Diagnostic accuracy of minimally invasive markers for detection of airway eosinophilia in asthma: a systematic review and meta-analysis

Daniël A Korevaar, Guus A Westerhof, Junfeng Wang, Jérémie F Cohen, René Spijker, Peter J Sterk, Elisabeth H Bel, Patrick M M Bossuyt

Lancet Respir Med 2015; 3: 290–300

Published Online March 20, 2015 http://dx.doi.org/10.1016/S2213-2600(15)00050-8

See Comment page 260

Department of Clinical Epidemiology, Biostatistics and Bioinformatics (D A Korevaar MD, J Wang MSc, J F Cohen PhD, Prof P M M Bossuyt PhD), **Department of Respiratory Medicine** (G A Westerhof MD, Prof P J Sterk PhD, Prof E H Bel PhD), **and Medical Library** (R Spijker MSc), **Academic Medical Centre, University of Amsterdam, Meibergdreef 9, Amsterdam, Netherlands; and Inserm U1153, Obstetrical, Perinatal and Pediatric Epidemiology Research Team, Center for Epidemiology and Biostatistics Sorbonne Paris Cité (CRESS), and Department of Pediatrics, Necker-Enfants Malades Hospital, Assistance Publique-Hôpitaux de Paris, Paris Descartes University, Paris, France** (J F Cohen)

Correspondence to: Dr Daniël A Korevaar, Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Centre, University of Amsterdam, Meibergdreef 9, 1105 AZ, Amsterdam, Netherlands
d.a.korevaar@amc.uva.nl

Summary

Background **Eosinophilic airway inflammation is associated with increased corticosteroid responsiveness in asthma, but direct airway sampling methods are invasive or laborious. Minimally invasive markers for airway eosinophilia could present an alternative method, but estimates of their accuracy vary.**

Methods **We did a systematic review and searched Medline, Embase, and PubMed for studies assessing the diagnostic accuracy of markers against a reference standard of induced sputum, bronchoalveolar lavage, or endobronchial biopsy in patients with asthma or suspected asthma (for inception to Aug 1, 2014). Unpublished results were obtained by contacting authors of studies that did not report on diagnostic accuracy, but had data from which estimates could be calculated. We assessed risk of bias with QUADAS-2. We used meta-analysis to produce summary estimates of accuracy.**

Findings **We included 32 studies: 24 in adults and eight in children. Of these, 26 (81%) showed risk of bias in at least one domain. In adults, three markers had extensively been investigated: fraction of exhaled nitric oxide (FeNO) (17 studies; 3216 patients; summary area under the receiver operator curve [AUC] 0·75 [95% CI 0·72–0·78]); blood eosinophils (14 studies; 2405 patients; 0·78 [0·74–0·82]); total IgE (seven studies; 942 patients; 0·65 [0·61–0·69]). In children, only FeNO (six studies; 349 patients; summary AUC 0·81 [0·72–0·89]) and blood eosinophils (three studies; 192 patients; 0·78 [0·71–0·85]) had been investigated in more than one study. Induced sputum was most frequently used as the reference standard. Summary estimates of sensitivity and specificity in detecting sputum eosinophils of 3% or more in adults were: 0·66 (0·57–0·75) and 0·76 (0·65–0·85) for FeNO; 0·71 (0·65–0·76) and 0·77 (0·70–0·83) for blood eosinophils; and 0·64 (0·42–0·81) and 0·71 (0·42–0·89) for IgE.**

Interpretation **FeNO, blood eosinophils, and IgE have moderate diagnostic accuracy. Their use as a single surrogate marker for airway eosinophilia in patients with asthma will lead to a substantial number of false positives or false negatives.**

Funding **None.**

Introduction

Historically, asthma control has been pursued by means of symptom and lung function monitoring.[1] Although asthma medications are effective in controlling the disease in most patients, a minority deteriorates despite maximum treatment. Non-eosinophilic asthma responds poorly to corticosteroid therapy, the standard treatment for suppressing airway inflammation. About half of patients with asthma seem to be persistently non-eosinophilic.[2]

Bronchoalveolar lavage and endobronchial biopsy are the reference standards for identifying the extent of eosinophilic airway inflammation, but these tests are invasive and expensive. Another option is induced sputum, which has been clinically useful in guiding asthma treatment.[3]

A Cochrane review showed that the frequency of asthma exacerbations is significantly lower in patients in whom inhaled corticosteroids are tailored based on sputum eosinophil levels, compared with those in whom management is based on traditional methods of asthma monitoring.[3] Recent guidelines recommend guiding treatment in severe asthma by sputum eosinophil counts in addition to clinical criteria in centres experienced in using this technique.[1,4] Sputum eosinophilia might also have prognostic value as a marker for persistent airflow limitation,[5] deteriorating asthma over time,[6] and responsiveness to future therapies specifically targeting eosinophilic inflammation, such as mepolizumab.[7]

Unfortunately, sputum induction is time-consuming, needs experienced laboratory personnel, and many patients are unable to produce adequate samples. Several minimally invasive markers of eosinophilic airway inflammation, such as fraction of exhaled nitric oxide (FeNO), blood eosinophils, and serum periostin, could have potential as a surrogate to replace sputum induction, but their accuracy to distinguish between patients with and without airway eosinophilia remains controversial.

We did a systematic review and meta-analysis to obtain summary estimates of the diagnostic accuracy[8] of markers for airway eosinophilia in patients with asthma.

Methods

Search strategy and selection

A medical information specialist (RS) developed searches in Medline, Embase, and PubMed without date or language restrictions (appendix). We included studies if they had assessed the diagnostic accuracy of one or more blood, serum, nasal lavage, or exhaled breath markers[9] (index test) in detecting airway eosinophilia (target condition) in patients with asthma or suspected asthma. Direct airway sampling methods (induced sputum, bronchoalveolar lavage, or endobronchial biopsy) were acceptable reference standards, independent of the threshold for positivity used. We excluded review articles. The searches were updated until Aug 1, 2014. Two independent reviewers (DAK, GAW) examined titles and abstracts of all search results. Full reports of studies that were considered potentially eligible by at least one of the reviewers were obtained and independently assessed for inclusion. Disagreements were resolved by consensus. One reviewer (DAK) also scanned reference lists of included articles, and searched trial registries (ClinicalTrials.gov, Current Controlled Trials, Netherlands Trial Register, and Australian New Zealand Clinical Trials Registry) for unpublished or ongoing studies.

To enrich the number of included studies, we also tried to identify unpublished data by contacting authors of published studies that had not reported on the diagnostic accuracy of a marker to detect airway eosinophilia, but seemed to have data from which accuracy estimates could be calculated (the "enrichment sample"). Studies were selected if they had done at least one index test and one reference standard, as defined above. Such studies were only eligible if they explicitly distinguished patients with airway eosinophilia from those without, included at least an arbitrary number of 50 patients with asthma, and were published before Jan 1, 2014. We contacted corresponding authors through email, and asked whether they were willing to calculate and share estimates of accuracy or to send their masked dataset.

Data extraction and quality assessment

One reviewer (DAK) did the data extraction, which was verified by a second reviewer (GAW). We identified the first author, country, journal, year of publication, recruitment setting, sample size, and characteristics of included patients (age, sex, BMI, atopy status, asthma severity, FEV_1 % predicted, smoking status, corticosteroid treatment status). We also extracted the index test(s), reference standard(s), test positivity thresholds, disease prevalence, accuracy estimates, and data for 2×2 tables presenting index test results by reference standard results for each reported threshold. If 2×2 tables were not reported, we attempted to reconstruct them from summary estimates or by contacting corresponding authors through email. If it appeared from an article that many markers had been assessed, but diagnostic accuracy data were not reported for all of them, we contacted authors to obtain these data. Two authors (DAK, GAW) independently assessed risk of bias and applicability concerns using QUADAS-2.[10]

Statistical analysis

Whenever we obtained datasets from studies in the enrichment sample, we assessed diagnostic accuracy as follows. First, we estimated the ability of each index test to discriminate between patients with and without airway eosinophilia by calculating the area under the receiver operating characteristic curve (AUC-ROC). Then, we selected the optimal cutpoint of sensitivity and specificity on the ROC curve using the Youden index, as has been done by almost all included diagnostic accuracy studies. Depending on the reference standard available, we repeated this analysis for each definition of airway eosinophilia used in the included studies. Patients with missing data on the index test or reference standard were excluded from the analysis for that specific marker. We analysed datasets using R v3.0.

See Online for appendix

We analysed studies in children and adults separately. To get a view of the overall diagnostic performance of each marker, we did a random effects meta-analysis of AUC estimates,[11] independent of the reference standard or definition of airway eosinophilia that had been used. Whenever a study reported more than one AUC estimate for one marker in the same group of patients, for example because the study relied on many definitions of airway eosinophilia, we included the highest AUC reported. If a study reported an AUC estimate in the total study group and in subgroups, we only included the estimate for the total study group. However, if a study reported on these estimates in subgroups only, and not in the total study group, we included the AUCs of all subgroups. If sufficient data were available (three or more studies), we repeated this meta-analysis for studies that had used the same reference standard and airway eosinophilia definition. We assessed statistical heterogeneity using the I^2 statistic.[12]

From each collected or reconstructed 2×2 table, we calculated estimates of sensitivity and specificity and 95% CIs. We used a hierarchical random effects model[8] to obtain summary estimates of sensitivity and specificity for studies that had used the same reference standard and airway eosinophilia definition. We did so whenever four or more tables were available. If articles provided data on direct, head-to-head comparisons of two or more markers, we assessed whether there were significant differences in accuracy between markers. Such direct comparisons ensure that differences in accuracy are not caused by heterogeneity across study populations. We used Deeks' funnel plot asymmetry test to assess risk of publication bias.[13] We used SAS v9.2 to fit the models.

Role of the funding source

There was no funding source for this study. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

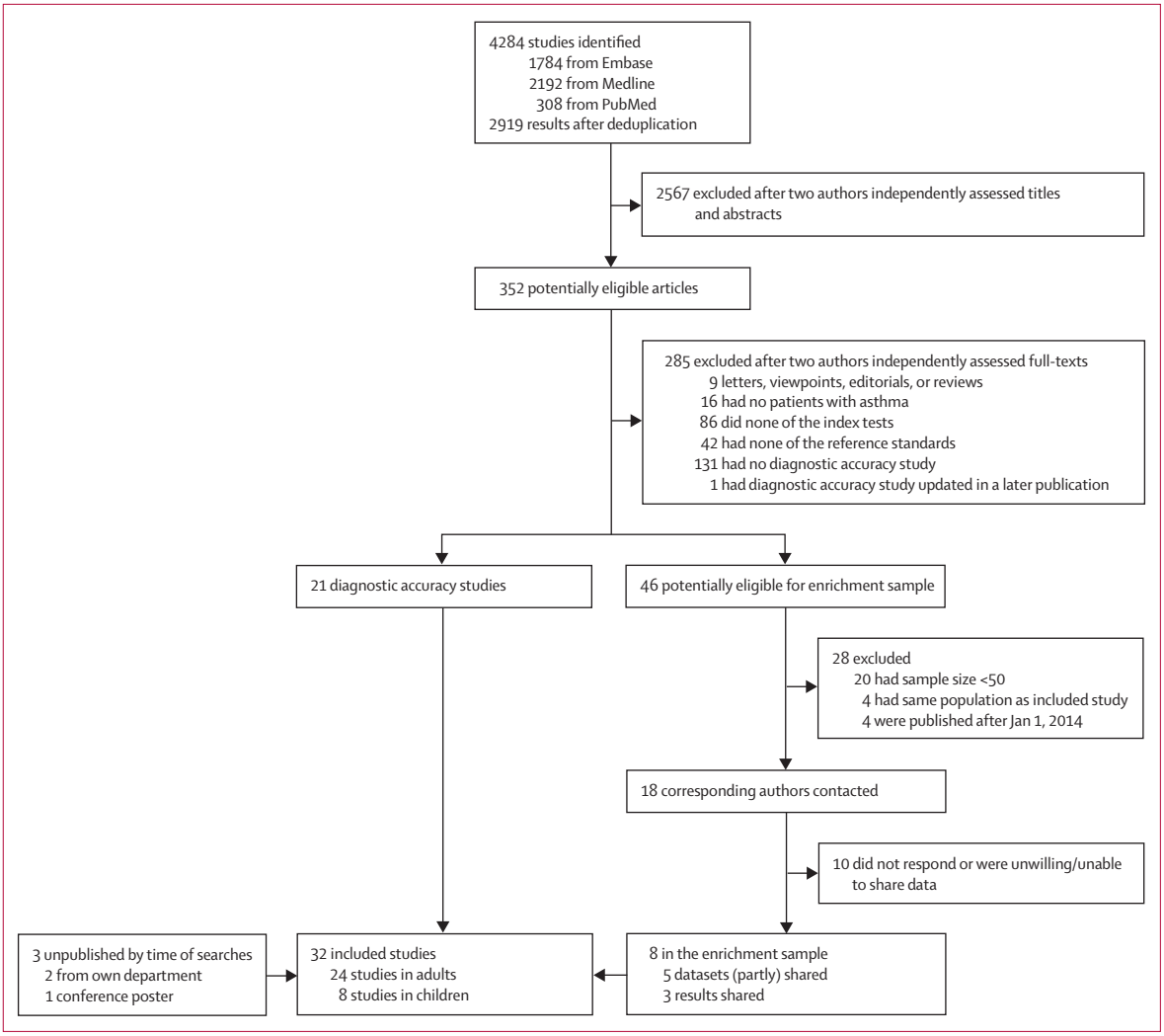

Figure 1: **Study selection**

	Studies in adults*				Studies in children*			
	Studies assessing marker (n)	AUCs included (n)	Patients (n)	AUC† (pooled 95% CI)	Studies assessing marker (n)	AUCs included (n)	Patients (n)	AUC† (pooled 95% CI)
FeNO	17	19	3216	0·75 (0·72–0·78)	6	5	349	0·81 (0·72–0·89)
Blood eosinophils	14	14	2405	0·78 (0·74–0·82)	3	3	192	0·78 (0·71–0·85)
Serum IgE	7	7	942	0·65 (0·61–0·69)	0	..	..	..
Serum periostin	2	3	204	0·65 (0·49–0·81)	0	..	..	..
Serum ECP	2	2	174	0·72 (0·64–0·81)	1	1	77	0·75‡
EBC pH	2	2	96	0·76 (0·63–0·90)	0	..	..	..
Exhaled VOCs	1	1	18	0·98‡	0	..	..	..
EBC model	1	1	53	0·69‡	0	..	..	..
Nasal lavage eosinophils	1	1	130	0·88‡	0	..	..	..

Detailed information on diagnostic accuracy data for individual studies can be found in the appendix. AUC=area under the receiver operator curve. FeNO=fraction of exhaled nitric oxide. ECP=eosinophil cationic protein. EBC=exhaled breath condensate. VOCs=volatile organic compounds. *Five different definitions of airway eosinophilia were used across studies, based on different thresholds for induced sputum, bronchoalveolar lavage, or endobronchial biopsy. †Results based on random effects meta-analysis. ‡Meta-analysis not possible as only one study reported on AUC.

Table 1: **Overall diagnostic performance of markers for detecting any airway eosinophilia**

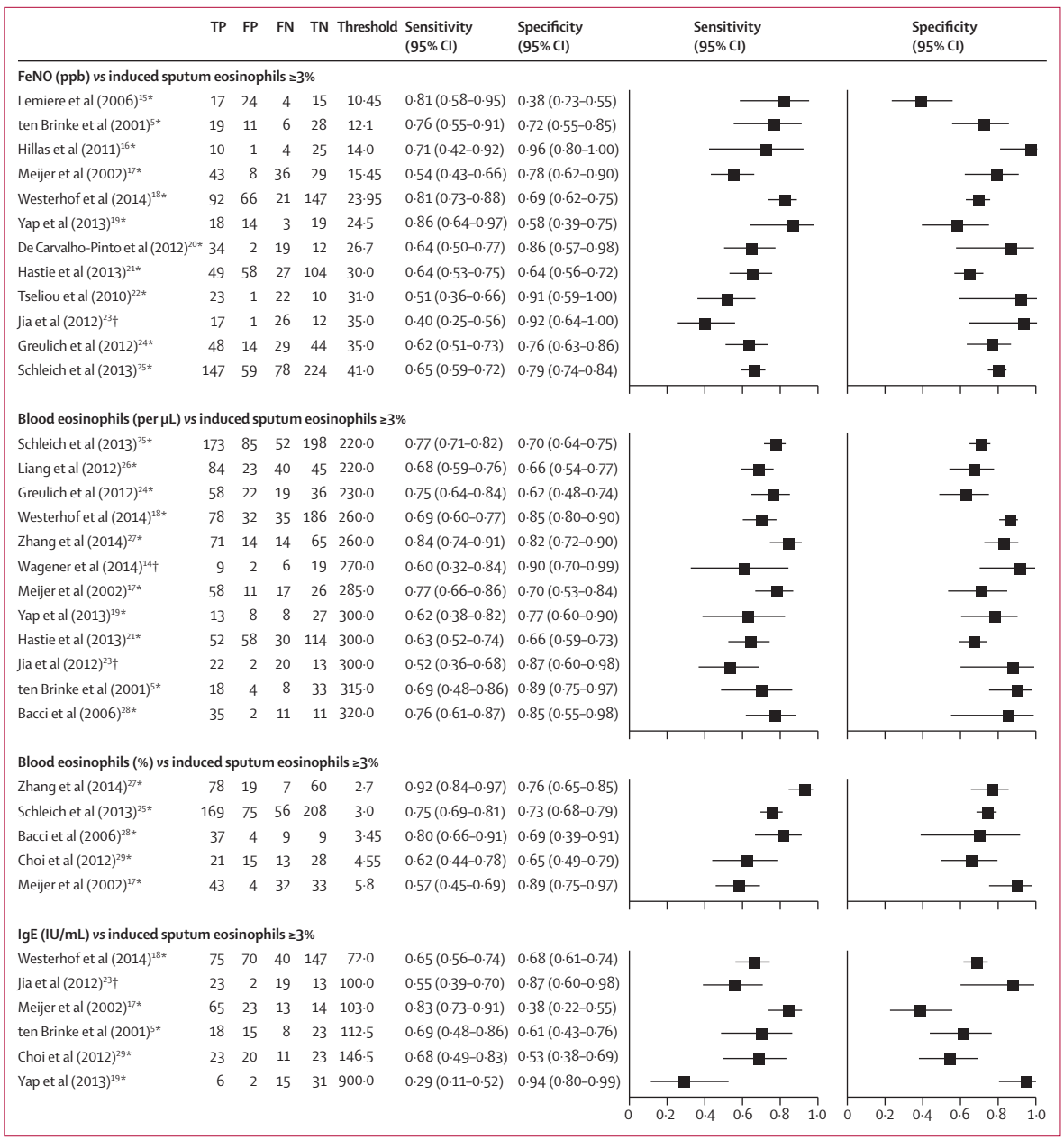

Figure 2: **Forest plots for detection of sputum eosinophils of 3% or more in adults**
Fraction of exhaled nitric oxide (FeNO), blood eosinophils, and IgE for detection of sputum eosinophils of 3% or more in adults. Studies are ordered by threshold. TP=true positive. FP=false positive. FN=false negative. TN=true negative. ppb=parts per billion. *Threshold based on optimal cutpoint between sensitivity and specificity on receiver operating characteristics curve. †Threshold selection arbitrary, based on results from previous studies, or unknown.

Results

The searches retrieved 2919 unique records, all of them providing titles or abstracts in English language. Among these, we found 21 eligible diagnostic accuracy studies (figure 1). Another 18 studies fulfilled the eligibility criteria for the enrichment sample. Contacting the authors of these studies led to eight additional inclusions. We also included data from two studies from our own department, and one more identified through a conference poster. No additional studies were identified by scanning reference lists and searching trial registries. Overall, we included 24 studies done in adults, and eight in children (appendix).

Detailed characteristics of included studies are provided in the appendix. All studies used a single set of inclusion criteria (cohort studies) and the number of patients included in the analysis of diagnostic accuracy varied from 24 to 566 in adults, and from 27 to 150 in children.

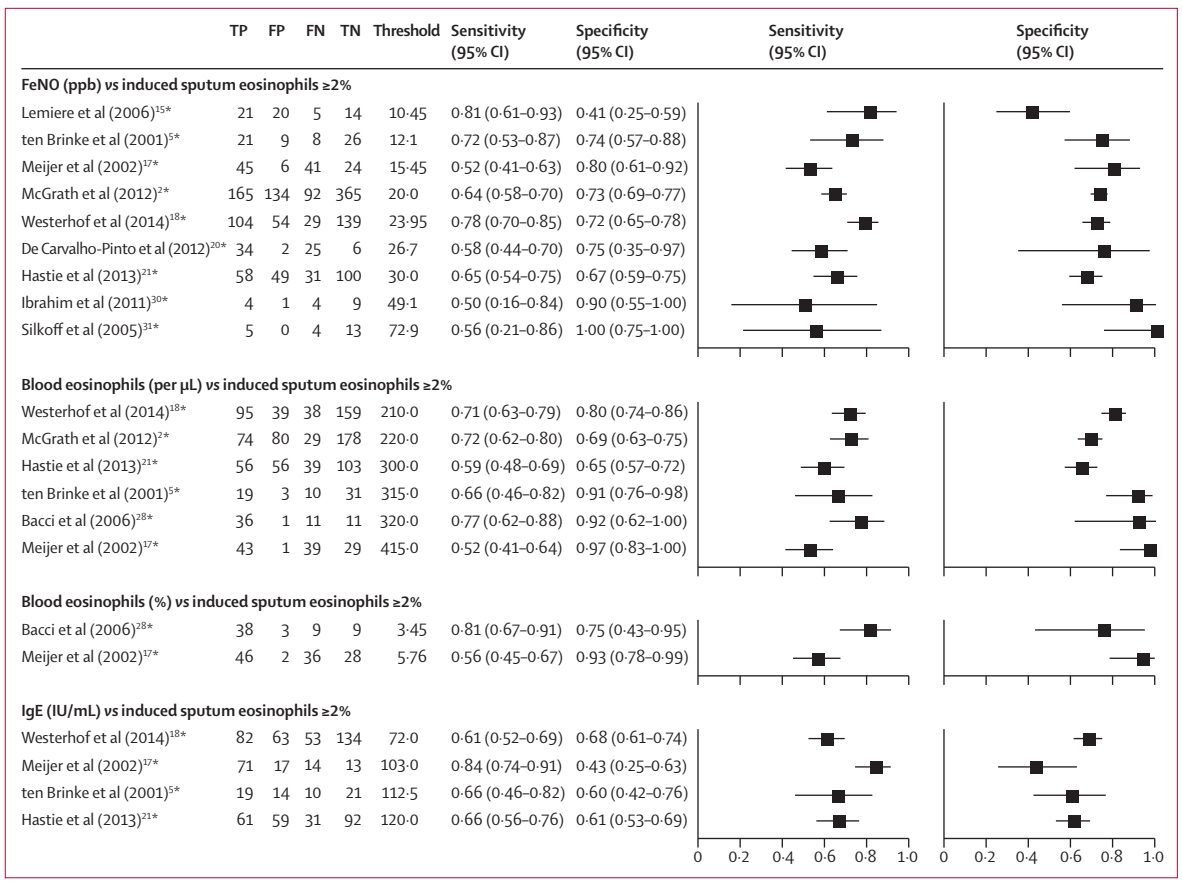

Figure 3: **Forest plots for detection of sputum eosinophils of 2% or more in adults**
FeNO, blood eosinophils, and IgE for detection of sputum eosinophils of 2% or more in adults. Studies are ordered by threshold. TP=true positive. FP=false positive. FN=false negative. TN=true negative. ppb=parts per billion. *Threshold based on optimal cutpoint between sensitivity and specificity on receiver operating characteristics curve.

The mean or median age ranged from 27·0 years to 59·8 years in adults, and from 6·8 years to 13·0 years in children. In all cases, study participants had been recruited in secondary care or tertiary care facilities and both males and females had been included. Studies in adults included patients with asthma of varying severity: mild-moderate (four studies, 17%), mild-severe (four studies, 17%), moderate-severe (four studies, 17%), severe (five studies, 21%), or not reported (seven studies, 29%). In children, asthma severity was mild (one study, 13%), mild-severe (one study, 13%), moderate-severe (one study, 13%), severe (two studies, 25%), or not reported (three studies, 38%). In adults, 12 studies (50%) included current non-smokers only, one study (4%) current smokers only, and 11 studies (46%) included both.

Two studies in adults (8%) assessed corticosteroid (inhaled or oral) untreated patients only, 11 studies (46%) assessed corticosteroid treated patients only, and 11 studies (46%) included both treated and untreated patients. In children, these numbers were one study (13%), three studies (38%), and four studies (50%), respectively. There were large between-study differences in atopy and asthma severity status.

In adults, 21 studies (88%) used only sputum as the reference standard, whereas two studies (8%) used sputum and endobronchial biopsy, and one study (4%) used bronchoalveolar lavage and endobronchial biopsy. In children, sputum was the reference standard in four studies (50%), bronchoalveolar lavage in two studies (25%), bronchoalveolar lavage and endobronchial biopsy in one study (13%), and sputum, bronchoalveolar lavage, and endobronchial biopsy in one study (13%).

Detailed results of the QUADAS-2 assessment are provided in the appendix. All but six studies (81%) showed risk of bias in at least one domain, often because thresholds for index test positivity had not been predefined (21 studies, 66%), or because more than 10% of the patients had been excluded because of missing reference standard results (14 studies, 44%). Additionally, methods for patient sampling (22 studies, 69%) or masking of the index test (20 studies, 63%), or masking of the reference standard (18 studies, 56%) were often unclear.

All diagnostic accuracy data for markers and reference standards are summarised in the appendix. Results of meta-analyses of AUC estimates are presented in table 1, with detailed results in the appendix.

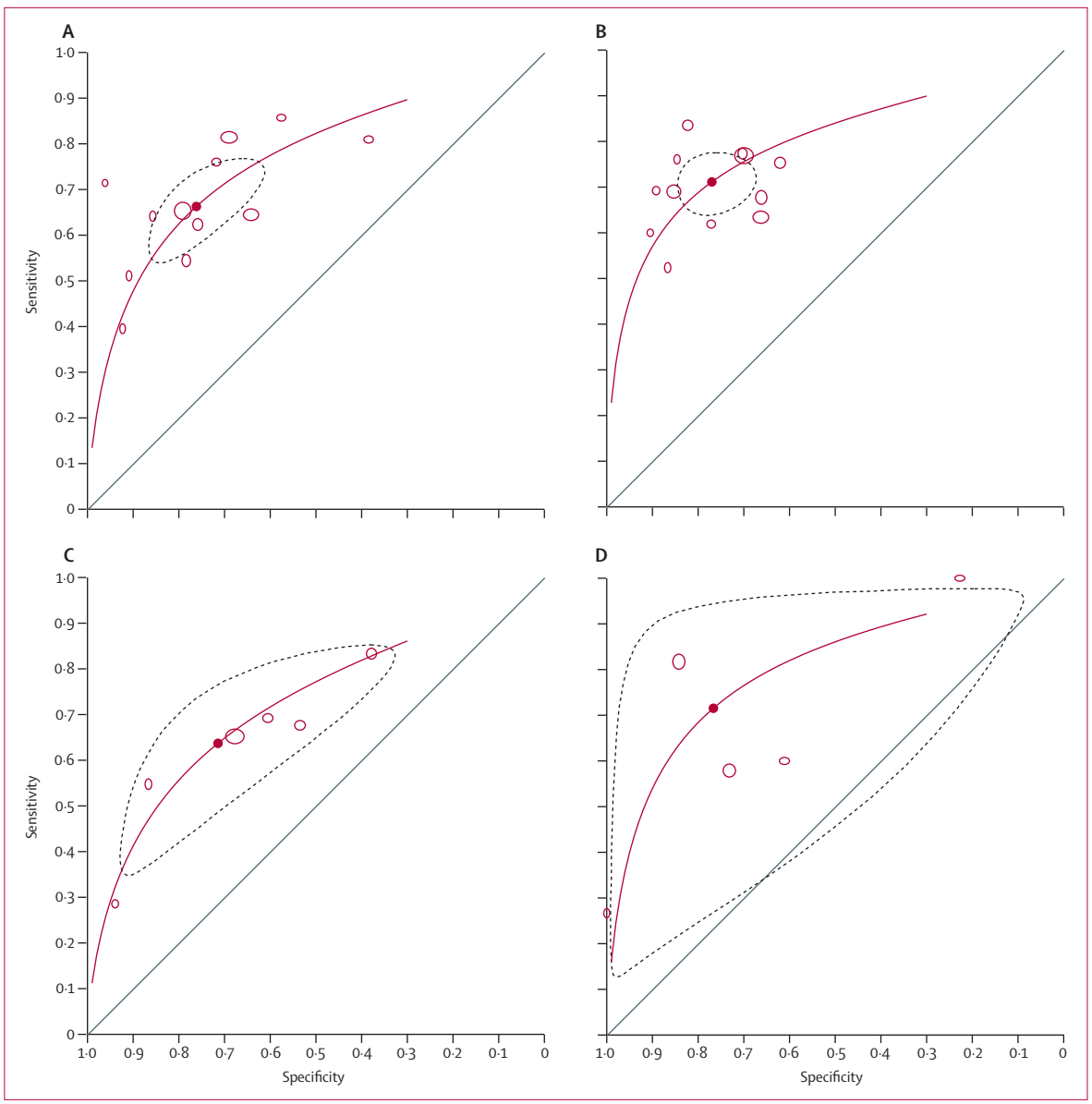

Figure 4: **Summary receiver operating characteristics curve for detecting sputum eosinophils of 3% or more in adults, and 2·5% or more or 3% or more in children**

(A) FeNO (ppb) in adults, (B) blood eosinophils (per μL) in adults, (C) IgE (IU/mL) in adults, and (D) FeNO (ppb) in children. Each open circle is the result from a single study. Closed circles are summary estimates. Dotted ellipses are 95% confidence regions around summary estimates. ppb=parts per billion.

In adults, five different definitions of airway eosinophilia had been used across studies, most often based on sputum eosinophils of 2% or more, or 3% or more. The prevalence of eosinophilia ranged from 20% to 88%. We obtained diagnostic accuracy data for nine markers, but only FeNO, blood eosinophils, total IgE, serum periostin, serum eosinophil cationic protein, and exhaled breath condensate pH had been investigated in more than one study (table 1). When we pooled data, independent of which reference standard or airway eosinophilia definition had been used, the summary AUC of these markers never exceeded 0·78. We found substantial heterogeneity in most analyses (appendix).

FeNO (17 studies, 3216 patients), blood eosinophils (14 studies, 2405 patients), and IgE (seven studies, 942 patients) have been investigated in more than two studies, with pooled AUC estimates of 0·75 (95% CI 0·72–0·78), 0·78 (0·74–0·82), and 0·65 (0·61–0·69), respectively. We repeated these meta-analyses for studies that had used sputum eosinophil values of 3% or more and 2% or more as the definition of airway eosinophilia (appendix), but the summary AUCs were barely affected: 0·74 (95% CI 0·70–0·78) and 0·73 (0·68–0·77),

	Sputum eosinophils ≥3%				Sputum eosinophils ≥2%			
	Studies (n)	Patients (n)	Sensitivity (95% CI)	Specificity (95% CI)	Studies (n)	Patients (n)	Sensitivity (95% CI)	Specificity (95% CI)
FeNO (ppb)	12	1720	0·66 (0·57–0·75)	0·76 (0·65–0·85)	9	1667	0·65 (0·55–0·74)	0·75 (0·62–0·84)
Blood eosinophils (per µL)	12	1967	0·71 (0·65–0·76)	0·77 (0·70–0·83)	6	1180	0·66 (0·56–0·75)	0·83 (0·62–0·94)
Blood eosinophils (%)	5	920	0·76 (0·52–0·90)	0·74 (0·67–0·80)	2	171	..	..
Serum IgE (IU/mL)	6	699	0·64 (0·42–0·81)	0·71 (0·42–0·89)	4	754	0·63 (0·36–0·84)	0·59 (0·37–0·79)

FeNO=fraction of exhaled nitric oxide. ppb=parts per billion.

Table 2: **Summary estimates of sensitivity and specificity for detecting sputum eosinophilia in adults**

	TP	FP	FN	TN	Threshold	Sensitivity (95% CI)	Specificity (95% CI)
Sivan et al (2009)[32]*	76	9	17	48	18·0	0·82 (0·72–0·89)	0·84 (0·72–0·93)
Lex et al (2005)[33]†	6	17	0	5	23·0	1·00 (0·54–1·00)	0·23 (0·08–0·45)
Lex et al (2006)[34]†	3	7	2	11	23·0	0·60 (0·15–0·95)	0·61 (0·36–0·83)
Fleming et al (2013)[35]†	22	11	16	30	35·0	0·58 (0·41–0·74)	0·73 (0·57–0·86)
Toyran et al (2014)[36]*	4	0	11	23	49·8	0·27 (0·08–0·55)	1·00 (0·85–1·00)

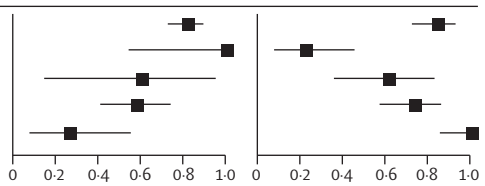

Figure 5: **Forest plots for detection of sputum eosinophils of 2·5% or more, or 3% or more in children**

FeNO for detection of sputum eosinophils of 2·5% or more, or 3% or more in children. Studies are ordered by threshold. TP=true positive. FP=false positive. FN=false negative. TN=true negative. *Threshold based on optimal cutpoint between sensitivity and specificity on receiver operating characteristics curve. †Threshold selection arbitrary, based on results from previous studies, or unknown.

respectively, for FeNO; 0·78 (0·73–0·83) and 0·78 (0·73–0·83) for blood eosinophils; and 0.63 (0·57–0·69) and 0·66 (0·62–0·70) for IgE.

Periostin showed promising performance in one study (AUC 0·84), but these results were not replicated in a second study (AUC 0·55).[14] Nasal lavage eosinophils (AUC 0·88) and a model based on exhaled volatile organic compounds (AUC 0·98) showed high accuracy, but were only investigated in single studies (table 1).

Three studies reported combinations of markers, but none of these showed a significant improvement in the diagnostic accuracy compared with single markers (data not shown).

Comparisons between published and unpublished diagnostic accuracy data for FeNO, blood eosinophils, and IgE are shown in the appendix. Adding unpublished data led to a substantial increase in precision, but did not affect summary estimates of AUCs.

Sufficient data in adults (four or more studies) to do the meta-analysis of sensitivity and specificity for studies that used the same airway eosinophilia definition, were only available for induced sputum as the reference standard. Forest plots of FeNO, blood eosinophils, and IgE for detecting sputum eosinophil values of 3% or more and 2% or more are presented in figures 2 and 3, with summary ROC curves given in figure 4. Almost all studies had used the optimal cutpoint of sensitivity and specificity on the ROC curve to define the positivity threshold of the markers. These thresholds varied widely. For example, the optimal threshold for FeNO to detect sputum eosinophils of 3% or more ranged from 10 to 41 parts per billion (ppb).

Summary estimates of sensitivity and specificity of FeNO, blood eosinophils, and IgE for detecting sputum eosinophils of 3% or more and 2% or more, obtained by meta-analysis, are presented in table 2. Sensitivity ranged from 0·63 to 0·76, and specificity ranged from 0·59 to 0·83. When pooling direct comparisons, FeNO was significantly more accurate than IgE in detecting sputum eosinophils of 2% or more (four studies; p=0·025), but not in detecting sputum eosinophils of 3% or more (five studies; p=0·34). Pooling of other direct comparisons (FeNO *vs* blood eosinophils and IgE *vs* blood eosinophils) showed no significant differences (data not shown).

Statistical testing for funnel plot asymmetry showed no evidence of publication bias (appendix). Forest plots of sensitivity and specificity of FeNO, blood eosinophils, and IgE for detecting sputum eosinophilia in subgroups based on smoking, treatment, and asthma severity status are shown in the appendix. Also in these subgroups, positivity thresholds of the markers varied considerably at the optimal cutpoint of sensitivity and specificity.

In children, five different definitions of airway eosinophilia had been used across studies, most often based on sputum eosinophil values of 2·5% or more (appendix). The prevalence of eosinophilia ranged from 21% to 81%. The diagnostic accuracy was assessed for three markers; two of them in more than one study (table 1): FeNO (six studies, 349 patients) and blood eosinophils (three studies, 192 patients) had pooled AUC estimates of 0·81 (95% CI 0·72–0·89) and 0·78 (0·71–0·85), respectively.

For children, the summary ROC curve and forest plot of FeNO for detecting sputum eosinophils of

2·5% or more, or 3% or more are presented in figures 4 and 5. Summary estimates of accuracy based on five studies (318 patients) were 0·72 (95% CI 0·24–0·95) for sensitivity and 0·77 (0·20–0·98) for specificity, again without evidence of publication bias (appendix).

Discussion

We systematically reviewed studies on the diagnostic accuracy of minimally invasive markers for detecting airway eosinophilia in asthma. In adults, FeNO, blood eosinophils, and total IgE have been extensively investigated, but their ability to distinguish between patients with and without airway eosinophilia is restricted, with summary estimates of AUC, sensitivity, and specificity never exceeding 0·8. Other markers, such as volatile organic compound analysis, were reported to be more accurate in single studies, but these results have not yet been replicated. Studies in children are scarce, but findings for FeNO and blood eosinophils are comparable with those in adults.

Several considerations deserve attention. Almost all studies showed risk of bias. These sources of bias are likely to overestimate diagnostic accuracy,[10] which would mean that the extracted accuracy estimates, although usually moderate, might be even too optimistic. Suboptimal reporting, a common phenomenon for diagnostic accuracy studies,[37] often withheld us from a proper assessment of risk of bias.

Failure to publish is a common phenomenon in diagnostic accuracy studies.[38] We aimed to reduce the risk of publication bias by searching trial registries, and by contacting authors of published studies that seemed to have data from which accuracy estimates could be calculated. This approach was successful. More than one-third of the included results were unpublished at the time of our searches. However, this approach also has its limitations. First, only a minority of diagnostic accuracy studies are registered.[39] Second, most of the included unpublished data came from studies that were, at least partially, reported and had included at least 50 patients. These studies might differ from smaller studies, or those that do not get published at all. Though we did not see any differences between accuracy estimates obtained from published and unpublished data (appendix) and we recorded no funnel plot asymmetry (appendix), we cannot completely exclude the possibility of reporting bias. Drivers of non-publication are unknown in diagnostic research, but it is likely that studies with lower accuracy estimates have lower chances of getting published than those with higher accuracy estimates. Should this be the case, this might have led to further overestimations of accuracy.

Overall, nine different definitions of airway eosinophilia were used across studies, based on different thresholds for eosinophilia in induced sputum, bronchoalveolar lavage, and endobronchial biopsy. These three airway compartments do not show strong correlations with regard to eosinophil counts.[40] Although the diagnostic accuracy of markers can vary across different eosinophilia definitions, we noted that the summary AUCs were stable when comparing studies using any definition of airway eosinophilia, sputum eosinophils of 3% or more, or sputum eosinophils of 2% or more. There was also substantial heterogeneity in the study population and test methods. Some studies only included smokers, for example, whereas others only included non-smokers, and at least four different FeNO measurement devices were used. Many studies analysed both patients with childhood and adult onset asthma. In the latter group, distinguishing asthma from COPD and asthma–COPD overlap syndrome can be problematic. The accuracy of markers can vary across these different subgroups. The prevalence of airway eosinophilia also differed substantially across studies. Diagnostic accuracy typically varies with clinical setting, context, and prevalence. Although the results from the individual studies show substantial heterogeneity, we felt it was safe to draw conclusions because AUCs for FeNO, blood eosinophils, and IgE consistently reflected moderate accuracy.

Combining markers with other clinical features in a prediction model is likely to improve diagnostic accuracy compared with single markers, but this has not been sufficiently investigated yet. All but three studies only reported on accuracy estimates of single markers. Since we did not have individual patient data, we were unable to further analyse the incremental value of combining markers.

The most robust evidence for the clinical value of detecting airway eosinophilia comes from a Cochrane review that showed the frequency of asthma exacerbations can be significantly reduced when tailoring inhaled corticosteroids on sputum eosinophilia.[3] For a marker to be able to replace induced sputum in this context, sensitivity and specificity should probably be at least 90%, so that at most 10% of all patients will be misclassified and, potentially, subjected to inappropriate clinical decisions. Our analysis shows that there are no single markers available with a large enough documented accuracy to fulfil these criteria. However, recent guidelines recommending the use of sputum eosinophil counts in severe asthma, acknowledge that the quality of evidence is very low.[1,4] Additionally, they do not recommend sputum-guided treatment in the general asthma population. Some of the markers assessed in this review on their own might have better potential in managing asthma than sputum eosinophil counts. This fact is shown by a recent study in which volatile organic compound analysis predicted corticosteroid responsiveness with greater accuracy than sputum eosinophils,[41] and by another study that showed good response to mepolizumab in patients with severe eosinophilic asthma as measured by blood eosinophils.[42] The latter study draws attention to the accumulating evidence for the potential role of blood eosinophils as a predictor of responsiveness to novel targeted therapies against eosinophilic airway inflammation.[7]

Moderate accuracy does not necessarily make the investigated markers useless. Markers can also be applied in a triage setting, for example, for ruling-out (high sensitivity required) or ruling-in (high specificity required) airway eosinophilia. In the case of high specificity, for example, those with a positive test result would be considered as eosinophilic, whereas those with a negative test result would need to undergo further testing (eg, sputum induction) because of a large number of false negatives due to a low sensitivity. Most included studies only reported on the optimal cutpoint between sensitivity and specificity, based on the Youden index. When a marker is not sufficiently accurate to replace the existing test, this optimal cutpoint is clinically not very practical because both sensitivity and specificity are typically suboptimal at this cutpoint. Therefore, it does not inform the reader about the ability of the marker to rule-in or rule-out airway eosinophilia. Furthermore, data-driven selection of an optimal cutpoint leads to over-optimistic estimates of sensitivity and specificity.[10] It could be more informative to report on sensitivity at a fixed high specificity (eg, 95%), or the other way around.

An American Thoracic Society guideline on the interpretation of FeNO for clinical applications strongly "recommends the use of FeNO in the diagnosis of eosinophilic airway inflammation".[43] It also strongly "recommends that low FeNO less than 25 ppb (20 ppb in children) be used to indicate that eosinophilic inflammation and responsiveness to corticosteroids are less likely", "that FeNO greater than 50 ppb (35 ppb in children) be used to indicate that eosinophilic inflammation and, in symptomatic patients, responsiveness to corticosteroids are likely", and "that FeNO values between 25 ppb and 50 ppb (20–35 ppb in children) should be interpreted cautiously and with reference to the clinical context". Our results challenge this concept. At FeNO thresholds below 25 ppb sensitivity ranges from 0·52 to 0·86 in adults (appendix). This finding means that of every 100 patients with asthma with airway eosinophilia tested by FeNO, up to 48 would be falsely considered as not having airway eosinophilia, and effective treatment might be withheld from them. In children, sensitivity for FeNO thresholds below 20 ppb ranges from 0·75 to 0·82, showing that up to 25 of every 100 patients with airway eosinophilia would be false negatives. Although these thresholds might be relevant in specific subgroups of asthma, these findings show that FeNO results should be interpreted with much more caution in the general asthma population than recommended by the American Thoracic Society.

It is not surprising that the markers assessed in our review generally were moderately accurate. The underlying biological mechanisms determining airway eosinophil counts are substantially different from those of some of the investigated markers.[44] Several studies also showed significant variability in blood eosinophils[45] and IgE[46] in the same patients with asthma over short periods of time. Some patients with asthma were shown to have persistently raised FeNO concentrations, not suppressed by corticosteroid treatment, and not reflecting raised sputum eosinophils.[47] Corticosteroid treatment significantly affects FeNO, blood eosinophils, IgE, and sputum eosinophils,[48] but the relative magnitude of this effect could vary across markers. Diagnostic accuracy might therefore be affected by treatment status. Also, many other factors, such as age, sex, reflux disease, smoking, and atopy, have been shown to affect FeNO concentrations.[49] This effect might also be the case with other markers and further compromises the identification of an accurate minimally invasive test for airway eosinophilia.

Similar reproducibility problems could apply to the reference standard and target condition. Although some studies showed that a threshold of 3% for sputum eosinophils is reproducible over time,[50] others found the phenotypic classification of asthma to change frequently, both spontaneously and in response to treatment.[51] Longitudinal studies that examined sputum cell counts in successive exacerbations found substantial heterogeneity in the type of inflammation within the same individuals.[52] Consequently, a diagnosis of eosinophilic asthma based on a single sputum sample might be questionable.

Based on our findings, we discourage the use of FeNO, blood eosinophils, or IgE as single surrogate tests for detecting airway eosinophilia in asthma. Our meta-analyses show that, at the optimal cutpoint, sensitivities and specificities of these markers for detecting sputum eosinophilia are moderate, and their use would lead to many false positives or false negatives. Future research will mainly need to focus on whether these markers can be applied as rule-in or rule-out tests, whether markers that were poorly investigated or clinical prediction models incorporating many markers together with other clinical data are more accurate, perhaps in specific settings or subgroups, and whether these markers on their own merits have potential in managing asthma.[48] A next step could be an extensive individual patient data project, combining existing datasets from observational asthma studies in which both clinical features, minimally invasive markers, and one or more reference standards for airway eosinophilia were assessed. Thresholds for ruling-in and ruling-out airway eosinophilia based on individual markers can then be reliably defined, and an optimal multivariable clinical prediction model can be developed. The clinical value of these findings can subsequently be investigated in terms of, for example, response to therapy or the reduction of exacerbations.

Contributors

DAK designed the study in consultation with GAW, JFC, PJS, EHB, and PMMB. RS designed the article search. DAK and GAW did the study selection, data extraction, and quality assessment. DAK and JW did the data analysis. DAK wrote the manuscript. GAW, JW, JFC, RS, EHB, and PMMB made critical revisions to the manuscript.

Declaration of interests
GAW reports support from Novartis to attend the annual congress of the European Respiratory Society, outside of the submitted work. PJS reports grants from Innovative Medicine Initiative, outside of the submitted work. EHB reports grants from GlaxoSmithKline and Chiesi, and personal fees from Novartis, Regeneron, CIPLA, GlaxoSmithKline, and Chiesi, outside of the submitted work. The remaining authors declare no competing interests.

Acknowledgments
We would like to thank the following collaborators who kindly shared their data or provided requested information: Joseph R Arron, Elena Bacci, Maarten van den Berge, Anneke ten Brinke, Andrew Bush, Shaoxi Cai, Regina M De Carvalho-Pinto, Hun Soo Chang, Timm Greulich, John V Fahy, Ana Luisa G Fernandes, Louise Fleming, Stephen J Fowler, Peter G Gibson, Annette T Hastie, Baharudin Ibrahim, Huib A M Kerstjens, Catherine Lemière, Stelios Loukides, Kelly W McGrath, Choon Sik Park, Ian D Pavord, Florence N Schleich, Antonius Schneider, Michael D Shields, Philip E Silkoff, Yakov Sivan, Muge Toyran, Elaine Yap, and Irene S L Zeng.

References

1 Global Initiative for Asthma. Global strategy for asthma management and prevention 2014. http://www.ginasthma.org (accessed Sept 23, 2015).
2 McGrath KW, Icitovic N, Boushey HA, et al. A large subgroup of mild-to-moderate asthma is persistently noneosinophilic. *Am J Respir Crit Care Med* 2012; **185:** 612–19.
3 Petsky HL, Kynaston JA, Turner C, et al. Tailored interventions based on sputum eosinophils versus clinical symptoms for asthma in children and adults. *Cochrane Database Syst Rev* 2007; **(2):** CD005603.
4 Chung KF, Wenzel SE, Brozek JL, et al. International ERS/ATS guidelines on definition, evaluation and treatment of severe asthma. *Eur Respir J* 2014; **43:** 343–73.
5 ten Brinke A, Zwinderman AH, Sterk PJ, Rabe KF, Bel EH. Factors associated with persistent airflow limitation in severe asthma. *Am J Respir Crit Care Med* 2001; **164:** 744–48.
6 Lovett CJ, Whitehead BF, Gibson PG. Eosinophilic airway inflammation and the prognosis of childhood asthma. *Clin Exp Allergy* 2007; **37:** 1594–601.
7 Pavord ID, Korn S, Howarth P, et al. Mepolizumab for severe eosinophilic asthma (DREAM): a multicentre, double-blind, placebo-controlled trial. *Lancet* 2012; **380:** 651–59.
8 Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM. Systematic reviews of diagnostic test accuracy. *Ann Intern Med* 2008; **149:** 889–97.
9 Kim MA, Shin YS, Pham le D, Park HS. Adult asthma biomarkers. *Curr Opin Allergy Clin Immunol* 2014; **14:** 49–54.
10 Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; **155:** 529–36.
11 Zhou XH, Obuchowski NA, McClish DK. Statistical methods in diagnostic medicine. New York: Wiley, 2002.
12 Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21:** 1539–58.
13 Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005; **58:** 882–93.
14 Wagener A, de Nijs S, Lutter R, et al. External validation of blood eosinophils, FeNO, and serum periostin as surrogates for sputum eosinophils in asthma. *Thorax* 2014; **70:** 115–20.
15 Lemière C, Ernst P, Olivenstein R, et al. Airway inflammation assessed by invasive and noninvasive means in severe asthma: eosinophilic and noneosinophilic phenotypes. *J Allergy Clin Immunol* 2006; **118:** 1033–39.
16 Hillas G, Kostikas K, Mantzouranis K, et al. Exhaled nitric oxide and exhaled breath condensate pH as predictors of sputum cell counts in optimally treated asthmatic smokers. *Respirology* 2011; **16:** 811–18.
17 Meijer RJ, Postma DS, Kauffman HF, Arends LR, Koeter GH, Kerstjens HAM. Accuracy of eosinophils and eosinophil cationic protein to predict steroid improvement in asthma. *Clin Exp Allergy* 2002; **32:** 1096–103.
18 Westerhof G, Korevaar D, Amelink M, et al. Biomarkers for sputum eosinophilia in mild-moderate and severe adult-onset asthma patients. *Eur Respir J* 2014; **44** (suppl 58)**:** P2053.
19 Yap E, Chua WM, Jayaram L, Zeng I, Vandal AC, Garrett J. Can we predict sputum eosinophilia from clinical assessment in patients referred to an adult asthma clinic? *Intern Med J* 2013; **43:** 46–52.
20 De Carvalho-Pinto RM, Cukier A, Angelini L, et al. Clinical characteristics and possible phenotypes of an adult severe asthma population. *Respir Med* 2012; **106:** 47–56.
21 Hastie AT, Moore WC, Li H, et al. Biomarker surrogates do not accurately predict sputum eosinophil and neutrophil percentages in asthmatic subjects. *J Allergy Clin Immunol* 2013; **132:** 72–80.
22 Tseliou E, Bessa V, Hillas G, et al. Exhaled nitric oxide and exhaled breath condensate pH in severe refractory asthma. *Chest* 2010; **138:** 107–13.
23 Jia G, Erickson RW, Choy DF, et al. Periostin is a systemic biomarker of eosinophilic airway inflammation in asthmatic patients. *J Allergy Clin Immunol* 2012; **130:** 647–54.
24 Greulich T, Haldar P, Agbetile J, et al. FENO and blood eosinophil count as surrogate markers of eosinophilic airway inflammation in patients with severe asthma. *Eur Respir J* 2012; **36:** 1005S.
25 Schleich FN, Manise M, Sele J, Henket M, Seidel L, Louis R. Distribution of sputum cellular phenotype in a large asthma cohort: predicting factors for eosinophilic vs neutrophilic inflammation. *BMC Pulm Med* 2013; **13:** 11.
26 Liang Z, Zhao H, Lv Y, et al. Moderate accuracy of peripheral eosinophil count for predicting eosinophilic phenotype in steroid-naive non-atopic adult asthmatics. *Intern Med* 2012; **51:** 717–22.
27 Zhang XY, Simpson JL, Powell H, et al. Full blood count parameters for the detection of asthma inflammatory phenotypes. *Clin Exp Allergy* 2014; **44:** 1137–45.
28 Bacci E, Cianchetti S, Bartoli M, et al. Low sputum eosinophils predict the lack of response to beclomethasone in symptomatic asthmatic patients. *Chest* 2006; **129:** 565–72.
29 Choi JS, Jang AS, Park SW, et al. Role of neutrophils in persistent airway obstruction due to refractory asthma. *Respirology* 2012; **17:** 322–29.
30 Ibrahim B, Basanta M, Cadden P, et al. Non-invasive phenotyping using exhaled volatile organic compounds in asthma. *Thorax* 2011; **66:** 804–09.
31 Silkoff PE, Lent AM, Busacker AA, et al. Exhaled nitric oxide identifies the persistent eosinophilic phenotype in severe refractory asthma. *J Allergy Clin Immunol* 2005; **116:** 1249–55.
32 Sivan Y, Gadish T, Fireman E, Soferman R. The use of exhaled nitric oxide in the diagnosis of asthma in school children. *J Pediatrics* 2009; **155:** 211–16.
33 Lex C, Payne DN, Zacharasiewicz A, et al. Sputum induction in children with difficult asthma: safety, feasibility, and inflammatory cell pattern. *Pediatr Pulmonol* 2005; **39:** 318–24.
34 Lex C, Ferreira F, Zacharasiewicz A, et al. Airway eosinophilia in children with severe asthma: predictive values of noninvasive tests. *Am J Respir Crit Care Med* 2006; **174:** 1286–91.
35 Fleming L, Tsartsali L, Wilson N, Regamey N, Bush A. Longitudinal relationship between sputum eosinophils and exhaled nitric oxide in children with asthma. *Am J Respir Crit Care Med* 2013; **188:** 400–02.
36 Toyran M, Bakirtas A, Dogruman-Al F, Turktas I. Airway inflammation and bronchial hyperreactivity in steroid naive children with intermittent and mild persistent asthma. *Pediatr Pulmonol* 2014; **49:** 140–47.
37 Korevaar DA, Wang J, van Enst WA, et al. Reporting diagnostic accuracy studies: some improvements after 10 years of STARD. *Radiology* 2014; published online Oct 27. 141160.
38 Korevaar DA, Ochodo EA, Bossuyt PM, Hooft L. Publication and reporting of test accuracy studies registered in ClinicalTrials.gov. *Clin Chem* 2014; **60:** 651–59.
39 Korevaar DA, Bossuyt PM, Hooft L. Infrequent and incomplete registration of test accuracy studies: analysis of recent study reports. *BMJ Open* 2014; **4:** e004596.
40 Grootendorst DC, Sont JK, Willems LN, et al. Comparison of inflammatory cell counts in asthma: induced sputum vs bronchoalveolar lavage and bronchial biopsies. *Clin Exp Allergy* 1997; **27:** 769–79.

41 van der Schee MP, Palmay R, Cowan JO, Taylor DR. Predicting steroid responsiveness in patients with asthma using exhaled breath profiling. *Clin Exp Allergy* 2013; **43:** 1217–25.

42 Bel EH, Wenzel SE, Thompson PJ, et al. Oral glucocorticoid-sparing effect of mepolizumab in eosinophilic asthma. *N Engl J Med* 2014; **371:** 1189–97.

43 Dweik RA, Boggs PB, Erzurum SC, et al. An official ATS clinical practice guideline: interpretation of exhaled nitric oxide levels (FENO) for clinical applications. *Am J Respir Crit Care Med* 2011; **184:** 602–15.

44 Ricciardolo FL, Sterk PJ, Gaston B, Folkerts G. Nitric oxide in health and disease of the respiratory system. *Physiol Rev* 2004; **84:** 731–65.

45 Spector SL, Tan RA. Is a single blood eosinophil count a reliable marker for "eosinophilic asthma?". *J Asthma* 2012; **49:** 807–10.

46 Mummadi SR, Hatipoglu US, Gupta M, Bossard MK, Xu M, Lang D. Clinically significant variability of serum IgE concentrations in patients with severe asthma. *J Asthma* 2012; **49:** 115–20.

47 Pijnenburg MW, Bakker EM, Lever S, Hop WC, De Jongste JC. High fractional concentration of nitric oxide in exhaled air despite steroid treatment in asthmatic children. *Clin Exp Allergy* 2005; **35:** 920–25.

48 Cowan DC, Taylor DR, Peterson LE, et al. Biomarker-based asthma phenotypes of corticosteroid response. *J Allergy Clin Immunol* 2014; published online Dec 6. DOI: 10.1016/j.jaci.2014.10.026.

49 Taylor DR. Advances in the clinical applications of exhaled nitric oxide measurements. *J Breath Res* 2012; **6:** 047102.

50 van Veen IH, Ten BA, Gauw SA, Sterk PJ, Rabe KF, Bel EH. Consistency of sputum eosinophilia in difficult-to-treat asthma: a 5-year follow-up study. *J Allergy Clin Immunol* 2009; **124:** 615–17.

51 Bacci E, Latorre M, Cianchetti S, et al. Transient sputum eosinophilia may occur over time in non-eosinophilic asthma and this is not prevented by salmeterol. *Respirology* 2012; **17:** 1199–206.

52 D'silva L, Cook RJ, Allen CJ, Hargreave FE, Parameswaran K. Changing pattern of sputum cell counts during successive exacerbations of airway disease. *Respir Med* 2007; **101:** 2217–20.