergy Study, a population-based cohort of 1184 children aged 13-16 years, of which 89 were symptomatic but not regularly inhaling corticosteroids [6]. That study found that less stringent cutoff values had higher sensitivity and specificity than those proposed in the algorithm. However, the Manchester study compared children with asthma defined by parent reports of wheeze and asthma treatment in the past 12 months plus a doctor diagnosis of asthma ever in life, to healthy children, excluding from the analysis all those with possible asthma. In clinical practice, we want to distinguish children with asthma from those with respiratory symptoms due to other causes, not from healthy children. In our clinical population, only 38 children out of 514 could be diagnosed with asthma based on the NICE algorithm (FEV1/FVC <70% and bronchodilator responsibility of ≥12%) in our clinical population. Nearly all (83%) would have needed an additional two weeks' peak expiratory flow (PEF) monitoring followed by a second visit to the outpatient clinic (Figure S2). In addition to the fact that this test is not used in most countries, it would also not be

practical for a busy outpatient clinic. In low income settings, however, PEF monitoring might be the only test available and its value should be studied further. We also found less stringent cut-off values of <84% for FEV1/FVC, ≥7% for bronchodilator reversibility, and ≥23ppb for FeNO to have higher sensitivity and specificity than the values proposed by the NICE algorithm: <70% for FEV1/FVC, ≥12% for bronchodilator reversibility, and ≥35ppb for FeNO. The accuracy of the GINA diagnostic algorithm has not previously been studied. At 90% its specificity is relatively high, but this would still lead to 10% overdiagnosis of asthma. The sensitivity was only 42%, which means that the GINA algorithm would identify less than half of the children diagnosed by paediatric pulmonologists.

A main weakness of this study, which is unavoidable, is that the reference standard for asthma diagnosis, the physician's diagnosis, drew upon patient history and diagnostic test results, of which we were assessing the accuracy. So, while we found that wheeze triggered by pollen is very specific for asthma (specificity 92%), we could phrase this finding also the other way around, i.e. that physicians are unlikely to give another diagnosis than asthma if parents report wheeze triggered by pollen. However, given the lack of a stand-alone diagnostic test for asthma, the physician's diagnosis based on a full patient history, physical examination, and diagnostic test results is closest to the true diagnosis [1, 23, 24]. The multicentre study adds heterogeneity in the tests and diagnoses, but it increases generalisability of the findings. Finally, some tests were not done in all children. This could have introduced a selection bias because the children who were tested could differ from those who were not. However, the percentage not tested was low, limiting the potential bias because we only evaluated tests done in more than 70% of the children.

The main strength of our study is that it was embedded in routine care and included the whole spectrum of children newly referred with suspected asthma to paediatric respiratory outpatient clinics. We could also restrict the analysis to a steroid-naïve population, with comparable findings. Finally, we could assess the diagnostic accuracy of body plethysmography, which has not been done before and might be a useful test to add to spirometry for the diagnosis of asthma.

Our findings highlight the need for better diagnostic algorithms combining respiratory symptoms and objective tests to diagnose asthma. The algorithms proposed by GINA and NICE guidelines do not seem to be very practical for an outpatient setting, and did not agree well with pulmonologists' diagnoses. Our findings confirm that the cut-offs used in the NICE algorithm are not appropriate for children. If we require a low FEV1/FVC ratio (less than 90% predicted) or significant bronchodilator responsiveness (>12%) to diagnose asthma, we are likely to underdiagnose asthma in children - particularly those who are already on ICS. This highlights the need to base diagnostic algorithms on clinical studies of appropriate age groups to generate evidence for threshold values of different tests relevant to the diagnosis of asthma in those groups—children in particular. The next step in research should take a systematic approach to assessing respiratory symptoms, allergy, FeNO, spirometry, body plethysmography, and bronchodilator reversibility tests in all children to develop an accurate diagnostic algorithm combining these tests.

Acknowledgements

We thank all families, lab technicians, and doctors participating in the Swiss Paediatric Airway Cohort. We thank Christopher Ritter for his editorial assistance.

References

[1] GINA. Global strategy for asthma management and prevention. https://ginasthma.org/wp-content/uploads/2019/04/wms-GINA-2019-report-V1.3-002.pdf. Date last updated: 2019. Date last accessed: December 2019

[2] BTS-SIGN. Britisch guideline on the management of asthma. https://www.sign.ac.uk/assets/sign153.pdf. Date last updated: September 2016. Date last accessed: June 2019.

[3] NICE. Guideline asthma diagnosis and monitoring https://www.nice.org.uk/guidance/ng80/evidence/full-guideline-asthma-diagnosis-and-monitoring-pdf-4656178047. Date last updated: November 2017. Date last accessed: June 2019.

[4] Looijmans-van den Akker I, van Luijn K, Verheij T. Overdiagnosis of asthma in children in primary care: a retrospective analysis. Br J Gen Pract. 2016;66:e152-7.

[5] Yang CL, Simons E, Foty RG, Subbarao P, To T, Dell SD. Misdiagnosis of asthma in schoolchildren. Pediatr Pulmonol. 2017;52:293-302.

[6] Murray C, Foden P, Lowe L, Durrington H, Custovic A, Simpson A. Diagnosis of asthma in symptomatic children based on measures of lung function: an analysis of data from a population-based birth cohort study. Lancet Child Adolesc Health. 2017;1:114-23.

[7] de Jong CCM, Pedersen ESL, Mozun R, Goutaki M, Trachsel D, Barben J, Kuehni CE. Diagnosis of asthma in children: the contribution of a detailed history and test results. Eur Respir J. 2019.

[8] Pedersen ESL, de Jong CCM, Ardura-Garcia C, Barben J, Casaulta C, Frey U, Jochmann A, Latzin P, Moeller A, Regamey N, Singer F, Spycher B, Sutter O, Goutaki M, Kuehni CE. The Swiss Paediatric Airway Cohort (SPAC). ERJ Open Res. 2018;4.

[9] Kuehni CE, Brooke AM, Strippoli MP, Spycher BD, Davis A, Silverman M. Cohort profile: the Leicester respiratory cohorts. Int J Epidemiology. 2007;36:977-85.

[10] Asher MI, Keil U, Anderson HR, Beasley R, Crane J, Martinez F, Mitchell EA, Pearce N, Sibbald B, Stewart AW, Strachan D, Weiland SK, Williams HC. International study of asthma and allergies in childhood (ISAAC): rationale and methods. Eur Respir J. 1995;8:483-91.

[11] American Thoracic Society. Standardization of Spirometry, 1994 Update. Am J Respir Crit Care Med. 1995;152:1107-36.

[12] American Thoracic Society. Recommendations for standardized procedures for the on-line and off-line measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide in adults and children-1999. Am J Respir Crit Care Med. 1999;160:2104-17.

[13] Braun-Fahrlaender C, Wüthrich B, Gassner M, Gritze I, Neu U, Varonier H. Prävalenz und Risikofaktoren einer allergischen Sensibilisierung bei Schulkindern in der Schweiz. Allergologie. 1999;22:54-64.

[14] Schiller B, Hammer J, Barben J, Trachsel D. Comparability of a hand-held nitric oxide analyser with online and offline chemiluminescence-based nitric oxide measurement. Ped Allergy Immunol. 2009;20:679-85.

[15] Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MS, Zheng J, Stocks J. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. Eur Respir J. 2012;40:1324-43.

[16] Grzelewski T, Witkowski K, Makandjou-Ola E, Grzelewska A, Majak P, Jerzynska J, Janas A, Stelmach R, Stelmach W, Stelmach I. Diagnostic value of lung function parameters and FeNO for asthma in schoolchildren in large, real-life population. Pediatr Pulmonol. 2014;49:632-40.

[17] Ma TT, Zhuang Y, Gong HY, Yii AC, Wang XY, Shi HZ. Predictive value of respiratory symptoms for the diagnosis of pollen-induced seasonal asthma among children and adults in Inner Mongolia. Ther Clin Risk Manag. 2017;13:967-74.

[18] Brouwer AF, Visser CA, Duiverman EJ, Roorda RJ, Brand PL. Is home spirometry useful in diagnosing asthma in children with nonspecific respiratory symptoms? Pediatr Pulmonol. 2010;45:326-32.

[19] Crispino Santos MC ACA. A brief questionnaire for screening asthma among children and adolescents in Rio de Janeiro, Brazil. Allergol et Immunopathol. 2005;33:20-6.

[20] Woo SI, Lee JH, Kim H, Kang JW, Sun YH, Hahn YS. Utility of fractional exhaled nitric oxide (F(E)NO) measurements in diagnosing asthma. Respir Med. 2012;106:1103-9.
[21] Eom SY, Lee JK, Lee YJ, Hahn YS. Combining spirometry and fractional exhaled nitric oxide improves diagnostic accuracy for childhood asthma. Clin Respir J. 2019 [Epub ahead of print].
[22] Korten I, Zacharasiewicz A, Bittkowski N, Hofmann A, Lex C. Asthma control in children: Body Plethysmography in addition to spirometry. Pediatr Pulmonol. 2019;54(8):1141-1148.
[23] Rutjes AW, Reitsma JB, Coomarasamy A, Khan KS, Bossuyt PM. Evaluation of diagnostic tests when there is no gold standard. A review of methods. Health Technol Assess. 2007;11:iii, ix-51.
[24] Lucas JS, Barbato A, Collins SA, Goutaki M, Behan L, Caudri D, Dell S, Eber E, Escudier E, Hirst RA, Hogg C, Jorissen M, Latzin P, Legendre M, Leigh MW, Midulla F, Nielsen KG, Omran H, Papon JF, Pohunek P, Redfern B, Rigau D, Rindlisbacher B, Santamaria F, Shoemark A, Snijders D, Tonia T, Titieni A, Walker WT, Werner C, Bush A, Kuehni CE. European Respiratory Society guidelines for the diagnosis of primary ciliary dyskinesia. Eur Respir J 2017; 49: 1601090.

Figure 1 Inclusion of study participants

Figure 2 Receiver operating characteristics (ROC) curve of clinical tests to diagnose asthma.

Test	**Unit**
Allergy test npos	increase of 1 positive skin prick or specific IgE test
FeNO	increase of 1 parts per billion (ppb)
FEV1	increase of 0.1 z-score
FEV1/FVC	increase of 1%
FEF50	increase of 0.1 z-score
sReff	decrease of 0.01 kPa·s
sRtot	decrease of 0.01 kPa·s
RV/TLC	increase of 5%
Bronchodilator reversibility (BDR), FEV_1	increase of 1% in FEV_1

Figure 3 Diagnostic accuracy of the algorithm proposed by the GINA guideline

Numbers in black: number of patients at this step. Numbers in red: number of patients with doctor diagnosed asthma at this step. Algorithm +: treat for asthma. Algorithm -: consider alternative diagnosis. FEV1: forced expiratory volume in 1 second. FVC: forced vital capacity. BDR: bronchodilator reversibility

*168 patients would need to repeat the spirometry and bronchodilator reversibility measurement or bronchial hyperresponsiveness testing during another visit.

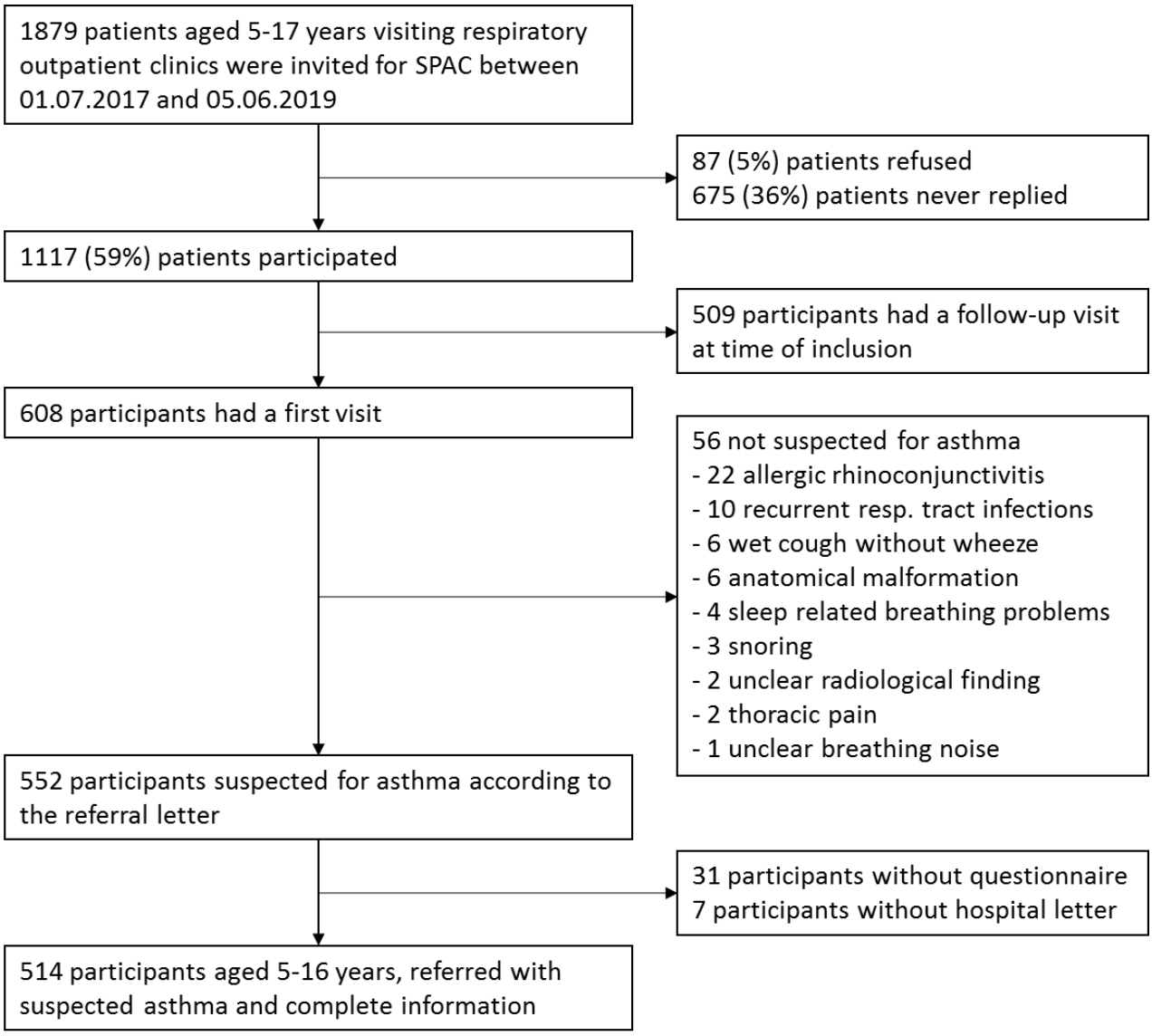


Figure 1 Inclusion of study participants

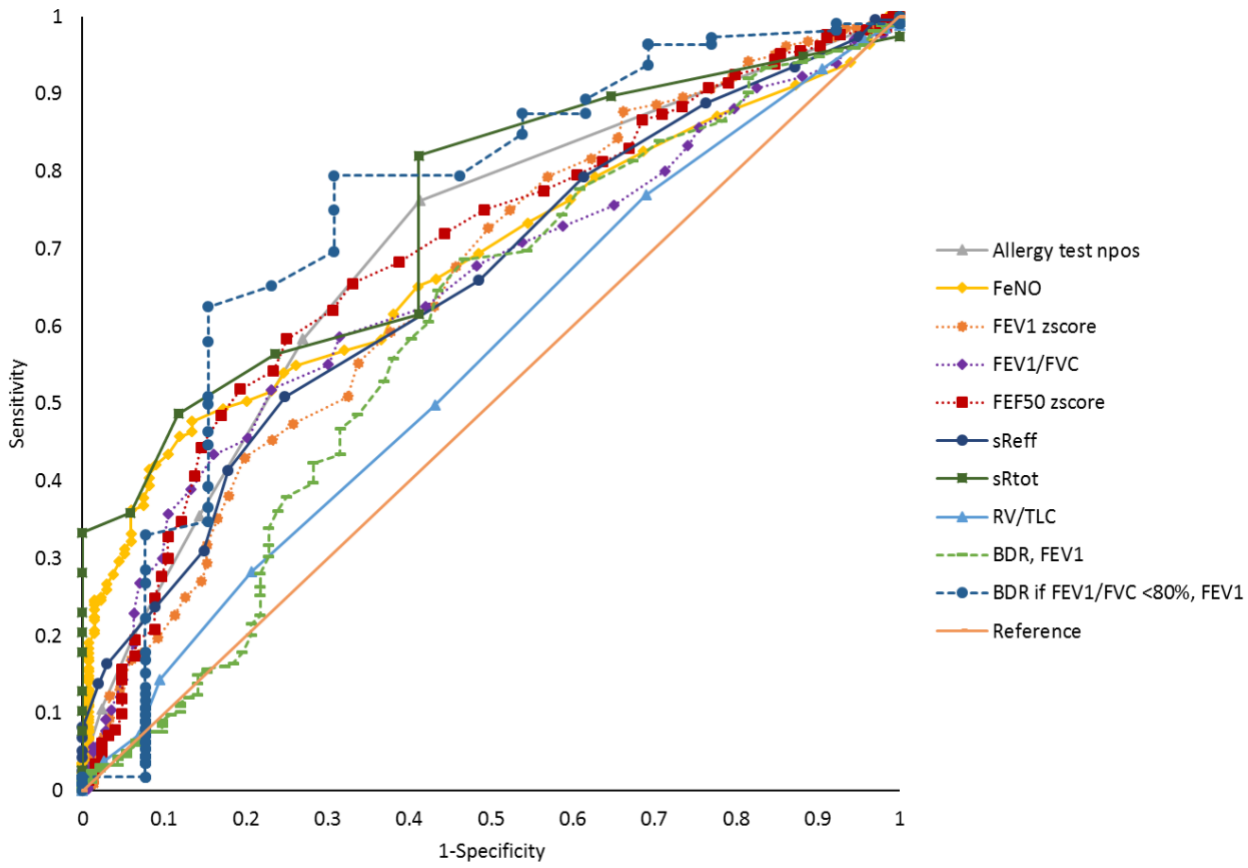


Figure 2 Receiver operating characteristics (ROC) curve of clinical tests to diagnose asthma.

Test	**Unit**
Allergy test npos	increase of 1 positive skin prick or specific IgE test
FeNO	increase of 1 parts per billion (ppb)
FEV1	increase of 0.1 z-score
FEV1/FVC	increase of 1%
FEF50	increase of 0.1 z-score
sReff	decrease of 0.01 kPa·s
sRtot	decrease of 0.01 kPa·s
RV/TLC	increase of 5%
Bronchodilator reversibility (BDR), FEV_1	increase of 1% in FEV_1

	Asthma +	Asthma -	Total
Algorithm +	81	10	91
Algorithm -	111	99	210
Total	192	99	301

Figure 3 Diagnostic accuracy of the algorithm proposed by the GINA guideline

Numbers in black: number of patients at this step. Numbers in red: number of patients with doctor diagnosed asthma at this step. Algorithm +: treat for asthma. Algorithm -: consider alternative diagnosis. FEV1: forced expiratory volume in 1 second. FVC: forced vital capacity. BDR: bronchodilator reversibility

*168 patients would need to repeat the spirometry and bronchodilator reversibility measurement or bronchial provocation test during another visit.

Supplementary material

Table S1. Diagnoses in children with suspected asthma after visiting the clinic* N=514

	Diagnoses n (%)	
Definite diagnoses		
Asthma	259	(50)
Cough not due to asthma[1]	71	(14)
Exercise limitation not due to asthma[2]	56	(11)
Allergic rhinoconjunctivitis	11	(2)
Non-CF bronchiectasis	1	(<1)
Probable diagnoses		
Asthma	97	(19)
Exercise limitation not due to asthma[2]	19	(4)

*Diagnosis at the first visit at a follow up visit within 3 months if the diagnosis was unclear

[1] Recurrent colds, post infectious cough, habitual cough, etc.

[2] Inducible laryngeal obstruction, dysfunctional breathing, functional symptoms, etc.

Table S2. Proportion of patients who performed diagnostic tests* per centre N=514

	Centres						Total
	A	B	C	D	E	F	
	N=60	N=196	N=83	N=75	N=15	N=85	N=514
Diagnostic tests	n%	n%	n%	n%	n%	n%	n%
Any allergy test[1]	60 (100)	182 (93)	74 (89)	55 (73)	15 (100)	81 (95)	467 (91)
FeNO	60 (100)	185 (94)	81 (98)	75 (100)	15 (100)	85 (100)	501 (97)
Spirometry	60 (100)	196 (100)	83 (100)	75 (100)	15 (100)	85 (100)	514 (100)
Body plethysmography	59 (98)	171 (87)	52 (63)	61 (81)	10 (67)	79 (93)	432 (84)
Bronchodilator reversibility	29 (48)	145 (74)	61 (73)	68 (91)	15 (100)	63 (74)	381 (74)
Any bronchial provocation test	13 (22)	86 (44)	23 (28)	23 (31)	7 (47)	58 (68)	210 (41)
Methacholine	7 (12)	71 (36)	2 (2)	1 (1)	-	48 (56)	129 (25)
Exercise	7 (12)	18 (9)	21 (25)	22 (29)	5 (33)	12 (14)	85 (17)
Mannitol	-	-	-	-	3 (20)	-	3 (1)

*At the first visit or at a follow up visit within 3 months

[1]Allergy test either done as above or results from <6 months ago reported in referral letter.

Table S3. Diagnostic test results in patients with and without asthma N=514

	Asthma					
	Definite asthma N=259		**Probable asthma** N=97		**Other diagnosis** N=158	
Diagnostic tests	median (IQR)		median (IQR)		median (IQR)	
Any allergy test, n(%)	248	(96)	93	(95)	124	(79)
≥1 positive test n(%)	201	(74)	59	(61)	52	(36)
Number of positive tests[1]	2	(1-3)	1	(0-3)	0	(0-2)
FeNO, n(%)	227	(88)	77	(80)	134	(84)
Parts per billion	25	(12-50)	14	(8-28)	11	(7-18)
Spirometry, n(%)	253	(98)	90	(93)	152	(96)
FEV1, z-scores	-0.5	(-1.4-0.1)	-0.2	(-1.0-0.5)	0.1	(-0.6-0.8)
FEV1/FVC	82	(76-88)	90	(84-95)	88	(84-93)
FEF50, z-score	-0.8	(-1.4--0.1)	0.1	(-0.7-0.9)	0.2	(-0.4-0.9)
Bodyplethysmography, n(%)	197	(76)	73	(75)	117	(74)
sReff, kPa·s[2]	1.0	(0.8-1.3)	0.8	(0.7-0.9)	0.8	(0.6-0.9)
sRtot, kPa·s[3]	1.3	(1.1-1.8)	1.4	(1.3-1.5)	1.1	(1.0-1.3)
RV/TLC	30	(25-36)	30	(26-36)	29	(24-34)
Bronchodilator reversibility, n(%)	215	(83)	59	(60)	92	(59)
Increase in FEV1 in %	13	(6-25)	10	(1-26)	6	(2-18)

[1]Defined as wheal size ≥3mm for mites, cat, grass, birch, mugwort and alternaria skin prick test and as ≥0.35kU/L for specific IgE test

[2] Reported by 4 centres

[3] Reported by 2 centres

Table S4. Diagnostic accuracy of diagnostic tests to diagnose asthma N=514 (sensitivity analysis: where children with "probable asthma" are subsumed to those with "no asthma" instead of those with "no asthma")

	A+T+ n	**A-T+** n	**A+T-** n	**A-T-** n	**Sens** % (95%CI)	**Spec** % (95%CI)	**PPV** % (95%CI)	**NPV** % (95%CI)	**YI**	**AUC**
Clinical tests										
Allergy test[1]										0.67
≥1 positive test*	201	111	47	108	81 (76-86)	49 (43-56)	64 (59-70)	70 (62-77)	0.30	
≥2 positive tests	154	79	94	140	62 (56-68)	64 (57-70)	66 (60-72)	60 (53-66)	0.26	
FeNO										0.70
≥20ppb	130	58	97	153	58 (51-64)	73 (66-78)	69 (62-76)	61 (55-67)	0.30	
≥21ppb	127	53	100	158	56 (49-63)	75 (68-81)	71 (63-77)	61 (55-67)	0.31	
≥23ppb*	119	44	108	167	52 (46-59)	79 (73-84)	73 (66-80)	61 (55-67)	0.32	
≥25ppb	114	41	113	170	50 (44-57)	81 (75-86)	74 (66-80)	60 (54-66)	0.31	
Spirometry										
FEV1										0.64
≤-0.2 z-score*	159	102	94	139	63 (57-69)	58 (51-64)	61 (55-67)	60 (53-66)	0.21	
≤-0.7 z-score	111	67	142	174	44 (38-50)	72 (66-78)	62 (55-69)	55 (49-61)	0.16	
≤-1.0 z-score	87	45	166	196	34 (29-41)	81 (76-86)	66 (57-74)	54 (49-59)	0.16	
FEV1/FVC										0.72
<80%	107	28	141	203	43 (37-50)	88 (83-92)	79 (71-86)	59 (54-64)	0.31	
<84%*	153	54	95	177	62 (55-68)	77 (71-82)	74 (67-80)	65 (59-71)	0.38	
<90%	202	127	46	104	81 (76-86)	45 (38-52)	61 (56-67)	69 (61-77)	0.26	
FEF50										0.73
≤-0.5 z-score*	131	46	87	155	60 (53-67)	77 (71-83)	74 (67-80)	64 (58-70)	0.37	
≤-1.0 z-score	83	27	135	174	38 (32-45)	87 (81-91)	75 (66-83)	56 (51-62)	0.25	
Bodyplethysm.										
sReff										0.68
≥0.9 kPa·s*	98	45	68	122	59 (51-67)	73 (66-80)	69 (60-76)	64 (57-71)	0.32	
≥1.0 kPa·s	81	33	85	134	49 (41-57)	80 (73-86)	71 (62-79)	61 (54-68)	0.29	
sRtot										0.64
≥1.0 kPa·s	28	18	4	6	88 (71-96)	25 (10-47)	61 (45-75)	60 (26-88)	0.13	
≥1.1 kPa·s *	25	14	7	10	78 (60-91)	42 (22-63)	64 (47-79)	59 (33-82)	0.20	
RV/TLC										0.53
≥25%*	150	134	45	52	77 (70-83)	28 (22-35)	53 (47-59)	54 (43-64)	0.05	
Bronchodilator rev.										0.60
FEV_1										
≥7% increase*	156	75	59	76	73 (66-78)	50 (42-59)	68 (61-74)	56 (47-65)	0.23	
≥10% increase	131	66	84	85	61 (54-67)	56 (48-64)	66 (59-73)	50 (43-58)	0.17	
≥12% increase	117	62	98	89	54 (48-61)	59 (51-67)	65 (58-72)	48 (40-55)	0.13	
Bronchodilator rev. if FEV1/FVC <80%[4]										0.70
≥7% increase*	83	10	18	14	82 (73-89)	58 (37-78)	89 (81-95)	44 (26-62)	0.41	
≥10% increase	68	8	33	16	67 (57-76)	67 (45-84)	89 (80-95)	33 (20-48)	0.34	
≥12% increase	60	7	41	17	59 (49-69)	71 (49-87)	90 (80-96)	29 (18-43)	0.30	

A+T+: children with asthma diagnosis and positive test result, A-T+: children without asthma diagnosis but positive test result, A+T-: children with asthma diagnosis but negative test result, A-T-: children without asthma and negative test result, Sens: sensitivity, Spec: specificity, PPV: positive predictive value, NPV: negative predictive value, YI: Youden's-Index: Sensitivity + Specificity -1, AUC: area under the curve, FeNO: fractional exhaled nitric oxide, ppb: parts per billion, FEV_1: forced expiratory volume in 1 second, FVC: forced vital capacity, FEF50: forced expiratory flow at 50% of FVC, sReff: specific effective airway resistance, sRtot: specific total airway resistance, RV: residual volume, TLC: total lung capacity, Bronchodilator rev.: bronchodilator reversibility

Displayed cut-offs chosen based on proposed cut-offs from previous publications

*Cut-off with maximum combined sensitivity and specificity (highest Youden's-Index)
[1] Number allergens for which the skin prick test is positive: wheal size ≥3 or the specific IgE test was positive: ≥0.35 kU/l.
[2] Reported by 4 centres
[3] Reported by 2 centres
[4] N= 126, cut-off chosen based on proposed cut-off from previous publications and guidelines

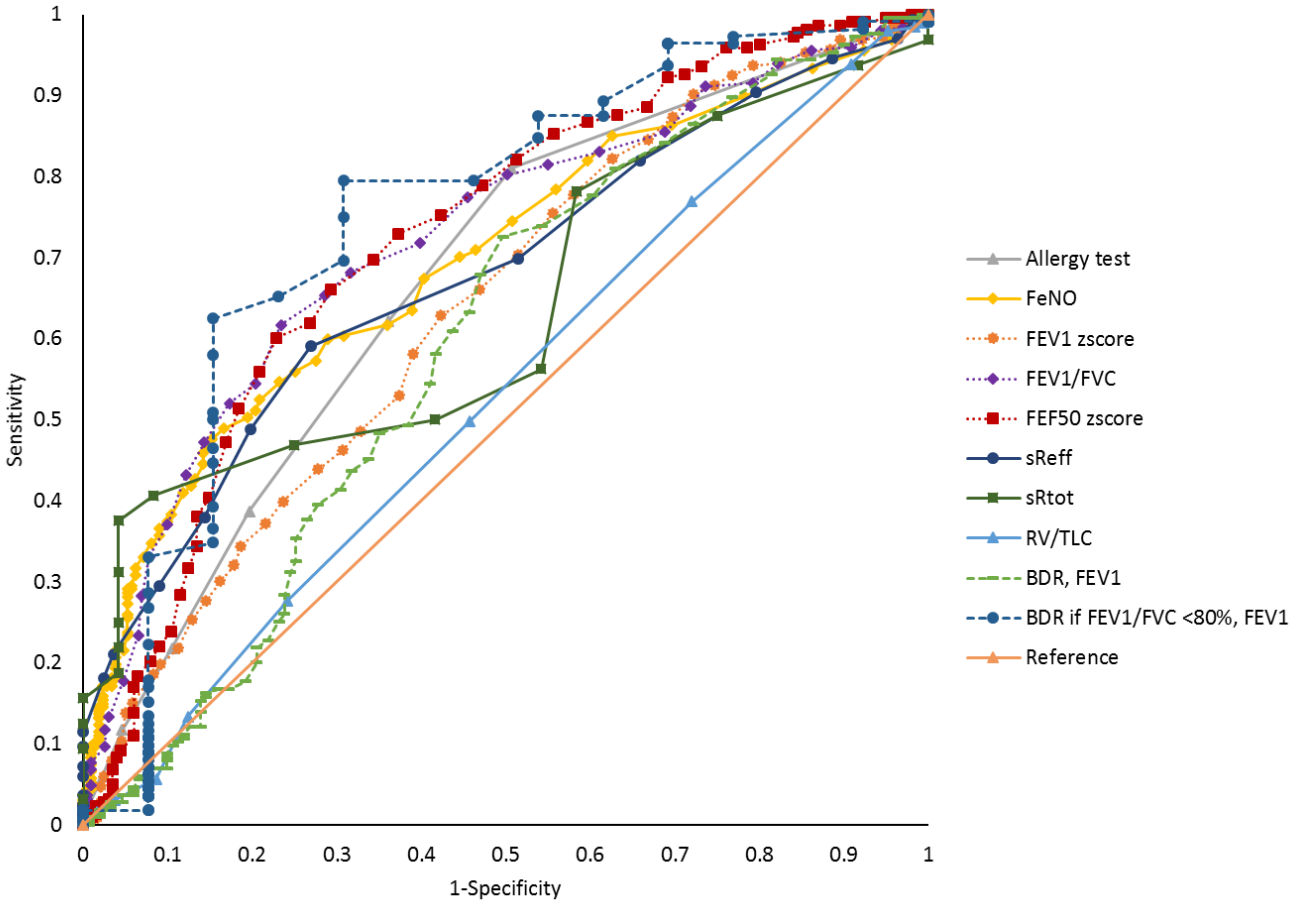


Figure S1 Receiver operating characteristics (ROC) curve of clinical tests to diagnose asthma. (Sensitivity analysis: "asthma" is defined as definite asthma and "not asthma" is defined as probable asthma or other diagnosis)

* Cut-off with maximum combined sensitivity and specificity

Test	**Unit**
Allergy test npos	increase of 1 positive skin prick or specific IgE test
FeNO	increase of 1 parts per billion (ppb)
FEV1	increase of 0.1 z-score
FEV1/FVC	increase of 1%
FEF50	increase of 0.1 z-score
sReff	decrease of 0.01 kPa·s
sRtot	decrease of 0.01 kPa·s
RV/TLC	increase of 5%
Bronchodilator reversibility (BDR), FEV_1	increase of 1% in FEV_1

Table S5. Diagnostic test results in steroid naive patients with and without asthma N=271

	Asthma					
	Definite asthma N=108		**Probable asthma** N=48		**Other diagnosis** N=115	
Diagnostic tests	median (IQR)		median (IQR)		median (IQR)	
Any allergy test, n(%)	101	(94)	46	(96)	86	(75)
≥1 positive test n(%)	87	(81)	31	(65)	39	(34)
Number of positive tests[1]	2	(1-4)	1	(0-3)	0	(0-2)
FeNO, n(%)	90	(83)	42	(88)	98	(85)
Parts per billion	33	(13-54)	17	(8-31)	13	(7-20)
Spirometry, n(%)	102	(94)	46	(96)	112	(97)
FEV1, z-scores	-0.5	(-1.3-0.1)	-0.1	(-1.0-0.5)	0.1	(-0.6-0.9)
FEV1/FVC	82	(77-88)	89	(83-92)	89	(86-95)
FEF50, z-score	-0.7	(-1.4-0.1)	0.1	(-0.7-0.9)	0.2	(-0.2-0.9)
Bodyplethysmography, n(%)	80	(74)	36	(75)	88	(77)
sReff, kPa·s[2]	1.0	(0.7-1.3)	0.8	(0.7-0.9)	0.8	(0.6-0.9)
sRtot, kPa·s[3]	1.3	(1.1-2.3)	1.4	(1.3-1.5)	1.0	(1.0-1.2)
RV/TLC	28	(24-33)	29	(25-37)	30	(24-34)
Bronchodilator reversibility, n(%)	85	(79)	30	(63)	66	(57)
Increase in FEV1 in %	13	(6-25)	5	(2-19)	7	(2-28)

[1]Defined as wheal size ≥3mm for mites, cat, grass, birch, mugwort and alternaria skin prick test and as ≥0.35kU/L for specific IgE test

[2] Reported by 4 centres

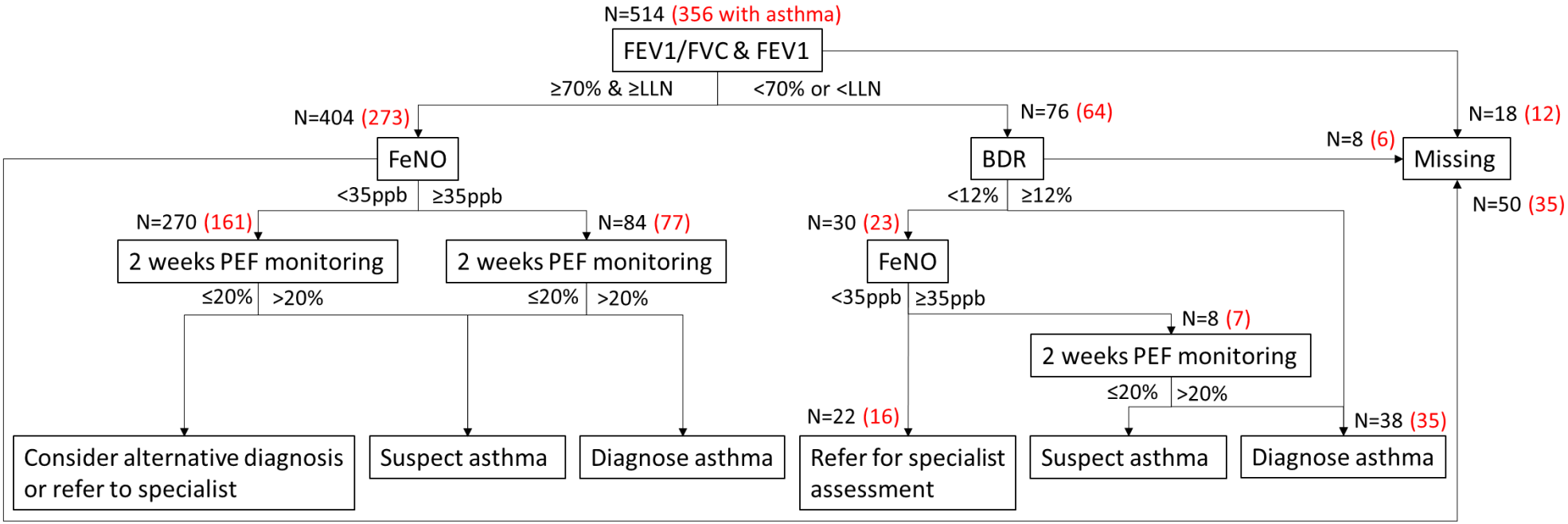
[3] Reported by 2 centres

Table S6. Diagnostic accuracy of diagnostic tests to diagnose asthma N=271 (sensitivity analysis: in steroid naïve children)

	A+T+ n	A-T+ n	A+T- n	A-T- n	Sens % (95%CI)	Spec % (95%CI)	PPV % (95%CI)	NPV % (95%CI)	YI	AUC
Clinical tests										
Allergy test[1]										0.71
≥1 positive test*	118	39	29	47	80 (73-86)	55 (44-65)	75 (68-82)	62 (50-73)	0.35	
≥2 positive tests	90	26	57	60	61 (53-69)	70 (59-79)	78 (69-85)	51 (42-61)	0.31	
FeNO										0.71
≥20ppb	75	25	57	73	57 (48-65)	74 (65-83)	75 (65-83)	56 (47-65)	0.31	
≥21ppb	72	23	60	75	55 (46-63)	77 (67-85)	76 (66-84)	56 (47-64)	0.31	
≥23ppb	69	17	63	81	52 (43-51)	83 (74-90)	80 (70-88)	56 (48-64)	0.35	
≥25ppb	66	15	66	83	50 (41-59)	85 (76-91)	81 (71-89)	56 (47-64)	0.35	
≥28ppb*	62	11	70	87	47 (38-56)	89 (81-94)	85 (75-92)	55 (47-63)	0.36	
Spirometry										
FEV1										0.65
≤-0.2 z-score	83	44	65	68	56 (48-64)	61 (51-70)	65 (56-74)	51 (42-60)	0.17	
≤-0.7 z-score*	60	23	88	89	41 (33-49)	79 (71-87)	72 (61-82)	50 (43-58)	0.20	
≤-1.0 z-score	42	17	106	95	28 (21-36)	85 (77-91)	71 (58-82)	47 (40-54)	0.13	
FEV1/FVC										0.69
<80%	47	6	100	99	32 (25-40)	94 (88-98)	87 (77-96)	50 (43-57)	0.26	
<86%*	88	26	59	79	60 (51-68)	75 (66-83)	77 (68-85)	57 (49-66)	0.35	
<90%	109	58	38	47	74 (66-81)	45 (35-56)	65 (58-72)	55 (44-66)	0.19	
FEF50										0.70
≤-0.3 z-score*	70	18	61	77	53 (45-62)	81 (72-88)	80 (70-87)	56 (47-64)	0.34	
≤-1.0 z-score	40	5	91	90	31 (23-39)	95 (88-98)	89 (76-96)	50 (42-57)	0.25	
Bodyplethysm.										
sReff										0.62
≥0.9 kPa·s*	47	20	54	55	47 (37-57)	73 (62-83)	70 (58-81)	50 (41-60)	0.20	
≥1.0 kPa·s	37	15	64	60	37 (27-47)	80 (69-88)	71 (57-83)	48 (39-58)	0.17	
sRtot										0.77
≥1.0 kPa·s	13	7	2	6	87 (60-98)	46 (19-75)	65 (41-85)	75 (35-97)	0.33	
≥1.5 kPa·s *	13	4	2	9	87 (60-98)	69 (39-91)	76 (50-93)	82 (48-98)	0.56	
RV/TLC										0.51
≥35%*	25	17	88	69	22 (15-31)	80 (70-88)	60 (43-74)	44 (36-52)	0.02	
Bronchodilator rev.										
FEV_1										0.56
≥7% increase*	75	32	40	34	65 (56-74)	52 (39-64)	70 (60-79)	46 (34-58)	0.17	
≥10% increase	66	27	49	39	57 (48-67)	59 (46-71)	71 (61-80)	44 (34-55)	0.16	
≥12% increase	59	25	56	41	51 (42-61)	62 (49-74)	70 (59-80)	42 (32-53)	0.13	
Bronchodilator rev. if FEV1/FVC <80%[4]										0.87
≥7% increase	37	1	8	4	82 (68-92)	80 (28-99)	97 (86-99)	33 (10-65)	0.62	
≥10% increase*	30	0	15	5	67 (51-80)	99 (48-99)	99 (88-99)	25 (9-49)	0.67	
≥12% increase	26	0	19	5	58 (42-72)	99 (48-99)	99 (87-99)	21 (7-42)	0.58	

A+T+: children with asthma diagnosis and positive test result, A-T+: children without asthma diagnosis but positive test result, A+T-: children with asthma diagnosis but negative test result, A-T-: children without asthma and negative test result, Sens: sensitivity, Spec: specificity, PPV: positive predictive value, NPV: negative predictive value, YI: Youden's-Index: Sensitivity + Specificity -1, AUC: area under the curve, FeNO: fractional exhaled nitric oxide, ppb: parts per billion, FEV_1: forced expiratory volume in 1 second, FVC: forced vital capacity, FEF50: forced expiratory flow at 50% of FVC, sReff: specific effective airway resistance, sRtot: specific total airway resistance, RV: residual volume, TLC: total lung capacity,

Bronchodilator rev.: bronchodilator reversibility
Displayed cut-offs chosen based on proposed cut-offs from previous publications
*Cut-off with maximum combined sensitivity and specificity (highest Youden's-Index)
[1] Number allergens for which the skin prick test is positive: wheal size ≥3 or the specific IgE test was positive: ≥0.35 kU/l.
[2] Reported by 4 centres
[3] Reported by 2 centres
[4] N= 51, cut-off chosen based on proposed cut-off from previous publications and guidelines



	Asthma +	Asthma -	Total
Algorithm +	35	3	38
Algorithm -	16	6	22
Total	51	9	60

Sens = 35/51 =0.69
Spec = 6/9 =0.67
PPV = 35/38 =0.92
NPV = 6/22=0.27

Figure S2 Diagnostic accuracy of the diagnostic algorithm proposed by the NICE guideline

Numbers in black: number of patients at this step. Numbers in red: number of patients with doctor diagnosed asthma at this step. Algorithm +: diagnose asthma or suspect asthma. Algorithm - : refer for specialist assessment or consider alternative diagnosis. FEV1: forced expiratory volume in 1 second. FVC: forced vital capacity. BDR: bronchodilator reversibility. FeNO: fractional exhaled nitric oxide. PEF: peak expiratory flow.

*362 patients would need 2 weeks PEF monitoring.