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

Carmen C.M. de Jong [1], Eva S.L. 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] and Claudia E. Kuehni [1,4]

Affiliations: [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.

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

@ERSpublications
Asthma diagnosis does not seem straightforward, even for experienced pulmonologists, and this highlights the need for new evidence-based guidance https://bit.ly/3dagdgv

Cite this article as: de Jong CCM, Pedersen ESL, Mozun R, *et al*. Diagnosis of asthma in children: findings from the Swiss Paediatric Airway Cohort. *Eur Respir J* 2020; 56: 2000132 [https://doi.org/10.1183/13993003.00132-2020].

ABSTRACT

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 the Global Initiative for Asthma (GINA) and the National Institute for Health and Care Excellence (NICE)) in the diagnosis of asthma in school-aged children.

Methods: We studied children aged 5–17 years who were referred consecutively to pulmonary outpatient clinics for evaluation of suspected asthma. Symptoms were assessed by parental questionnaire. The investigations included specific IgE measurement or skin prick tests, measurement of exhaled nitric oxide fraction (F_{eNO}), spirometry, body plethysmography and bronchodilator reversibility (BDR). 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 predictive value (PPV), negative predictive value (NPV) and area under the curve (AUC).

Results: Among 514 participants, 357 (70%) were diagnosed with asthma. The combined sensitivity and specificity was highest for any wheeze (sensitivity=75%, specificity=65%), dyspnoea (sensitivity=56%, specificity=76%) and wheeze triggered by colds (sensitivity=58%, specificity=78%) or by exercise (sensitivity=55%, specificity=74%). Of the diagnostic tests, the AUC was highest for specific total airway resistance (sR_{tot}; AUC=0.73) and lowest for the residual volume (RV)/total lung capacity (TLC) ratio (AUC=0.56). The NICE algorithm had sensitivity=69% and specificity=67%, whereas the GINA algorithm had sensitivity=42% and specificity=90%.

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

This article has an editorial commentary: https://doi.org/10.1183/13993003.02687-2020

This article has supplementary material available from erj.ersjournals.com

Received: 22 Jan 2020 | Accepted after revision: 26 May 2020

Introduction

Diagnosing asthma in children remains a challenge because respiratory symptoms, such as wheeze and cough, are not specific and vary over time, while asthma lacks an effective stand-alone diagnostic test. Clinically, physicians diagnose asthma based on a combination of symptoms, physical examinations and diagnostic tests [1–3]. Among available tests, spirometry and body plethysmography, in combination with bronchodilator reversibility (BDR) testing, demonstrate reversible airway obstruction. Bronchial provocation tests measure bronchial hyper-responsiveness (BHR), exhaled nitric oxide fraction (F_{eNO}) 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-treatment and over-treatment [3–5]. Systematic literature reviews on asthma diagnosis in children, performed as part of the ongoing task force of the European Respiratory Society (ERS), 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, F_{eNO}, and BHR testing by methacholine or exercise [7]. However, because the study of the NICE algorithm was based on the general population, while 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 (those of 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 from 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 who were 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 F_{eNO}, spirometry, body plethysmography and BDR testing. When indicated, children also underwent BHR testing (by methacholine, exercise, or mannitol), either during the same visit or at a follow-up visit within 3 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 (*i.e.* 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 Cohorts questionnaires [9, 10]. Dyspnoea was assessed using the question "Hatte Ihr Kind in den letzten 12 Monaten Atemnot?" [In the past 12 months did your child have episodes of breathlessness?]. All diagnostic tests were performed according to published guidelines [11, 12]. Children were asked to withhold medication before the visit: short-acting β_2-agonists (SABAs) for 8 h, inhaled

corticosteroids (ICS), leukotriene antagonists and long acting β_2-agonists (LABAs) for 24 h, and antihistamines and sodium cromoglycate for >72 h.

Skin prick tests and specific immunoglobulin E (IgE) measurements were used to measure atopy. Skin prick tests were performed using histamine, saline, birch, grass, mugwort, alternaria, cat hair, and house dust mites (*Dermatophagoides pteronyssinus* and *Dermatophagoides farinae*) [13]. A wheal size of ⩾3 mm was considered positive where the positive control (histamine) had a wheal size of ⩾3 mm and the negative control (saline) had a wheal size of <3 mm. Specific IgE levels for birch, grass, mugwort, alternaria, cat hair and house dust mites were measured in serum samples using the fluorescence enzyme immunoassay/immunocap (ThermoFisher Scientific, Uppsala, Sweden). IgE levels were considered positive at the detection threshold (⩾0.35 $kU \cdot L^{-1}$).

F_{eNO} was measured in doublets before spirometry using a portable multi-gas analyser (NIOX MINO, Aerocrine AB, Solna, Sweden) in St. Gallen and an F_{eNO} analyzer (ANALYZER CLD 88 sp, Eco Medics AG, Duernten, Switzerland) 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. A Jaeger masterscope (Erich Jaeger GmbH, Würzburg, Germany) using JLAB software version 4.34 was used in Basel and St. Gallen, while a MasterScreen Pneumo spirometer (Vyaire Medical, Chicago, IL, USA) using Sentrysuite software was used in Bern, Zurich, Aarau and Lucerne. Spirometry was performed in triplicate by experienced lung function technicians, who performed quality control during the process and recorded the best measurement. The flow–volume curve was checked by the responsible paediatric pulmonologist. Spirometry results (forced expiratory volume in 1 s (FEV_1) and forced expiratory flow at 50% of FVC ($FEF_{50\%}$)) are expressed as z-scores based on the Global Lung Function Initiative (GLI) 2012 reference standards [15] and as proportions (the FEV_1/forced

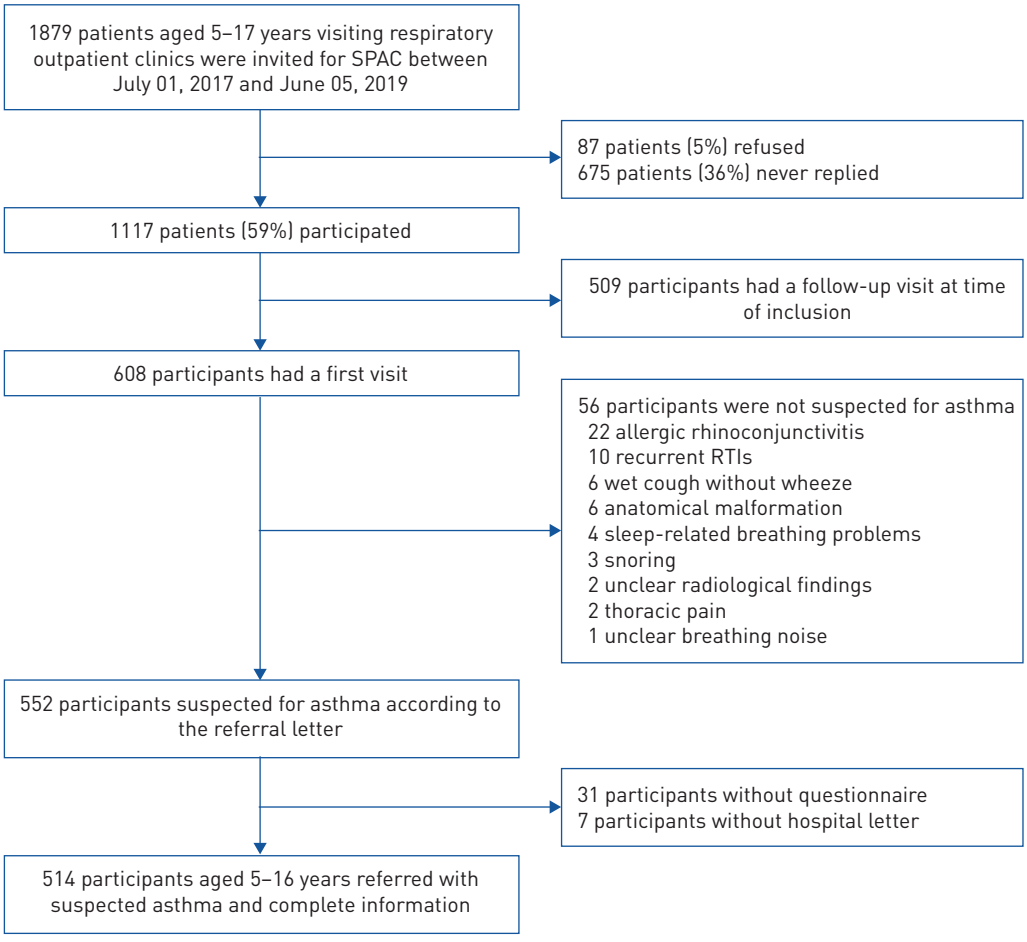


FIGURE 1 Flow diagram illustrating the inclusion of study participants. SPAC: Swiss Paediatric Airway Cohort; RTI: respiratory tract infection.

vital capacity (FVC) ratio). The results of body plethysmography are expressed in kPa·s for specific effective airway resistance (sR_{eff}) and specific total airway resistance (sR_{tot}) and as a proportion (the residual volume (RV)/total lung capacity (TLC) ratio).

BDR tests were performed for FEV_1 ⩽90%, $FEF_{75\%}$ ⩽67%, $FEF_{50\%}$ ⩽67%, or $FEF_{25\%}$ ⩽67% (Lucerne, Zurich and Aarau); for sR_{eff} >180% or $FEF_{50\%}$ <80% (Bern); or for all patients (St. Gallen and Basel). All centres gave salbutamol (400 µg; Ventolin pressurised metered-dose inhaler *via* spacer) to assess BDR. Spirometry was repeated in duplicate after 10 min (Lucerne and Basel), 15 min (St. Gallen, Aarau and Bern) or 20 min (Zurich). BDR was calculated by the following equation: (post-bronchodilator FEV_1–pre-bronchodilator FEV_1)·100/pre-bronchodilator FEV_1.

Statistical analysis

We calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), Youden's Index (sensitivity+specificity–1), area under the curve (AUC) and 95% confidence interval (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 value. We assessed the diagnostic accuracy of tests if they were performed in at least 70% of the children. To assess the diagnostic accuracy of BDR we performed a subanalysis in children with obstructive lung function (FEV_1/FVC <80%) [1, 16]. We performed an initial sensitivity analysis in which we classified children with "probable asthma" as having "no asthma". We performed a second sensitivity analysis using only steroid naïve children. We applied the asthma diagnosis algorithms from GINA and NICE to assess how they would have performed in a clinical setting and calculated sensitivity, specificity, PPV and NPV. We used STATA version 15 (StataCorp LLC, College Station, TX, USA) 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 reported wheeze and/or cough. Asthma had been diagnosed in 356 (69%) of the participants. Table 1 presents the full characteristics of the participants. Frequent other diagnoses included exercise limitation (56 children, 11%) and cough (71 children, 14%), both not due to asthma (supplementary table S1).

TABLE 1 Characteristics of the study participants (n=514)

Characteristic	Result
Age years	9 (7–12)
Male sex	294 (57)
BMI z-score	0.3 (−0.4 to −1.1)
Respiratory symptoms in the last 12 months	
Any wheeze	317 (62)
More than three 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)
SABA alone	152 (30)
ICS±SABA	114 (22)
ICS+LABA	129 (25)

Data are presented as n (%) or median (IQR). BMI: body mass index; SABA: short-acting β_2-agonist; ICS: inhaled corticosteroids; LABA: long-acting β_2-agonist; IQR: interquartile range.

TABLE 2 Diagnostic accuracy of respiratory symptoms to diagnose asthma (n=514)

Respiratory symptoms in the past 12 months	A+S+	A−S+	A+S−	A−S−	Sensitivity %	Specificity %	PPV %	NPV %	Youden's index[#]
Any wheeze	262	55	90	100	74 (70–79)	65 (56–72)	83 (78–87)	53 (45–60)	0.39
More than three 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
Wheeze triggered by house dust	76	5	259	150	23 (18–28)	97 (93–99)	94 (86–98)	37 (32–42)	0.19
Wheeze triggered by 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

Data are presented as n or n (95% CI). A+S+: children with an asthma diagnosis and reported symptoms; A−S+: children without an asthma diagnosis but with reported symptoms; A+S−: children with an asthma diagnosis but without reported symptoms; A−S−: children without an asthma diagnosis and without reported symptoms. PPV: positive predictive value; NPV: negative predictive value; CI: confidence interval. [#]: sensitivity+specificity−1.

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 value (0.39) for asthma (table 2). Specificity was highest for frequent attacks (more than three per 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).

Diagnostic accuracy of tests to diagnose asthma

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

The cut-off values with the best diagnostic accuracy were one or more positive test for the allergy tests, ⩾23 ppb for F_{eNO}, ⩽−0.7 z-score for FEV_1, <84% for FEV_1/FVC, ⩽−0.3 z-score for $FEF_{50\%}$, ⩾0.9 kPa·s for sR_{eff}, ⩾1.1 kPa·s for sR_{tot}, ⩾25% for RV/TLC and ⩾7% increase in FEV_1 for the BDR test (table 3). The diagnostic accuracy (AUC) was highest for sR_{tot} (0.74), allergy tests (0.70), $FEF_{50\%}$ (0.69) and F_{eNO} (0.68). The accuracy was lowest for RV/TLC (0.56) and the BDR test (0.60). However, BDR had the highest accuracy (0.75) when we analysed only data from children with FEV_1/FVC <80% (figure 2).

Sensitivity analysis

In the first 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 FEV_1 (from ⩽−0.7 z-score to ⩽−0.2 z-score) and $FEF_{50\%}$ (from ⩽−0.3 z-score to ⩽−0.5 z-score). The diagnostic accuracy (AUC) remained highest for $FEF_{50\%}$ (0.73) and F_{eNO} (0.70) and increased for FEV_1/FVC (from 0.65 to 0.72), and sR_{eff} (from 0.66 to 0.68). The accuracy remained lowest for RV/TLC (0.53) and the BDR test (0.60) (supplementary table S4 and supplementary figure S1).

In the second sensitivity analysis, which included only ICS naïve children, asthma was diagnosed in 156 of the 271 children (58%) (supplementary table S5). The cut-offs with highest combined sensitivity and specificity changed for F_{eNO} (from ⩽23 ppb to ⩽28 ppb), FEV_1/FVC (from 84% to 86%), sR_{tot} (from 1.1 kPa·s to 1.5 kPa·s) and RV/TLC (from ⩾25% to ⩾35%). The diagnostic accuracy remained highest for sR_{tot} (0.77), the allergy tests (0.71), $FEF_{50\%}$ (0.70) and F_{eNO} (0.71), as in the main analysis. The accuracy was still lowest for RV/TLC (0.51) and the BDR test (0.56) (supplementary 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 were

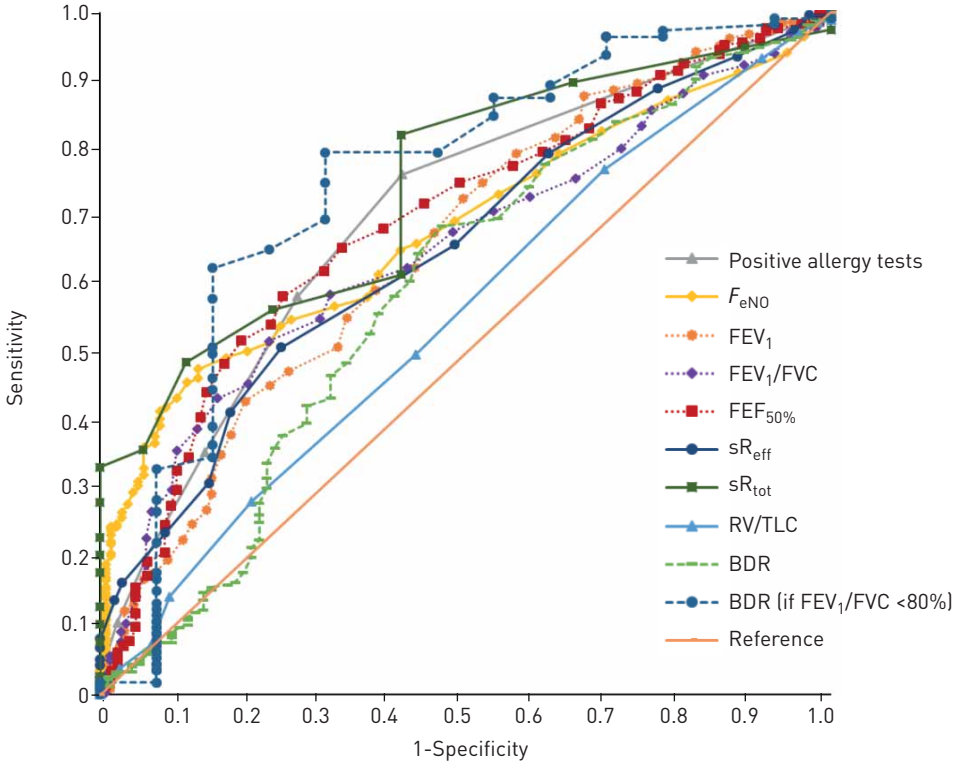
TABLE 3 Diagnostic accuracy of diagnostic tests to diagnose asthma (n=514)

Clinical test	A+T+	A−T+	A+T−	A−T−	Sensitivity %	Specificity %	PPV %	NPV %	Youden's index[#]	AUC
Positive allergy test[¶]										0.70
⩾1[+]	260	52	81	74	76 (71–81)	59 (50–67)	83 (79–87)	48 (40–56)	0.35	
⩾2	199	34	142	92	58 (53–64)	73 (64–81)	85 (80–90)	39 (33–46)	0.31	
F_{eNO} ppb										0.68
⩾20	157	31	147	103	52 (46–57)	77 (69–84)	84 (77–89)	41 (35–48)	0.29	
⩾21	153	27	151	107	50 (45–56)	80 (72–86)	85 (79–90)	41 (35–48)	0.30	
⩾23[+]	145	18	159	116	48 (42–53)	87 (80–92)	89 (83–93)	42 (36–48)	0.34	
⩾25	139	16	165	118	46 (40–52)	88 (81–93)	90 (84–94)	42 (36–48)	0.34	
Spirometry										
FEV_1 z-score										0.66
⩽ −0.7[+]	148	30	195	121	43 (38–49)	80 (73–86)	83 (77–88)	38 (33–44)	0.23	
⩽ −1.0	109	23	234	128	32 (27–37)	85 (78–90)	83 (75–89)	35 (30–41)	0.17	
FEV_1/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	
$FEF_{50\%}$ z-score										0.69
⩽ −0.3[+]	171	31	122	93	58 (52–64)	75 (66–82)	85 (79–89)	43 (37–50)	0.33	
⩽ −1.0	96	13	197	111	33 (27–38)	90 (83–94)	88 (80–93)	36 (31–42)	0.22	
Body plethysmography										
sR_{eff}[§] kPa·s·L^{-1}										0.66
⩾0.9[+]	118	25	114	76	51 (44–57)	75 (66–83)	83 (75–88)	40 (33–47)	0.26	
⩾1.0	96	18	136	83	41 (35–48)	82 (73–89)	84 (76–90)	38 (31–45)	0.24	
sR_{tot}[f] kPa·s·L^{-1}										0.74
⩾1.0	35	11	4	6	90 (76–97)	35 (14–62)	76 (61–87)	60 (26–88)	0.25	
⩾1.1[+]	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	
BDR % increase in FEV_1										0.60
⩾7[+]	188	43	86	49	69 (63–74)	53 (43–64)	81 (76–86)	36 (28–45)	0.22	
⩾10	160	37	114	55	58 (52–64)	60 (49–70)	81 (75–86)	33 (26–40)	0.18	
⩾12	145	34	129	58	53 (47–59)	63 (52–73)	81 (74–86)	31 (24–38)	0.16	
BDR (if FEV_1/FVC <80%)[##] % increase in FEV_1										0.75
⩾7[+]	89	4	23	9	79 (71–87)	69 (39–91)	96 (89–99)	28 (14–47)	0.49	
⩾10	73	3	39	10	65 (56–74)	77 (46–95)	96 (89–99)	20 (10–34)	0.42	
⩾12	65	2	47	11	58 (48–67)	85 (55–98)	97 (90–99)	19 (10–31)	0.43	

Data are presented as n or n (95% CI). The cut-offs displayed were chosen based on the proposed cut-offs from previous publications. A+T+: children with an asthma diagnosis and positive test results; A−T+: children without an asthma diagnosis but with positive test results; A+T−: children with an asthma diagnosis and negative test results; A−T−: children without an asthma diagnosis but with negative test results; PPV: positive predictive value; NPV: negative predictive value; AUC: area under the curve; F_{eNO}: exhaled nitric oxide fraction; FEV_1: forced expiratory volume in 1 s; FVC: forced vital capacity; $FEF_{50\%}$: forced expiratory flow at 50% of FVC; sR_{eff}: specific effective airway resistance; sR_{tot}: specific total airway resistance; RV: residual volume; TLC: total lung capacity; BDR: bronchodilator reversibility; CI: confidence interval. [#]: sensitivity+specificity−1; [¶]: number of allergens for which the skin prick test is positive (wheal size ⩾3 or the specific IgE test was positive (⩾0.35 kU·L^{-1})); [+]: cut-off with maximum combined sensitivity and specificity (highest Youden's index value); [§]: reported by four centres; [f]: reported by two centres; [##]: cut-off chosen based on proposed cut-off from previous publications and guidelines (n=126).

diagnosed with asthma by our paediatric pulmonologist (PPV 89%). Of the 210 children we could pass through the algorithm to the step "consider alternative diagnosis", 111 were diagnosed with asthma and 99 were not (NPV 47%). The sensitivity of the algorithm was 42% and the specificity was 90%. The GINA algorithm would have been inconclusive in 168 children, as they ended up 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 able to pass only 38 children through to the step "diagnose asthma" (supplemental figure S2). Of these 38 children, 35 were diagnosed with asthma by the paediatric pulmonologists (PPV 92%). Of the 22 children whom we could pass through to the step "refer for specialist assessment", 16 were diagnosed with asthma and six were not (NPV 27%). The sensitivity was thus 69% and the specificity 67%. However, 362 children



Test	Unit
Positive allergy tests	Increase of one positive skin prick or specific IgE test
F_{eNO}	Increase of 1 ppb
FEV_1	Increase of 0.1 z-score
FEV_1/FVC	Increase of 1%
$FEF_{50\%}$	Increase of 0.1 z-score
sR_{eff}	Decrease of 0.01 kPa·s
sR_{tot}	Decrease of 0.01 kPa·s
RV/TLC	Increase of 5%
BDR	Increase of 1% in FEV_1

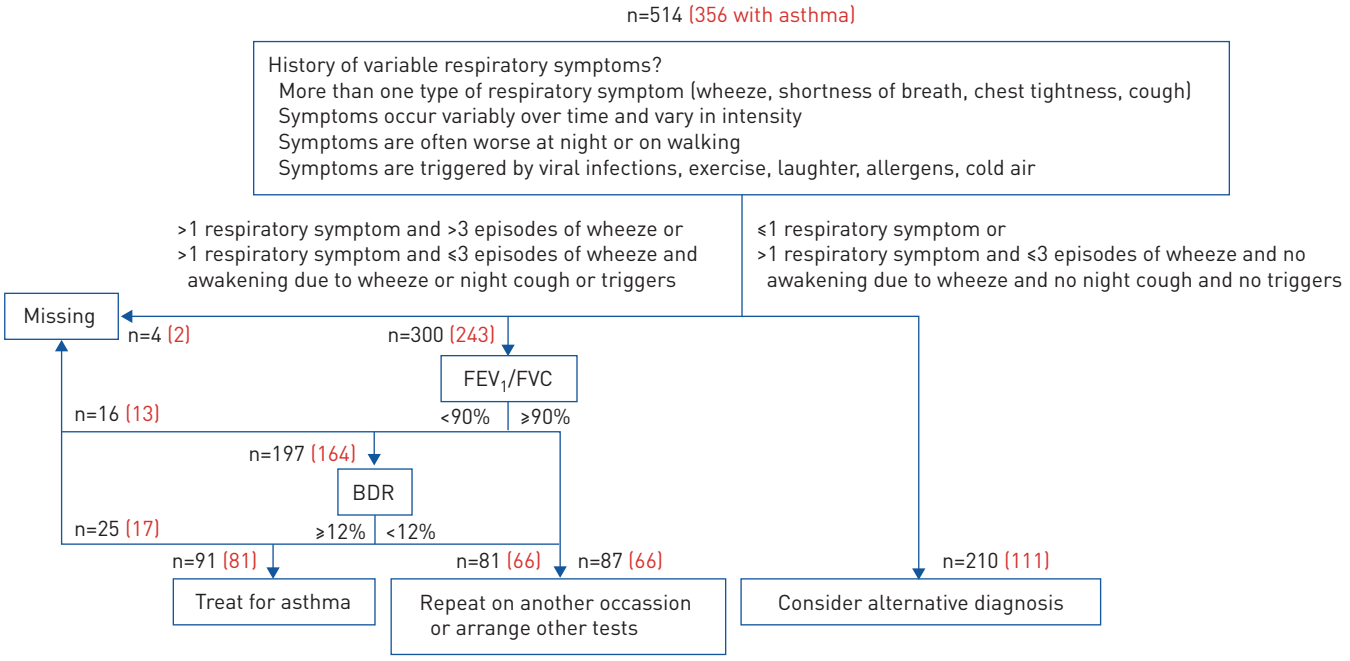
FIGURE 2 Receiver operating characteristic (ROC) curve of clinical tests to diagnose asthma. F_{eNO}: exhaled nitric oxide fraction; FEV_1: forced expiratory volume in 1 s; FVC: forced vital capacity; $FEF_{50\%}$: forced expiratory flow at 50% of FVC; sR_{eff}: specific effective airway resistance; sR_{tot}: specific total airway resistance; RV: residual volume; TLC: total lung capacity; BDR: bronchodilator reversibility.

(83%) were stuck at "2 weeks of PEF monitoring" (where PEF: peak expiratory flow). From this step onwards we could not apply the NICE diagnostic algorithm.

Discussion

This study found that the commonly available tests, used either 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=74%, specificity=65%) and the diagnostic accuracy (AUC) was highest for sR_{tot} (0.74), a positive allergy test (0.70), $FEF_{50\%}$ (0.69), F_{eNO} (0.68) and the BDR test in children with FEV_1/FVC <80% (0.75). Diagnostic accuracy was lowest for RV/TLC (0.56). The NICE algorithm relied too much upon 2 weeks' PEF monitoring, which was indicated in 83% of children according to the algorithm and 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 referred consecutively for evaluation of possible asthma [7, 17–19]. They all found that reported wheeze



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

Sensitivity=81/192=0.42
Specificity=99/110=0.90
PPV=81/91=0.89
NPV=99/210=0.47

FIGURE 3 Diagnostic accuracy of the algorithm proposed by the Global Initiative for Asthma (GINA) guideline. Numbers in black are the number of patients at this step. Red numbers in parentheses are the number of patients with doctor diagnosed asthma at this step. algorithm+: treat for asthma; algorithm−: consider alternative diagnosis; FEV_1: forced expiratory volume in 1 s; FVC: forced vital capacity; BDR: bronchodilator reversibility; PPV: positive predictive value; NPV: negative predictive value.

was sensitive (range 75–86%) but not specific (range 64–73%) and that frequent wheeze and awakening due to dyspnoea were specific (range 84–90%) but not sensitive (range 33–54%), which is in line with our findings. Another way to phrase this would be that if frequent wheeze (for example) 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 highly [7]. Also in that study and as reported by Woo *et al.* [20], we confirmed that a positive skin prick test was sensitive (range 68–90%) but not specific (range 32–40%). The AUC for F_{eNO} in our study (0.68) was lower than that in our previous study (0.79) [7] as well as that in a Korean study (0.80) by Eom *et al.* [21]. In steroid naïve children we expected the diagnostic accuracy for F_{eNO} to be higher, as treatment with ICS reduces F_{eNO} levels. This was not the case and could be due to two reasons: 1. the ICS naïve children were less severely affected and episodes were mostly triggered by colds (F_{eNO} 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 ICS use in the last 12 months and while this included current use it did not allow for a separate analysis of it). FEV_1/FVC had low diagnostic accuracy in all studies, while $FEF_{25–75\%}$ (0.81) and $FEF_{50\%}$ (0.69) seemed to perform better. Differences in the AUC between the Korean and our study could be due to their exclusion of children with unclear asthma. A study by Korten *et al.* [22] found that adding sR_{tot} and RV/TLC measured during body plethysmography to spirometry measures improved the agreement with the asthma control test in a cohort of 145 asthmatic children from Germany. In our study, sR_{tot} had the highest AUC to diagnose asthma; however, if we wished 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 cut-off values had higher

sensitivity and specificity than those proposed in the algorithm. However, the Manchester study compared healthy children to children with asthma (defined by parental reports of wheeze and asthma treatment in the past 12 months, plus a doctor's diagnosis of asthma ever in life), 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 and not from healthy children. In our clinical population, only 38 children out of 514 could be diagnosed with asthma based on the NICE algorithm (FEV_1/FVC <70% and BDR ⩾12%). Nearly all (83%) would have needed an additional 2 weeks' PEF monitoring, followed by a second visit to the outpatient clinic (supplemental 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. However, in low income settings, PEF monitoring might be the only test available and its value should be studied further. We also found less stringent cut-off values (<84% for FEV_1/FVC, ⩾7% for BDR and ⩾23 ppb for F_{eNO}) to have higher sensitivity and specificity than the values proposed by the NICE algorithm (<70% for FEV_1/FVC, ⩾12% for BDR and ⩾35 ppb for F_{eNO}). The accuracy of the GINA diagnostic algorithm has not been studied previously and although its specificity is relatively high at 90% this would still lead to a 10% overdiagnosis of asthma. In addition, its sensitivity was only 42%, meaning 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, draws upon patient history and the 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 also phrase this finding 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 the closest to the true diagnosis available [1, 23, 24]. The multicentre study adds heterogeneity in the tests and diagnoses, but it increases the generalisability of the findings. Finally, some tests were not performed in all children, which could have introduced a selection bias as 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 performed 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 to paediatric respiratory outpatient clinics with suspected asthma. 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 the GINA and NICE guidelines do not seem to be very practical for an outpatient setting and do 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 FEV_1/FVC ratio (<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, in order to generate evidence for threshold values in different tests that are relevant to diagnosis in those age groups (this is particularly true in the case of children). The next step in research should be to take a systematic approach to assessing respiratory symptoms, allergy, F_{eNO}, spirometry, body plethysmography and BDR tests in all children, in order to develop an accurate diagnostic algorithm combining these tests.

Acknowledgements: We thank all the families, laboratory technicians and doctors who participated in the Swiss Paediatric Airway Cohort. We thank Christopher Ritter (Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland) for his editorial assistance.

Author contributions: C.E. Kuehni and C.C.M. 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.C.M. de Jong analysed the data and drafted the manuscript. E.S.L. Pedersen and C. Ardura-Garcia supported the statistical analysis and gave input for interpretation of the data. All authors critically revised the manuscript and approved the final draft as submitted.

Conflict of interest: C.C.M. de Jong has nothing to disclose. E.S.L. Pedersen has nothing to disclose. R. Mozun has nothing to disclose. D. Müller-Suter has nothing to disclose. A. Jochmann has nothing to disclose. F. Singer reports grants from the Schweizerische Gesellschaft für Cystische Fibrose (CFCH) and personal fees from Novartis and Vertex, outside the submitted work. C. Casaulta has nothing to disclose. N. Regamey has nothing to disclose. A. Moeller has nothing to disclose. C. Ardura-Garcia has nothing to disclose. C.E. Kuehni has nothing to disclose.

Support statement: This work was supported by the Swiss National Science Foundation (award 32003B_162820), the Lung League St. Gallen and the Allergiestiftung Ulrich Müller-Gierok. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Global Initiative for Asthma (GINA). *Global strategy for asthma management and prevention, 2019*. Available from: https://ginasthma.org/wp-content/uploads/2019/06/GINA-2019-main-report-June-2019-wms.pdf Date last accessed: August 07, 2019.
2 British Thoracic Society (BTS), Scottish Intercollegiate Guidelines Network (SIGN). British guideline on the management of asthma: a national clinical guideline. www.sign.ac.uk/media/1048/sign158.pdf Date last updated: July 2019. Date last accessed: August 07, 2019.
3 National Institute for Health and Care Excellence (NICE). Asthma: diagnosis, monitoring and chronic asthma management (guideline NG80). www.nice.org.uk/guidance/ng80/resources/asthma-diagnosis-monitoring-and-chronic-asthma-management-pdf-1837687975621 Date last updated: February 12, 2020. Date last accessed: August 07, 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–e157.
5 Yang CL, Simons E, Foty RG, *et al.* Misdiagnosis of asthma in schoolchildren. *Pediatr Pulmonol* 2017; 52: 293–302.
6 Murray C, Foden P, Lowe L, *et al.* 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: 114–123.
7 de Jong CCM, Pedersen ESL, Mozun R, *et al.* Diagnosis of asthma in children: the contribution of a detailed history and test results. *Eur Respir J* 2019; 54: 1901326.
8 Pedersen ESL, de Jong CCM, Ardura-Garcia C, *et al.* The Swiss Paediatric Airway Cohort (SPAC). *ERJ Open Res* 2018; 4: 00050-2018.
9 Kuehni CE, Brooke AM, Strippoli MP, *et al.* Cohort profile: the Leicester respiratory cohorts. *Int J Epidemiology* 2007; 36: 977–85.
10 Asher MI, Keil U, Anderson HR, *et al.* 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: 1107–1136.
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: 2104–2117.
13 Braun-Fahrlaender C, Wüthrich B, Gassner M, *et al.* Prevalence and risk factors of allergic sensitization in Swiss school children [Prävalenz und Risikofaktoren einer allergischen Sensibilisierung bei Schulkindern in der Schweiz]. *Allergologie* 1999; 22: 54–64.
14 Schiller B, Hammer J, Barben J, *et al.* Comparability of a hand-held nitric oxide analyser with online and offline chemiluminescence-based nitric oxide measurement. *Ped Allergy Immunol* 2009; 20: 679–685.
15 Quanjer PH, Stanojevic S, Cole TJ, *et al.* Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.
16 Grzelewski T, Witkowski K, Makandjou-Ola E, *et al.* Diagnostic value of lung function parameters and F_{eNO} for asthma in schoolchildren in large, real-life population. *Pediatr Pulmonol* 2014; 49: 632–640.
17 Ma TT, Zhuang Y, Gong HY, *et al.* 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: 967–974.
18 Brouwer AF, Visser CA, Duiverman EJ, *et al.* Is home spirometry useful in diagnosing asthma in children with nonspecific respiratory symptoms? *Pediatr Pulmonol* 2010; 45: 326–332.
19 Santos MC, Cunha AA. A brief questionnaire for screening asthma among children and adolescents in Rio de Janeiro, Brazil. *Allergol et Immunopathol* 2005; 33: 20–26.
20 Woo SI, Lee JH, Kim H, *et al.* Utility of fractional exhaled nitric oxide (F_ENO) measurements in diagnosing asthma. *Respir Med* 2012; 106: 1103–1109.
21 Eom SY, Lee JK, Lee YJ, *et al.* Combining spirometry and fractional exhaled nitric oxide improves diagnostic accuracy for childhood asthma. *Clin Respir J* 2020; 14: 21–28.
22 Korten I, Zacharasiewicz A, Bittkowski N, *et al.* Asthma control in children: body plethysmography in addition to spirometry. *Pediatr Pulmonol* 2019; 54: 1141–1148.
23 Rutjes AW, Reitsma JB, Coomarasamy A, *et al.* Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technol Assess* 2007; 11: 50. https://njl-admin.nihr.ac.uk/document/download/2001504
24 Lucas JS, Barbato A, Collins SA, *et al.* European Respiratory Society guidelines for the diagnosis of primary ciliary dyskinesia. *Eur Respir J* 2017; 49: 1601090.