Diagnostic Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Care Excellence

Title: Measurement of exhaled nitric oxide concentration in asthma; NIOX MINO and NObreath

Produced by ScHARR, University of Sheffield

Authors

Ms Sue Harnan, Research Fellow, Health Economics and Decision Science, ScHARR

Dr Paul Tappenden, Reader in Health Economic Modelling, Health Economics and Decision Science, ScHARR

Dr Munira Essat, Research Associate, Health Economics and Decision Science, ScHARR

Dr Tim Gomersall, Research Associate, Health Economics and Decision Science, ScHARR

Dr Jon Minton, Research Associate, Health Economics and Decision Science, ScHARR

Dr Ruth Wong, Information Specialist, Health Economics and Decision Science, ScHARR

Professor Ian Pavord, Consultant Physician and Honorary Professor of Medicine,Department of Respiratory Medicine, Thoracic Surgery and Allergy, University Hospitals of Leicester NHS Trust

Professor Mark Everard, McCusker Foundation Prof of Paediatric Respiratory Medicine, University of Western Australia, School of Paediatrics & Child Health, Princess Margaret Hospital, Australia

Dr Rod Lawson, Repiratory Physician, Department of Respiratory Medicine, Royal Hallamshire Hospital

Correspondence to Ms Sue Harnan, Research Fellow
Health Economics and Decision Science
ScHARR
Regent Court
30 Regent Street
Sheffield S1 4DA

Date completed 31 August 2013

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 12/60/01.

Declared competing interests of the authors

None.

Acknowledgements

The authors wish to thank Dr Clifford Godley, Dr Daryl Freeman, Professor Hasan Ashad and Dr John White, who provided clinical advice, and either commented on the draft report or sections of the report. They also wish to thank Andrea Shippam (Programme Administrator, ScHARR) for providing administrative support and in preparing and formatting the report. The authors also wish to thank Rachid Rafia for providing an internal peer review of the economic analysis and Eva Kaltenthaler for providing an internal peer review of the clinical review.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Harnan S, Tappenden P, Essat M, Gomersall T, Minton J, Wong R, Pavord I, Everard M, Lawson R. Measurement of exhaled nitric oxide concentration in asthma; NIOX MINO and NObreath. *Health Technol Assess*

Contributions of authors

Sue Harnan acted as Principal Investigator for this assessment. Ruth Wong developed the electronic search strategies. Sue Harnan, Munira Essat and Tim Gommersall undertook the reviews of clinical effectiveness. Paul Tappenden and Jon Minton undertook the review of existing health economic

analyses. Paul Tappenden designed, developed and analysed the *de novo* EAG models. Professor Ian Pavord, Professor Mark Everard and Dr Rod Lawson provided clinical guidance, and commented on and edited the report.

About ScHARR

The School of Health and Related Research (ScHARR) is one of the nine departments that comprise the Faculty of Medicine, Dentistry and Health at the University of Sheffield. ScHARR specialises in health services and public health research, and the application of health economics and decision science to the development of health services and the improvement of public health.

The ScHARR Technology Assessment Group (ScHARR-TAG) synthesises research on the clinical effectiveness and cost-effectiveness of healthcare interventions for the NIHR Health Technology Assessment Programme on behalf of a range of policy makers, including the National Institute for Health and Care Excellence (NICE). ScHARR-TAG is part of a wider collaboration of a number of units from other regions including Health Economics Research Unit and Health Services Research Unit, University of Aberdeen; Southampton Health Technology Assessment Centre (SHTAC), University of Southampton; Liverpool Reviews & Implementation Group (LRiG), University of Liverpool; Peninsular Technology Assessment Group (PenTAG), University of Exeter; the NHS Centre for Reviews and Dissemination, University of York; Warwick Evidence, University of Warwick; the BMJ Technology Assessment Group (BMJ-TAG), BMJ Evidence Centre and Kleijnen Systematic Reviews Ltd.

Table of contents

TABLES

FIGURES

BOX

1. DEFINITION OF TERMS AND LIST OF ABBREVIATIONS

Technical terms and abbreviations are used throughout this report. The meaning is usually clear from the context, but a glossary is provided for the non-specialist reader.

Term	Definition
Airway hyperresponsiveness	Synonymous with bronchial hyperresponsiveness and an indicator of asthma. Usually assessed using a bronchial challenge test. In a bronchial challenge test an agent like histamine or methacholine is inhaled. If these agents trigger bronchospasm at a significantly lower threshold than normal, then the individuals are considered to have airway hyperresponsiveness.
Airway reversibility	Airway obstruction which improves when a bronchodilator or corticosteroids are taken.
Antihistamines	A drug which inhibits the action of histamine in the body, and so may be effective in treating allergic asthma.
Area under the curve (AUC)	A measure of the diagnostic accuracy of a technology based on the geometric inspection of a ROC plot, which plots true positive rate against false positive rate. A technology with perfect diagnostic accuracy will have an AUC of 1, a technology which is no better than chance will have an AUC of 0.5, and a technology which miscategorises on every occasion will have an AUC of zero.
Atopy/atopic disorder	A predisposition towards the development of some forms of allergic hypersensitivity. Atopy is considered to be a risk factor for asthma.
Attrition bias	A statistical bias caused by systematic differences in rates of attrition in the control and intervention arms of a study. For example, the intervention may make some patients receiving it better, but cause others to experience severe side effects and be more likely to leave the study.
Bland-Altman plot	Also known as a difference plot. A plot used to estimate the level of agreement between two devices or assays used for measuring the same thing. Observations are paired, and the mean of the paired observations are plotted against the difference in estimates between the two devices for the same observation.
Bronchoconstriction	Constriction of the airways in the lungs due to the action of surrounding smooth muscle, airway inflammation or excessive production of mucus due to allergy or irritation from air friction, overcooling or drying of the airways.

	It is characterised by coughing, wheezing and shortness of breath.
Chemiluminescence	A broad range of methods in which light is emitted as a result of a chemical reaction. Used to detect the presence and level of nitric oxide in exhaled breath.
Chronic Obstructive Pulmonary Disorder (COPD)	A lung disease in which airflow is persistently poor due to lung tissue damage and dysfunction of the small airways. Some treatment for COPD is similar to that for asthma, but unlike asthma, COPD is usually acquired rather than inherited, and the prognosis and HRQoL is poorer.
Cut-off	A value within a range of values used, in a binary categorisation exercise, to categorise observations into one of two mutually exclusive groups. With respect to the FeNO devices considered in this assessment, the cut-off is typically the number of parts per billion of nitric oxide in exhaled breath; those with values above the threshold are considered 'positive' and below the threshold are considered 'negative'.
Detection bias	Detection bias refers to systematic differences between groups in how outcomes are determined. This usually occurs as a result of preconceptions about treatment efficacy. As such, blinding (or masking) of outcome assessors may reduce the risk that knowledge of which intervention was received, rather than the intervention itself, affects outcome measurement. Blinding of outcome assessors can be especially important for assessment of subjective outcomes, such as degree of postoperative pain. The outcome assessor can be the patient themselves, where outcomes are self-assessed.
Diagnostic accuracy	The effectiveness of a diagnostic test in correctly categorising patients as either 'positive' or 'negative'. There are several ways this can be expressed, for example, the area under the curve (AUC), or the sum of sensitivity and specificity.
Exacerbation	A worsening of symptoms that may be acute or sub-acute. In the case of asthma, this can also be termed an “asthma attack”. Symptoms include shortness of breath, wheezing, cough, and chest tightness. Exacerbations also lead to decreases from baseline in lung function, such as FEV_1.
Extended Dominance	The state when a strategy under study is both less effective and more costly than a linear combination of two other strategies with which it is mutually Exclusive

False negative (FN)	An individual that has been incorrectly categorised as a member of the category 'negative' in a binary categorisation exercise where the only other possible classification is 'positive'. For example, someone who has asthma but has been categorised as not having asthma.
False positive (FP)	An individual that has been incorrectly categorised as a member of the category 'positive' in a binary categorisation exercise where the only other possible classification is 'negative'. For example, a patient incorrectly diagnosed with asthma.
Forced expiratory volume in the first second (FEV_1)	The volume of air expelled by a patient within the first second.
Fraction of Exhaled Nitric Oxide (FeNO)	Also known as fractional exhaled nitric oxide. The proportion of nitric oxide in exhaled breath.
ICS responsiveness	The degree to which asthma condition improves in response to treatment with inhaled corticosteroids.
Index test	A diagnostic test whose sensitivity and specificity is assessed by comparing its categorisations (positive, negative) with another diagnostic test, known as a reference standard, which is assumed to have perfect sensitivity and specificity. In this assessment, the index test is FeNO.
Juniper Score	A quality of life measure for patients with asthma
Negative predictive value (NPV)	The probability that a patient who has been categorised as 'negative' really is negative.
Peak Expiratory Flow (PEF) rate	Maximum rate of expiration of breath, as measured by a peak flow meter (PFM). Considered a measure of lung function.
Pearson correlation	A measure ranging between -1 and 1 indicating the degree and direction of linear dependence between two variables. Values with magnitude close to zero indicate no/very low correlation; and values with magnitude close to one indicate very high correlation.
Performance bias	A statistical bias caused by control and treatment groups receiving different standards of care, or exposure to factors other than the interventions of interest.
Positive predictive value (PPV)	The probability that a patient who has been categorised as 'positive' really is positive.
Receiver operating	A graphic which plots the joint sensitivity and specificity of a diagnostic test

characteristic (ROC) plot	at a range of cut-off thresholds.
Reference standard	A diagnostic test used to estimate the sensitivity and specificity of another diagnostic test, known as an index test. The reference standard is assumed to have perfect sensitivity and specificity, and so where both tests categorise something differently, the index test categorisation is assumed to be incorrect (either a false negative or false positive).
Reporting bias	Reporting bias refers to systematic differences between reported and unreported findings. In any given study, analyses with statistically significant differences between intervention groups are more likely to be reported than non-significant differences. Also known as outcome reporting bias or selective reporting bias. Reporting bias can also occur when results are reported in such a way that they cannot be included in a meta analysis.
Selection bias	Systematic differences in the baseline characteristics of the intervention and control groups. Randomisation should result in study groups with similar baseline characteristics, but can be subverted if there is a lack of allocation concealment (preventing foreknowledge of forthcoming allocations).
Sensitivity	The proportion of 'positives' within a population undergoing diagnostic testing who are identified as such.
Simple Dominance	Where a given treatment alternative is less effective and more expensive than its comparator.
Specificity	The proportion of 'negatives' within a population undergoing diagnostic testing who are identified as such.
Spirometry	Lung function tests based on the measurement of exhaled air under controlled conditions, using a device called a spirometer.
Standardised mean difference (SMD)	A summary statistic showing the difference between two groups, calculated as the difference in mean outcomes between two groups divided by the standard deviation of scores for all study participants. This can be used to meta-analyse data for an outcome that has been measured using different metrics.
True negative (TN)	An individual that has been correctly cateogorised as a member of the category 'negative' in a binary categorisation exercise where the only other possible classification is 'positive'. For example, someone who has been correctly identified as not asthmatic.

True positive (TP)	An individual that has been correctly categorised as a member of a category 'positive' in a binary categorisation exercise where the only other possible classification is 'negative'. For example, someone who has been correctly diagnosed with asthma.

Abbreviation	**Term/Definition**
ACQ	Asthma Control Questionnaire
ACT	Asthma Control Test
AQLQ-M	Asthma Quality of Life Questionnaire-Marks
ASUI	Asthma Symptoms Utility Index
ATS	American Thoracic Society
AUC	Area under the Curve
BTS	British Thoracic Society
CCRCT	Cochrane Central Register of Controlled Trials
CDSR	Cochrane Database of Systematic Reviews
CE	Conformité Européenne
CEAC	Cost effectiveness acceptability Curve
COPD	Chronic obstructive pulmonary disease
CPCI-S	Conference Proceedings Citation Index - Science
CRD	Centre for Reviews and Dissemination
DAR	Diagnostic assessment report
DARE	Database of Abstracts of Reviews of Effects
DSA	Deterministic Sensitivity Analysis
EAG	External assessment group
EIB	Exercise-induced bronchoconstriction
Eos	Induced sputum eosinophil count
EQ-5D	Euroqol 5 dimensions
ERS	European Respiratory Society
FeNO	Fractional exhaled nitric oxide (also known as fraction of exhaled nitric oxide and exhaled nitric oxide (ENO)
FEV_1	Forced expiratory volume in first second
FEV_1%	Percentage of predicted forced expiratory volume in first second
FEV_1/FVC	Forced expiry volume in first second over forced vital capacity
FN	False negative
FP	False positive
GINA	Global initiative for asthma guideline
GM	Geometric mean
GP	General practitioner
GPRD	General Practice Research Database
HRQoL	Health related quality of life
HTA	Health Technology Assessment

HUI-3	Health Utilities Index Mark 3
IAD	Inflammatory airway disease
ICER	Incremental cost effectiveness ratio
ICS	Inhaled corticosteroids
IgE	Immunoglobulin E
IQR	Interquartile range
ITT	Intention to treat
LABA	Long-acting beta agonist
LRTS	Lower respiratory tract symptoms
LTRA	Leukotriene receptor antagonist
mAQLQ	Mini Asthma Quality of Life Questionnaire
MAUDE	FDA Manufacturer and User Facility Device
MCT	Methacholine challenge test
mL/s	Millilitres per second
MSGP	Mortality Statistics in General Practice
NHS	National Health Service
NHS EED	NHS Economic Evaluation Database
NICE	National Institute for Health and Care Excellence
NPV	Negative predictive value
NR	Not reported
OAD	Obstructive airways disease
ONS	Office for National Statistics
OCS	Oral corticosteroids
PEF	Peak expiratory flow
PEFR	Peak expiratory flow rate
ppb	Parts per billion
PPV	Positive predictive value
PSA	Probabilistic sensitivity analysis
PSS	Personal social services
QALY	Quality adjusted life year
QoL	Quality of life
QUADAS-2	Quality assessment tool for diagnostic accuracy studies - second revision
RCT	Randomised controlled trial
RR	Relative risk
RR	Risk ratio
SABA	Short-acting ß2-agonists
SCIE	Science Citation Index Expanded
SCM	Specialist committee members
SD	Standard deviation
SE	Standard error
SEM	Standard error of the mean
SF-12	Short form 12
SF-36	Short form 36

SF-6D	Short form 6 dimensions
SIGN	Scottish Intercollegiate Guidelines Network
TN	True negative
TP	True positive
UK	United Kingdom
VBA	Visual basic for applications

2. EXECUTIVE SUMMARY

2.1 Background

Asthma is a chronic disorder of the airways, caused primarily by inflammatory processes and bronchoconstriction. It is characterised by airflow obstruction and increased responsiveness of the airways to various stimuli. Symptoms of asthma include recurrent episodes of wheezing, breathlessness, chest tightness and coughing. Typical asthma symptoms tend to be variable, intermittent and worse at night. Asthma is commonly triggered by viral respiratory infections, exercise, or external factors, such as smoke, a change in weather conditions and allergens, for instance pollen, mould and house dust mites. There is no cure for asthma, although people may experience long periods of remission. Poorly controlled asthma can have a significant impact on the quality of life of the affected individual and their family.

In 2011, an estimated 5.4 million people in the UK were receiving treatment for asthma. It should be noted however that as there is no definitive, objective test for the diagnosis of asthma, there is significant over- and under-diagnosis of the condition. Despite its high prevalence, deaths resulting from asthma are generally rare. In 2011, the Office for National Statistics reported that there were 1,041 reported deaths due to asthma in England and Wales; approximately two-thirds of these were in women. Around four in five reported asthma deaths occur in adults over the age of 65 years.

Detailed guidelines on the diagnosis and management of asthma have been published and updated by the British Thoracic Society (BTS) and the Scottish Intercollegiate Guidelines Network (SIGN). The diagnosis of asthma is a clinical one and there is no standardised definition of the condition, nor is there a single gold standard recommendation on how it should be diagnosed. Central to all definitions of asthma in adults is the presence of symptoms (wheezing, breathlessness, chest tightness, and cough) and of variable airflow obstruction measured through objective tests of lung function (such as peak expiratory flow (PEF) rate and forced expiratory volume in the first second (FEV_1) divided by forced vital capacity (FVC) known as the Tiffeneau-Pinelli index (FEV_1/FVC) and percentage predicted FEV_1 (calculated as a percentage of the predicted FEV_1 for a person of the same height, sex and age without diagnosed asthma)). Variability of PEF and FEV_1, either spontaneously or in response to therapy, is a characteristic feature of asthma. The diagnosis of asthma in children is clinical, based on recognising a characteristic pattern of episodic symptoms in the absence of an alternative explanation. Lung function tests are less useful due to variability and the inability of very young children to perform these tests reliably. For both children and adults, the BTS/SIGN guidelines indicate that the severity of asthma should be judged according to symptoms and the amount of medication required to control symptoms.

The management of asthma aims to control symptoms (including nocturnal symptoms and exercise-induced asthma), prevent exacerbations and achieve the best possible lung function, with minimal side effects of treatment. For both children and adults, asthma is monitored and managed in primary care by routine clinical reviews on at least an annual basis. These clinical reviews include (but are not limited to) assessment of patient's symptom score using a validated questionnaire, exacerbations, oral corticosteroid use, time off school or work, growth, inhaler technique and in adults, lung function assessed by spirometry of peak expiratory flow. Patients are managed in a stepwise manner, with escalation of medication until control is reached.

This assessment concerns the potential role of nitric oxide monitors (FeNO) in the diagnosis and management of asthma. High FeNO levels in people with symptoms suggestive of asthma, such as coughing and wheezing, may suggest that the patient has eosinophilic asthma that could be treated with inhaled corticosteroids (ICS). In individuals already diagnosed with asthma, changes in FeNO levels can indicate how well a patient is responding to ICS-based medication, whether medication is being adhered to, and whether the dosage of medication should be increased or decreased (titrated, or step-up/step-down adjustment). This assessment concerns the diagnostic accuracy, clinical effectiveness and cost-effectiveness of three handheld FeNO monitors: NIOX MINO (Aerocrine), NIOX VERO (Aerocrine) and NObreath (Bedfont Scientific).

2.2 Objectives

The aim of the assessment is to assess the clinical effectiveness and cost-effectiveness of FeNO measurement in people with asthma. This can be separated into two distinct questions:

1. What is the clinical and cost-effectiveness of FeNO testing in the diagnosis of asthma in adults and children?
2. What is the clinical and cost-effectiveness FeNO testing in the management and monitoring of asthma in adults and children?

2.3 Methods

The assessment report is comprised of two main parts: (1) an assessment of the clinical evidence relating to FeNO in the diagnosis and management of asthma, and; (2) an assessment of the cost-effectiveness of FeNO versus standard care in the diagnosis and management of asthma.

2.3.1 Clinical evidence review

Two systematic reviews and one rapid review were conducted concurrently to identify clinical evidence relevant to the decision problem.

- *Rapid review of equivalence of FeNO devices.* As there was not sufficient evidence from primary research studies which used the three devices that are the focus of this appraisal, a review of the equivalence of NIOX MINO, NIOX VERO and NObreath to other FeNO chemiluminescent devices was conducted. This equivalence review aimed to establish whether FeNO measurement devices could be considered to be equivalent in their measurements to one another, and so whether studies that used other devices could helpfully inform this appraisal.
- *Systematic review of diagnostic accuracy of FeNO for asthma.* The ideal study design would recruit patients with symptoms of asthma, have a cohort design or randomise them to diagnosis using FeNO or diagnosis using other methods and follow them to clinical outcomes. Such studies are known as end-to-end studies and demonstrate the ability of the test to improve patient outcomes. In the absence of such studies, diagnostic cohort studies represent the next best level of evidence, with modelling of clinical outcomes based on the numbers of patients classed as true-positive (TP), true-negative (TN), false-positive (FP) and false-negative (FN). Below this are correlation studies. All levels of evidence were searched for in this review; lower levels of evidence were consulted where the higher levels of evidence were not identified. Where available, three pairs of sensitivity and specificity were selected: those that produced the highest sum of sensitivity and specificity; those that had the highest sensitivity for rule-in scenarios; those that had the highest specificity for rule-out scenarios. In rule-in scenarios, patients testing positive are assumed to have asthma, and those testing negative go on to have further tests for asthma. In rule-out scenarios, those who test negative are assumed not to have asthma, and those who test positive go on to have further tests for asthma.
- *Systematic review of the efficacy of FeNO-guided management of asthma.* Owing to the existence of previous systematic reviews of RCT evidence in adults and in children, only RCT evidence was searched for in this review, with additional interrogation of the database for data on subgroups where RCT evidence was not found.

All three systematic reviews were undertaken according to robust high quality methodology.

2.3.2 Cost-effectiveness assessment

The cost-effectiveness assessment of FeNO includes two components: a systematic review of existing economic analyses and the development of two de novo health economic models.

- *Systematic review of the cost-effectiveness of FeNO for the diagnosis and/or management of asthma.* A systematic review was undertaken to identify all existing economic analyses of FeNO testing for asthma; this includes published studies as well as evidence submitted by the manufacturers of NIOX MINO, NIOX VERO and NObreath. The section includes a critical appraisal of the available evidence and a summary of methodological problems and concerns relating to these analyses.
- *Development of two de novo models.* Independent health economic models were developed to assess the incremental cost-effectiveness of FeNO versus standard care in the diagnosis and management of asthma.

2.4 Results

2.4.1 Clinical effectiveness results

Rapid review of equivalence of FeNO devices: A total of 27 studies met the inclusion criteria and were included in the review. Studies relating to all three devices were located, with the data relating to NIOX VERO being submitted by the sponsors. Whilst there was often good correlation between FeNO measurement devices, equivalence of readings could not necessarily be assumed in all situations. Whilst many studies concluded that the comparability of measurements between devices was within clinically acceptable limits, others went on to produce correction equations to correct for systematic bias in measurements. There was also no common justified definition of clinically acceptable differences, and 95% limits of agreement were sometimes very wide (around 20pbb). There seemed to be a generally consistent observation of poorer equivalence between FeNO devices at higher FeNO levels. The direction of disagreement varied between studies and comparator devices.

However, as there is mostly a high degree of correlation between measurements across all devices, estimates of sensitivity and specificity are likely to be a reasonable indication of potential diagnostic accuracy of using FeNO to guide diagnosis and management, but the derived cut-off points are not likely to be interchangeable between devices. As such, for the purpose of this assessment, sensitivities and specificities will be assumed to be interchangeable, but it cannot be assumed that the cut-off points that should be used to achieve them will be the same in each device, and there is still some doubt as to whether the same diagnostic accuracy would be achievable in all devices. This is an important issue that should be considered in the interpretation of the diagnostic accuracy review and the findings of the health economic analysis assessment presented within this report.

Test failure rates were generally low in all devices in adults, the highest reported rate being 3.3%. In children, there may be some problems using the NIOX MINO device in younger children, with failure rates ranging from 5.5% to 27%. One study used NObreath with children and reported no test failures.

Systematic review of diagnostic accuracy of FeNO for asthma: A total of 24 studies met the inclusion criteria and were included in the review. 20 studies were conducted with adults/all ages and four with children. This review grouped studies that were similar to one another in terms of the position of the patients in the UK pathway and the reference standards used. These groups were: adults presenting with symptoms of asthma versus most of or the entire UK pathway; a subset of adults presenting with symptoms of asthma versus airway hyper-responsiveness; difficult to diagnose patients versus airway hyper-responsiveness; patients with chronic cough who were difficult to diagnose, versus ICS responsiveness; children with symptoms of asthma versus various reference standards.

No meta-analysis was conducted in any group as clinical heterogeneity between studies was generally extremely high. Estimates of cut-off points, sensitivity and specificity were not consistent within groups and ranged widely when used as a rule-in test, rule-out test and when considering the highest sum of sensitivity and specificity. Given the wide ranging estimates of sensitivity and specificity, together with heterogeneous cut-off points, it is difficult to draw any firm conclusions as to the diagnostic accuracy of FeNO in any situation and at any given cut-off point. Interestingly, there did not appear to be an obvious difference in the diagnostic accuracy of FeNO versus the whole or parts of the UK pathway in patients who present with symptoms of asthma compared to the diagnostic accuracy of FeNO versus airway hyper-responsiveness in patients who are difficult to diagnose. The large variation in estimates within groups may obscure any true underlying differences in the accuracy of FeNO between groups and versus different reference standards.

However, some limited observations can be made. It would appear that FeNO was more often able to reach 100% specificity than 100% sensitivity, and that ranges of specificity were generally tighter. This may indicate it has best potential for consistency as a rule-in test, though whether this is clinically and cost effective will depend on the resulting balance of consequences for those who are TP, TN, FP and FN. It would also appear that FeNO cut-off points should probably be lower in children than in adults.

In addition to the above, two studies were found that reported results for FeNO in conjunction with another test in adults, one in those difficult to diagnose, and one in patients of all ages with symptoms of asthma. In both cases, the addition of another test to the diagnostic protocol resulted in a change in

diagnostic accuracy, but as this involved the usual trade off between sensitivity and specificity it is difficult to tell if this represents an increase or decrease in clinical and cost-effectiveness.

No cohort studies were found that provided evidence relating to the subgroups defined *a priori*, namely pregnancy, the elderly and smokers/environmental tobacco exposure. As such, lower levels of evidence were consulted, and the results of these studies should be interpreted with some caution.

- Smokers: FeNO appeared to be able to distinguish between asthmatics and non-asthmatics in adult smokers with similar accuracy as in non-smokers and ex-smokers. It would seem likely that FeNO is generally lower in smokers, and it may be useful to consider a patient's smoking status when interpreting results, or to select lower cut-off points for smokers. Limited data in children support the same conclusion as for adults.
- The elderly: A case control study indicated that FeNO is unlikely to be a useful test in the diagnosis of asthma in the elderly.
- Pregnant women: A cross-sectional study indicated that pregnancy does not alter FeNO levels in asthmatics or non-asthmatics, and that FeNO can distinguish between asthmatic and non-asthmatic pregnant and healthy women.

Systematic review of the efficacy of FeNO-guided management of asthma: Four studies in adults were identified. There were high levels of heterogeneity in multiple study characteristics and outcome definitions, and as such it was not possible to draw any firm conclusions as to which step-up/step-down protocol or cut-off points offer the best efficacy. All studies reported a fall in exacerbation rates per person year, though it appeared that this was mostly driven by mild and moderate exacerbations and was not statistically significant in all but one study (Syk *et al*). Pooled analysis showed a non-significant trend in favour of the intervention group for severe exacerbations and a statistically significant decrease in exacerbations in the intervention groups when considering the composite outcomes of any severity of exacerbation. The effects on ICS use were heterogeneous; two studies showed statistically significant decreases in ICS use in the FeNO-guided management groups, one study showed a minor increase, and another showed very similar levels of use in each arm. This may indicate that some step-up step-down protocols were better at decreasing ICS use than others, or may be due to the characteristics of the study populations. Pooled analysis showed a statistically non-significant trend towards decreased ICS use. HRQoL was infrequently reported; the two studies used versions of the AQLQ to measure quality of life and both showed no effect in the global score, but one investigated domains and found a statistically significant difference in the symptoms score.

Despite the heterogeneity in results, and the lack of statistically significant findings, it would seem possible to conclude that, in adults, FeNO-guided management of some or most designs is likely, during the first year of management, to result in a non-significant trend towards better management overall with either a small or zero reduction in ICS use. There was no evidence relating to whether these effects would be maintained over a longer time period.

Five studies in children were identified. One study appeared to recruit a group of patients who were well controlled whilst two others recruited patients who appeared to be poorly controlled. Both reported fewer severe exacerbations in the intervention arm, but not statistically significantly so. All studies reported a decrease in exacerbations (however defined) in the intervention arm, but only one reported a statistically significant reduction. The effects on ICS use were heterogeneous, with two studies showing a statistically significant increase in ICS use, one showing no difference, one being difficult to interpret and one further study not reporting this outcome. HRQoL was only reported within one study, although insufficient details were reported to draw conclusions.

Due to the high levels of heterogeneity in multiple study characteristics and outcome definitions, it was not possible to draw any firm conclusion as to which step-up/step-down protocol or cut-off points offer the best efficacy for management. However, it would seem possible to conclude that FeNO-guided management of most descriptions is likely, during the first year of management, to result in non-statistically significant trends towards better management overall. It is unclear whether ICS use is likely to increase or decrease, and this may depend on the details of the step-up step-down protocols or the characteristics of the patients recruited to the trials in terms of control and severity. There was no evidence relating to whether these effects would be maintained over a longer time period.

2.4.2 Cost-effectiveness results

There is very limited available evidence concerning the cost-effectiveness of FeNO testing for the diagnosis and/or management of asthma. The systematic review presented in this chapter identified one published UK model of FeNO testing in the diagnostic setting and one published model of FeNO testing in the management setting. Both models were published within the same paper (Price *et al)*. Aerocrine submitted a model of FeNO testing for diagnosis and a model of FeNO testing for management; these models were similar to, but not the same as, the published Price *et al* models. The existing economic diagnostic models indicate that NIOX MINO is likely to be cost saving in comparison to other tests routinely used in the diagnosis of asthma, but may be more expensive than standard diagnostic tests when used in conjunction with other tests. Neither diagnostic models capture the health consequences

associated with correct or incorrect diagnostic outcomes, hence these models do not provide any information regarding the economic trade-off between potential additional health gains resulting from the more accurate diagnosis of asthma and the health loss associated with displacing of existing services. The existing management models indicate that NIOX MINO produces more health gain at a lower cost than guidelines alone. The EAG critique of these management models highlighted a number of problems including the use of a short time horizon, the selective use of efficacy evidence from different sources, assumptions regarding equivalence between sputum count monitoring and FeNO, and invalid assumptions regarding the health losses associated with exacerbations. No economic evidence was submitted by the manufacturers for either NIOX VERO or NObreath.

The EAG developed two *de novo* models. The first model assesses the cost-effectiveness of FeNO testing using NIOX MINO, NIOX VERO and NObreath in addition to, or in place of existing tests, as compared against other diagnostic options commonly used in the diagnosis of asthma. The second model assesses the cost-effectiveness of NIOX MINO, NIOX VERO and NObreath plus guidelines versus guidelines alone for the management of asthma.

The EAG diagnostic model suggests that across the diagnostic options included in the economic analysis, the expected difference in QALY gains is likely to be very small. Airway hyperresponsiveness (MCT) is expected to produce the greatest QALY gain. All options which include NIOX MINO or NIOX VERO are expected to be dominated as their marginal per-test cost is higher than that for NObreath (assuming a device lifetime of 10-years). In the base case analysis, all options except airway hyperresponsiveness (MCT) and FeNO plus bronchodilator reversibility testing are expected to be ruled out by simple dominance. The incremental cost-effectiveness of airway hyperresponsivenss (MCT) versus FeNO (NObreath) plus bronchodilator reversibility is expected to be £1.125million per QALY gained. The results of the analysis are particularly sensitive to assumptions about the duration of time required to resolve misdiagnoses, assumptions about health losses incurred by patients who are false-negative, the costs of asthma management and the use of "rule-in" and "rule-out" diagnostic decision rules.

The EAG management model was evaluated across two subgroups – (i) children and (ii) adults. Studies from the clinical effectiveness review were selected for the model, based on similarity to UK practice and patient populations. Sensitivity analyses were conducted using alternative studies to test the stability of results in other populations and versus different comparators. Within both the children and adult subgroup base case analyses, FeNO testing is expected to produce a small incremental QALY gain compared to guidelines alone. In both subgroups, NIOX MINO and NIOX VERO are expected to be dominated as

their marginal per-test cost is higher than that for NObreath. Within the children subgroup, the incremental cost-effectiveness of guidelines plus FeNO monitoring using NObreath versus guidelines alone is expected to be approximately £45,200 per QALY gained. Within the adult subgroup, FeNO monitoring using NObreath versus guidelines alone is expected to cost approximately £2,100 per QALY gained. A similarly favourable result was produced within a further analysis based on a subgroup of women who are pregnant. Importantly, these positive results are not held when alternative trials are used to inform the analysis. The results in the children and adult subgroups are particularly sensitive to assumptions regarding changes in ICS use over time, the number of nurse visits for FeNO monitoring and the duration over which FeNO monitoring is assumed to impact upon exacerbations and ICS use.

2.5 Discussion

Strengths of the assessment

The assessment includes systematic reviews of equivalence of devices, diagnostic accuracy, management efficacy and test failures which have been undertaken according to robust and high quality methods.

The scope of the assessment was agreed by NICE and the SCM during an extensive scoping exercise.

The existing economic evidence base models have been formally critiqued using the Drummond checklist and assessed in terms of adherence of the individual studies to the NICE Reference Case.

The two economic models have been developed to a high standard, based on the decision problem rather than being limited by the available empirical evidence. Both EAG models explicitly address the trade-off between expected additional health gains resulting from the more accurate diagnosis of asthma and the health loss associated with displacing existing services. Whilst many of the parameters included in these models are subject to considerable uncertainty, the use of a modelling framework helps elucidate which parameters are likely to be most important for decision-making.

The assessment report has been peer reviewed by NICE, other experienced HTA researchers and leading experts in the diagnosis and management of inflammatory airways diseases.

Limitations and uncertainties of the assessment

This assessment is subject to several limitations and uncertainties. It is important to note that these limitations are principally sourced in the evidence base, rather than the methods used to interrogate and evaluate it. Overall, the evidence base for this assessment was not of the highest quality. No end to end

studies were found which estimated the clinical utility of FeNO in the diagnosis of asthma, and no studies were found which used NIOX VERO or NObreath. As such, clinical validity studies were included and a review of the equivalence of devices was conducted. This leads to the following limitations:

- The benefits and harms associated with diagnosis of asthma using FeNO have been estimated based on modelling of the consequences of being true-positive, true-negative, false-positive and false-negative. This includes a large number of assumptions and extrapolations many of which cannot be substantiated with empirical evidence.
- The equivalence of devices is assumed, and this may not hold true in practice. As such, FeNO cut-off values as reported in the primary research may not be applicable to measurements using other devices.
- NObreath will always dominate other devices as its efficacy has been assumed to be equivalent, but its unit cost is less.

No study provided estimates relating to the additional diagnostic value of FeNO to the whole UK diagnostic pathway. This limits the scope of the economic analysis.

No short-term diagnosis of asthma is 100% accurate, and as such all diagnostic studies included in the review had a flawed reference standard. However, in the absence of any alternative, these reference standards were considered to be 100% accurate. A better reference standard would have been long term follow-up of patients. Only one study used such a reference standard (Sivan 2009).

None of the management studies in children included a step-up/step-down protocol that allowed ICS to be stepped down on the basis of FeNO alone. This will limit the degree to which ICS use can be reduced and means that one of the major putative benefits of FeNO management has not actually been assessed empirically: the identification of ICS non-responsive asthmatics who can be taken off ICS therapy with no loss of control.

The EAG diagnostic model is based on evidence identified through the systematic review of FeNO. The diagnostic accuracy of other non-FeNO comparators (spirometry, airways reversibility (MCT) and bronchodilator reversibility) was based on comparative studies identified through the review process. It is possible that other studies not identified within the review could be considered relevant to the model. The use of the Hunter *et al* case control study does however mean that all non-FeNO diagnostic options are assessed consistently within the same study.

The EAG diagnostic model and the Price/Aerocrine diagnostic models draw a number of naïve indirect comparisons across studies; this is a limitation of the evidence base rather than the assessment. It does however limit the confidence that can and should be placed on the findings of these diagnostic models.

The EAG management model is based on short-term evidence of the comparative efficacy of FeNO versus guidelines. The extrapolation of these benefits to the longer-term is subject to considerable uncertainty. Again, this limitation reflects the evidence base rather than the model itself.

Generalisability of the findings

Generalisability of evidence relating to FeNO in the diagnosis of asthma

The clinical evidence was heterogeneous in terms of clinical characteristics and results, and studies were selected for modelling based on their similarity to UK practice, and similarity to the subgroups of interest as defined in the protocol (i.e. those difficult to diagnose, or the wider population of those presenting with symptoms of asthma). As such, no single study can be generalised to the whole population and this should be noted when interpreting the results of this assessment.

Some of the subgroups of interest to the appraisal were not modelled. These groups were the elderly, pregnant women, and smokers/those exposed to environmental tobacco smoke. This was due to limitations in the identified evidence. Only inferences as to the generalisability of results from other studies to these populations can be made.

The EAG model is "blunt" in that it assumes that all misdiagnoses are assumed to be later corrected by subsequent tests. The model is not specific about what these tests are.

In addition, all but one of the studies used to inform the diagnostic accuracy parameters (Sivan *et al*) was undertaken in adults. As a consequence, the EAG model does not fully capture differences in the likely diagnostic pathways between children and adult subgroups.

Generalisability of evidence relating to FeNO in the management of asthma

Each study selected for use in the model had its own merits in terms of generalisability.

- Shaw *et al* followed UK practice in terms of the comparator arm management strategy. It also recruited a population from primary care and included mild to severe asthmatics regardless of atopic status. Smokers were excluded, so it is not clear if the results can be generalised to the UK smoking population. It was also not clear which FeNO device was used.

- Smith *et al* recruited what is likely to be a population with mild to moderately severe asthma and used a different step-up/step-down protocol in both the control and intervention arms. It is unclear to what extent this study could be generalisable to the UK population, but it nevertheless provides some insight into the impact that different but plausible efficacy inputs have on the cost-effectiveness estimates.
- Syk *et al* is most notable for having recruited only atopic patients, only non-smokers and only mild to moderate asthmatics. This study is unlikely to have wide generalisability. Again, it nevertheless provides some insight into the impact that different but plausible efficacy inputs have on the cost-effectiveness estimates.

In children, the two studies were modelled. They were selected largely because these two studies reported the most complete sets of data, and recruited different populations. Again, each study has its own merits in terms of generalisability.

- Szefler *et al* had the lowest risk of bias amongst the studies available. It also recruited patients who were difficult to treat, one of the subgroups identified in the scope as being of especial interest, and so generalisability may be limited to this group. The step-up/step down protocol within this trial did not however allow for ICS to be decreased on the basis of low FeNO alone, making it less likely that a decrease in ICS will be seen in the intervention arm in comparison to some other protocols. Therefore, the generalisability of this study largely depends on what type of step-up/step-down protocol is likely to be adopted in the UK.
- Pijnenburg *et al* adopted inclusion criteria which were likely to result in a population of asthmatics who have more stable disease. The step-up/step-down protocol also does not allow for ICS to be decreased on the basis of low FeNO alone, requiring that symptoms are also low. As such, the generalisability of this study is also largely depends on what type of step-up/step-down protocol is likely to be adopted in the UK.

One study was found which recruited pregnant women. The management strategy allowed step-down on the basis of FeNO alone. This study can be generalised within the population of pregnant women.

Equivalence of devices

- As the equivalence of devices is not assured, the generalisability of these results to all three devices is also not assured.
- It is thought that estimates of diagnostic accuracy and efficacy in managing asthma are probably achievable by all devices, as correlation between measurements is good. However,

the actual value that should be used as a cut-off in diagnosis and management is much more difficult to generalise, and further research may be required to estimate the most appropriate values

2.6 Conclusions

Implications for service provision

There is considerable uncertainty associated with all analyses within this assessment. This is largely due to the limitations of the evidence base.

Studies using the devices that are the focus of this review were not available for all analyses, and in the absence of an alternative, equivalence has been assumed between devices. However, there is not a strong indication across the literature to support this assumption.

The clinical evidence relating to the use of FeNO for diagnosis of asthma is highly heterogeneous and difficult to interpret in the context of the insertion of FeNO into a diagnostic pathway. This is compounded by a lack of certainty as to the equivalence of devices used in the primary research studies to the devices that are the focus of this assessment.

The health economic analysis indicates that FeNO could have value in both the diagnostic and management settings. In particular, the diagnostic model indicates that FeNO plus bronchodilator reversibility dominates many other diagnostic tests and may render airway reversibility cost-ineffective. In the management setting, FeNO-guided management has the potential to appear cost-effective, although this is largely dependent on the expected duration over which it continues to impact upon medication decisions. The conclusions drawn from both models require strong technical value judgements with respect to several aspects of the decision problem in which little or no empirical evidence exists.

Suggested research priorities

This appraisal has been limited by several key evidence gaps that would benefit from further research. It could be argued that this technology is currently under-researched and that any conclusions drawn at this stage may be unduly affected by this lack of evidence. However, some of the problems with the evidence base seem intractable in terms of practicalities, and it could also be argued that the available evidence does point towards some benefits to the technology, albeit benefits that are difficult to quantify with certainty.

Some key problems and suggested research priorities are listed below.

- The equivalence of devices is not assured. There are several ways this problem could be addressed, none of which offer a panacea:
    1. Additional extensive equivalence testing of all devices in relation to one another to ascertain what is driving the heterogeneity in study results. This may be expensive and time-consuming, and may still reveal high levels of disagreement between studies owing to the evidence of variability between devices of the same design.
    2. A network meta-analysis of the existing evidence. This was precluded in this project owing to time and resource constraints. There is likely to be a high degree of uncertainty in any such analysis, based on current evidence, and its results may not be useful.
    3. Derivation and validation studies conducted using the devices in question to develop unique cut-off points for each device for management and diagnosis. This may also be expensive and time-consuming.
    4. Explore the option of using intra-subject relative change to assess control when managing asthma. There is already evidence relating to this approach, but it appears to be in comparatively early stages of development. This is not likely to be a useful option in diagnosis.
- Cut-off values are highly variable and are largely based on derivation studies not validation studies. This problem is related to problems with the equivalence of devices. Possible research priorities relating to this include
    1. Large validation studies (possibly preceded by derivation studies) for cut-off values in all populations of interest, using a number of available devices. Whilst expensive and time consuming, these studies could be very valuable.
- The clinical utility of diagnosis of asthma using FeNO compared to current practice is not informed by direct evidence. Possible research priorities relating to this include
    1. A study which charts the clinical utility of diagnosis of asthma using FeNO versus diagnosis of asthma using current guidelines, against a reference standard of long term follow-up of diagnosis to correct for the mis-diagnoses of both diagnostic approaches.
- It is unclear which step-up/step-down protocol offers the best efficacy. Possible research priorities relating to this include
    1. RCT studies which compare different management protocols to one another. It may be that different protocols are necessary in different populations.

2. Studies which aim to derive the best cut-off points for management protocols. This may be influenced by the specifics of the step-up/step-down protocols.

- It is unclear how treatment effects will progress over time. Long term studies following patients for a number of years could address this evidence gap.

3. BACKGROUND

3.1 Condition and aetiology

3.1.1 Introduction

Asthma is a chronic disorder of the airways, caused primarily by inflammatory processes and constriction of the smooth muscle in airway walls (bronchoconstriction). It is characterised by airflow obstruction and increased responsiveness of the airways to various stimuli. Symptoms include recurrent episodes of wheezing, breathlessness, chest tightness and coughing. Typical asthma symptoms tend to be variable, intermittent and worse at night. Asthma is commonly triggered by viral respiratory infections, exercise, or external factors, such as smoke, a change in weather conditions and allergens, for instance pollen, mould and house dust mites.

Asthma usually develops in childhood but may start at any age. It runs in some families, but many people with asthma have no other family members affected. In adults, asthma is more common in women than in men.[1] There is no cure for asthma, although people may experience long periods of remission. Poorly controlled asthma can have a significant impact on the quality of life of the affected individual and their family. However, there may be variation in an individual's perception of the symptoms and how he or she adapts to the condition over time. Clinical measures such as lung function may not correlate with an individual's quality of life scores, but if asthma is well controlled, near-maximal scores on quality of life instruments can be achieved.

3.1.2 Classification of asthma

There are several ways of categorising different types of asthma; amonst these are:

(i) Intrinsic and extrinsic asthma

Asthma can be divided into extrinsic (external cause) and intrinsic (when no causative agent can be found). Extrinsic asthma is triggered by allergens, hence it is also termed as allergic asthma. In extrinsic asthma, the immune system reacts to substances such as pollen and produces antibodies. Individuals with a predisposition to developing such allergies are said to be atopic and may develop any combination of the triad of hay fever, eczema and asthma. In the case of asthma, the allergic reaction is observed in bronchi and bronchioles which results in the production of excess mucus that obstruct the air passage. Extrinsic asthma is commonly seen in children. About ninety percent of childhood asthma cases are due to specific allergens. Individuals with a family history of atopy are at a higher risk of developing extrinsic asthma.

Intrinsic asthma is a non-seasonal, non-allergic form of asthma, which usually first occurs at a later point in life compared against allergic asthma. Intrinsic asthma tends to be chronic and persistent rather than episodic. It is not related to specific allergens and may be provoked by inhalation of chemicals such as cigarette smoke or cleaning agents, non-steroidal anti-inflammatory drugs (NSAIDS), chest infections, emotion, exercise, cold air, food preservatives or various other non-specific irritants.

(ii) Eosinophilic and non-eosinophilic asthma (neutrophilic asthma)

Asthma can also be categorised as eosinophilic or non-eosinophilic. There is some evidence that eosinophils may play an important proinflammatory role in the pathogenesis of asthma,[2,3] though the remains some uncertainty around this and other pathogenic mechanisms associated with asthma. Eosinophils are found in the airways of asthmatics but not healthy subjects and are believed to be related to exacerbations. It has also been noted that suppression of eosinophil infiltration is often associated with amelioration of symptoms,[2] but that the relationship is not close. Poor inflammation control is most closely related to the risk of future exacerbations. The presence of eosinophils may be used to direct treatment as patients without eosinophilic inflammation are thought to be less responsive to ICS treatment.[4] High levels of eosinophils are correlated with high levels of fractional exhaled nitric oxide (FeNO) and it is thought that FeNO could be used as a biomarker of eosinophilic inflammation, and therefore of ICS responsiveness.[5,6] However, the presence of eosinophils is not always a marker of severity of disease; fatal asthma may be associated with neutrophilia rather than eosinophilia.[7] Targeting the type of inflammation may be a better guide to treatment than measures of disease severity alone; for instance glucocorticosteroids are typically very effective in eosinophilic inflammation but less so if the inflammation is neutrophilic.

(iii) Eosinophilic and non-eosinophilic airway disease

Eosinophilic inflammation occurs in both asthma and COPD, and in both cases the appropriate treatment is ICS.[6] There is a view held by some clinicians that rather than a diagnosis of asthma, a diagnosis of responsiveness to ICS (irrespective of diagnostic label (asthma or COPD)) may be a more helpful approach in terms of directing treatment, reducing costs and reducing exacerbations.[6] However, this form of classification has not yet been officially adopted in the British Thoracic Society and the Scottish Intercollegiate Guidelines Network[8] guidelines (BTS/SIGN guidelines) and this report will focus on the diagnosis of asthma as described in these guidelines.

iv) Molecular approaches to classifying asthma phenotypes

There is an increasing trend to characterise asthma by molecular and cellular factors to enable more targeted and personalised therapy. Such efforts are ongoing and specific phenotypes and the implications of these are not yet fully elucidated.[9]

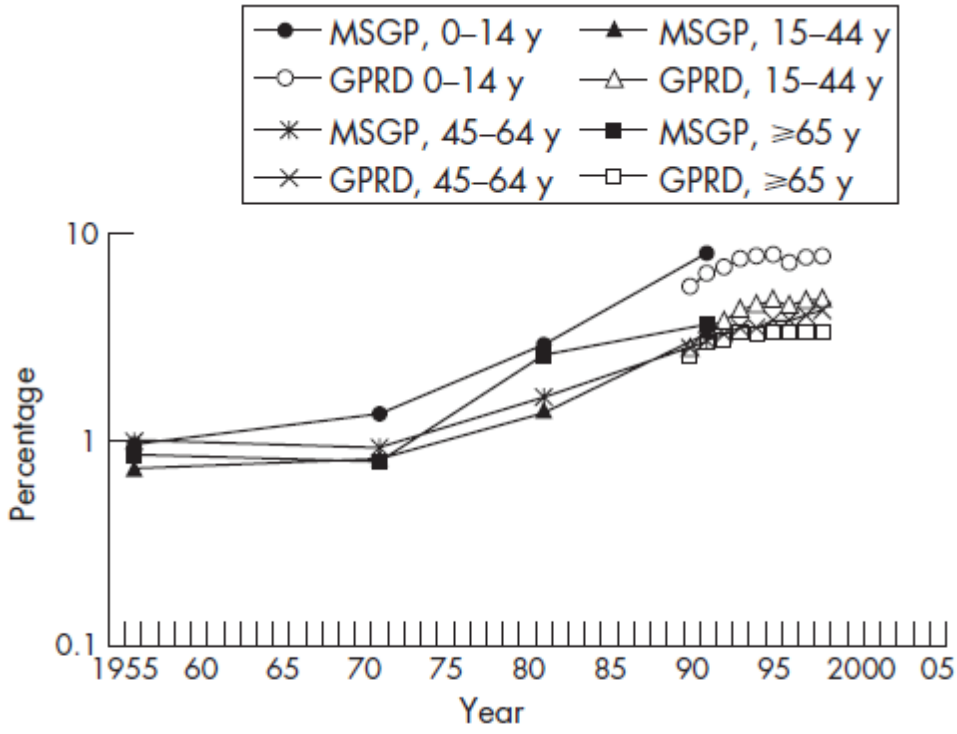
v) Exercise-induced asthma/bronchoconstriction (EIB)

Most patients with asthma will experience EIB, but approximately 11% of the population without other forms of asthma also experience this. It is characterised by a reduction in FEV_1 >10% after exercise and can be treated pharmacologically with short-acting ß2-agonists (SABA) or leukotriene receptor antagonists (LTRA), and non-pharmacologically with a light warm-up before vigorous exercise for example. However, the exact mechanisms behind EIB are not fully understood, but may include neural and biochemical mediators.[10]

3.2 Prevalence of asthma

In 2011, it was reported that 5.4 million people in the UK were receiving treatment for asthma. Of these, 1.1 million were children (1 in 11) and 4.3 million were adults (1 in 12). The UK has among the highest prevalence rates of asthma symptoms in children worldwide. In adults, occupational asthma, for instance due to allergens from animals, flour or grain, may afflict up to 20% of the workforce exposed to the sensitiser. An analysis of routine UK databases undertaken by Anderson *et al*[11] indicates that the prevalence of asthma in all age groups has risen substantially between 1955 and 2004 (see Figure 1).

Figure 1: Patients consulting general practitioners for asthma per 100,000 population, England and Wales, 1955-1988[11]



GPRD – General Practice Research Database; MSGP – Mortality Statistics in General Practice; y – year

Estimates of the prevalence of doctor-diagnosed asthma by age and sex are presented in Table 1, taken from the Health Survey for England 2011.[12]

Based on data from the 2011 Health Survey for England,[12] the prevalence of lifetime doctor-diagnosed asthma was 16% among men and 17% among women, and decreased with age for both sexes. At the time of the survey, approximately nine per cent of men and ten per cent of women were classed as currently having asthma, as they had experienced symptoms of asthma, or were controlling their symptoms with medication, in the previous 12-months. The proportion of respondents with asthma in the last 12-months did not vary by age group in either sex. Of those individuals who had doctor-diagnosed asthma, 30% of men and 39% of women had experienced an asthma attack in the previous 12-months. Of these patients, 42% of men and 52% of women had experienced symptoms during the day in the last week, 22% of men and 29% of women reported that their symptoms interfered with their usual activities in the last week, and 19% of men and 28% of women reported difficulties with sleep in the last week.[12]

Table 1: Prevalence of doctor-diagnosed asthma by age and sex. Health Survey for England 2011[12]

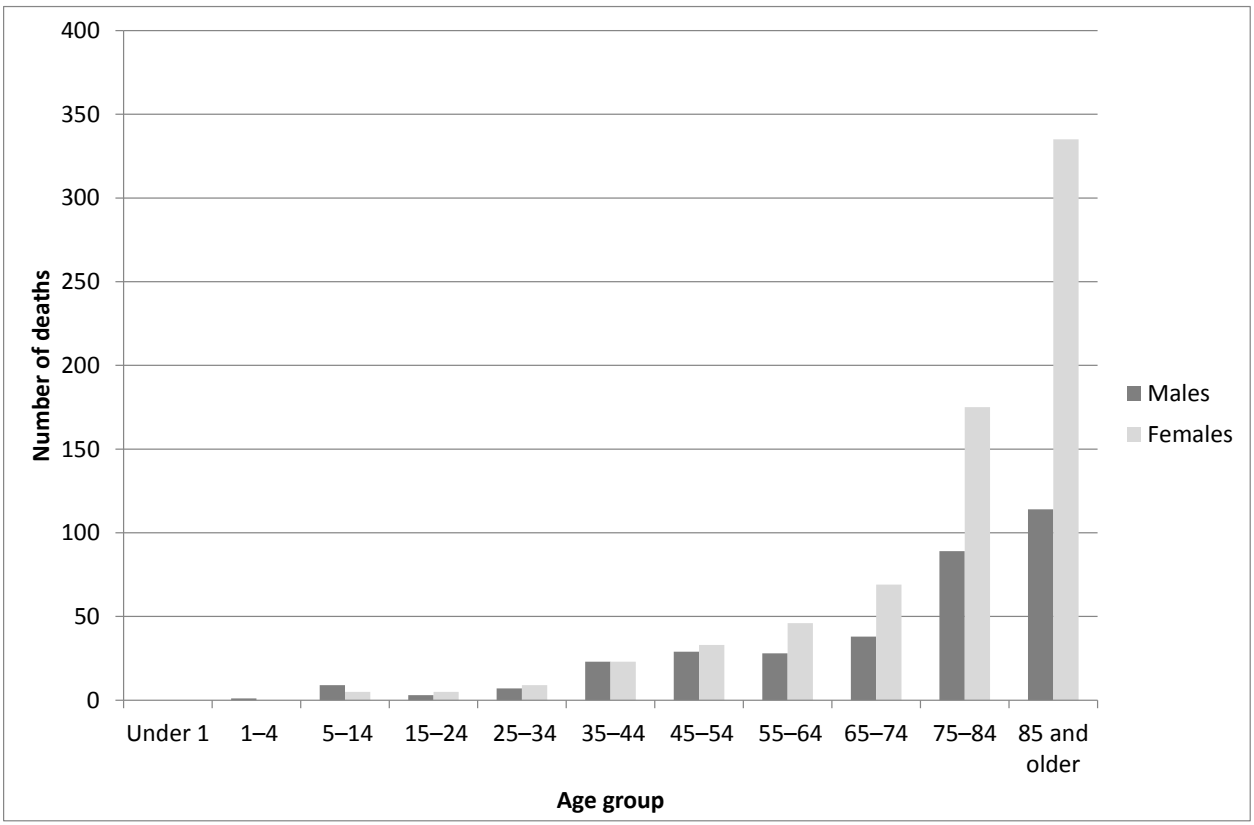
Prevalence of doctor-diagnosed asthma, by age and sex								
Aged 16 and over								*2010*
Asthma	**Age group**							**Total**
	16-24 %	25-34 %	35-44 %	45-54 %	55-64 %	65-74 %	75+ %	**%**
Men								
Ever								
Self-reported doctor-diagnosed asthma	25	20	16	12	13	13	9	**16**
Doctor-diagnosed asthma and in last 12 months								
Symptoms of asthma	6	7	7	4	5	5	4	**5**
No symptoms, asthma controlled with medications	4	3	3	3	3	5	4	**3**
Current asthma: with symptoms of asthma or taking medication	*10*	*10*	*10*	*7*	*8*	*9*	*8*	***9***
No symptoms and no medication for asthma	90	90	90	93	92	91	92	**91**
Women								
Ever								
Self-reported doctor-diagnosed asthma	21	20	17	16	15	16	14	**17**
Doctor-diagnosed asthma and in last 12 months								
Symptoms of asthma	7	7	7	7	8	5	5	**7**
No symptoms, asthma controlled with medications	4	3	4	3	2	7	5	**4**
Current asthma: with symptoms of asthma or taking medication	*10*	*10*	*11*	*10*	*10*	*12*	*10*	***10***
No symptoms and no medication for asthma	90	90	89	90	90	88	90	**90**
Bases (unweighted)								
Men	*378*	*493*	*642*	*624*	*642*	*518*	*402*	***3699***
Women	*476*	*695*	*820*	*874*	*722*	*566*	*563*	***4716***
Bases (weighted)								
Men	*644*	*701*	*754*	*720*	*608*	*429*	*318*	***4175***
Women	*610*	*686*	*760*	*730*	*630*	*470*	*441*	***4328***

Any data on prevalence of asthma is subject to the problems associated with diagnosing asthma. As there is no definitive, objective test, there is significant over and under diagnosis of the condition.

3.3 Asthma mortality

In England and Wales, deaths resulting from asthma are rare. In 2011, the Office for National Statistics (ONS) reported that there were 1,041 reported deaths due to asthma in England and Wales (www.ons.gov.uk). Approximately two-thirds (67.2%) of these were in women. Almost seventy-nine per cent of all asthma deaths were in adults over the age of 65 years.

Figure 2: Registered deaths due to asthma, England and Wales, 2011



As noted elsewhere,[13] audit and case-control studies[14-18] indicate that risk factors for death can be separated into four categories: (1) disease severity, (2) medical care factors both prior to and during the fatal episode, (3) health behaviour such as reduced concordance with prescribed medication, poor inhaler technique and reduced contact with primary care services and (4) adverse psychosocial factors. Shepherd *et al*[13] suggest that given this categorisation, a proportion of asthma-related deaths are preventable, especially in patients under the age of 65 years.

3.4 Impact of the health problem

3.4.1 Impact of asthma on patients

The principal symptoms of asthma are wheezing attacks and episodic shortness of breath. An acute onset of symptoms is known as an exacerbation. Coughing, which worsens at night, may also be a symptom. Asthma exacerbations tend to vary considerably in terms of frequency and duration. Some people experience one or two per year lasting for a few hours, whilst others have exacerbations lasting for weeks, or experience them more frequently. Exacerbations may be precipitated by a wide range of triggers, as described in Section 3.1.2. Asthma is a major cause of impaired quality of life and may impact upon a patient's work, recreational activities, physical activities and emotions. However, whilst patients' health-related quality of life (HRQoL) may be impacted upon by poor asthma control and the incidence of exacerbations, it has been noted elsewhere that meeting clinical treatment goals may not result in noticeable changes in a patient's quality of life.[13]

In the long-term, asthma may lead to permanent airflow obstruction and associated loss of quality of life, especially where it is persistent or poorly controlled.[19] Asthma also has a substantial impact on a patient's ability to work and study and has been estimated to result in at least 12.7 million lost working days per year.[12] Many patients will undergo regular monitoring, and will be required to take medication for the rest of their life. There have been concerns that long term ICS use may reduce growth rates in children, though evidence is conflicting, and it appears that any reduction in growth may be transient, with patients eventually achieving normal adult heights.[20,21]

3.4.2 Burden on the NHS

Given the high prevalence of people with asthma, asthma treatment represents a significant cost to the NHS. The Health Survey for England 2010 estimated that direct healthcare costs associated with asthma are estimated to be £1billion per year. In addition, estimates from 2002 indicate that GP prescriptions alone are approximately £600 million per year.

As asthma is an incurable condition, treatment or at the least monitoring is usually required for the remainder of the patient's lifetime. However, as the diagnosis of asthma is not definitive there is the potential for misdiagnoses to go undetected for many years or even an entire lifetime. Misdiagnosis can occur when a patient appears to respond to treatment, but in fact has experienced a natural resolution of the symptoms of another underlying condition such as a cold, a respiratory infection or allergy. In these cases, patients will appear well controlled, and a treating physician may simply assume that treatment is working. The BTS/SIGN guidelines recommend that patients who are well controlled should "step down"

their therapy doses. This could result in a patient being taken off treatment altogether and their diagnosis being reconsidered. However, clinical input to this review suggests that step-down of doses does not always occur as treatment is relatively cheap per patient, and physicians are cautious not to risk exacerbations. As such, there may be long-term unnecessary NHS expenditure associated with these misdiagnoses. Similarly, both over-treatment and under-treatment of patients who have been correctly diagnosed with asthma may be sources of substantial NHS expenditure. Under-treatment may increase costs to the NHS as poor control may lead to an increased rate of severe exacerbations which require additional primary care management and acute hospital admissions. Over-treatment may increase costs to the NHS because a patient may be able to receive the same level of symptom control with less medication, and so the condition could have been treated as effectively at lower cost.

3.5 Guideline for the diagnosis and management of asthma

Detailed guidelines on the diagnosis and management of asthma have been published and updated by the British Thoracic Society (BTS) and the Scottish Intercollegiate Guidelines Network (SIGN).[8] These guidelines are referred to as the BTS/SIGN guidelines throughout the remainder of this report.

3.5.1 Diagnosis of asthma

The diagnosis of asthma is a clinical one and there is no standardised definition of the condition. Central to all definitions in adults is the presence of symptoms (wheezing, breathlessness, chest tightness, and cough) and of variable airflow obstruction measured through objective tests of lung function (such as peak expiratory flow (PEF) rate and forced expiratory volume in the first second (FEV_1) divided by forced vital capacity (FVC) known as the Tiffeneau-Pinelli index (FEV_1FVC)) and percentage predicted FEV_1 (calculated as a percentage of the predicted FEV_1 for a person of the same height, sex and age without diagnosed asthma). Variability of PEF and FEV_1, either spontaneously or in response to therapy, is a characteristic feature of asthma. The BTS/SIGN guidelines[8] indicate that the severity of asthma should be judged according to symptoms and the amount of medication required to control symptoms.

More recently, descriptions of asthma have included airway hyper-responsiveness and airway inflammation. It is unclear how these features relate to each other, how they are best measured and how they contribute to the clinical manifestations of asthma. Figures 3 and 4 present the diagnostic pathways for children and adults respectively as they currently stand.[8]

Diagnosis in children is clinically-based on recognising a characteristic pattern of episodic symptoms in the absence of an alternative explanation. Lung function tests are less useful due to variability and the

inability of very young children to perform these tests reliably. According to the BTS/SIGN guidelines,[8] clinical features that increase the probability of asthma include:

- More than one of the following symptoms - wheeze, cough, difficulty breathing, chest tightness - particularly if these are frequent and recurrent; are worse at night and in the early morning; occur in response to, or are worse after, exercise or other triggers, such as exposure to pets; cold or damp air, or with emotions or laughter; or occur apart from colds
- Personal history of atopic disorder
- Family history of atopic disorder and/or asthma
- Widespread wheeze heard on auscultation
- History of improvement in symptoms or lung function in response to adequate therapy.

If asthma is suspected, an initial clinical assessment should be carried out to estimate the probability of asthma. According to the BTS/SIGN guidelines,[8] based on initial clinical assessment, an individual child can be classed into one of three groups:

- High probability – diagnosis of asthma likely
- Low probability – diagnosis other than asthma likely
- Intermediate probability – diagnosis uncertain

For children identified as having a low probability of asthma, a more detailed investigation and specialist referral should be considered. For children with a high probability of asthma, a trial of treatment should be started immediately with review at six to eight weeks. Where the response is good, ICS dose should be reassessed every 6 months. Those with a poor response to treatment should undergo more detailed investigations.

There is insufficient evidence at first consultation to make a firm diagnosis of asthma in some children, particularly those below the age of 4 to 5 years.[8] For these children who can perform spirometry and for whom airway obstruction is evident, change in forced expiratory flow volume or peak expiratory flow monitoring should be assessed in response to an inhaled bronchodilator and/or the response to a trial of treatment for a specified period.

In children with an intermediate probability of asthma who can perform spirometry and have no evidence of airway obstruction, tests for atopic status, assessment of bronchodilator reversibility and if possible, bronchial hyper-responsiveness using methacholine, exercise or mannitol should be considered, though

these latter three would be performed in secondary care. In such cases, specialist referral should always be considered.

Other investigations to support a diagnosis of, or alternatively rule out asthma in children include tests of eosinophilic airway inflammation using induced sputum or exhaled nitric oxide concentrations, tests of atopy by skin-prick test or blood eosinophilia and chest x-ray or other imaging techniques to investigate other causes.

Diagnosis in adults is also based on the clinical history and includes the recognition of a characteristic pattern of symptoms and signs and the absence of an alternative explanation for them. However, in contrast to the diagnostic pathway for children, in adutls, spirometry is performed at the first consultation to assess the presence and severity of airflow obstruction.

As in the diagnosis of children, adults are also classified as having a high, low or intermediate probability of asthma. Chest x-ray and specialist referral may be considered in any patient presenting atypically or with additional symptoms or signs.

Figure 3: Diagnosis of asthma in children, BTS/SIGN guidelines[8]

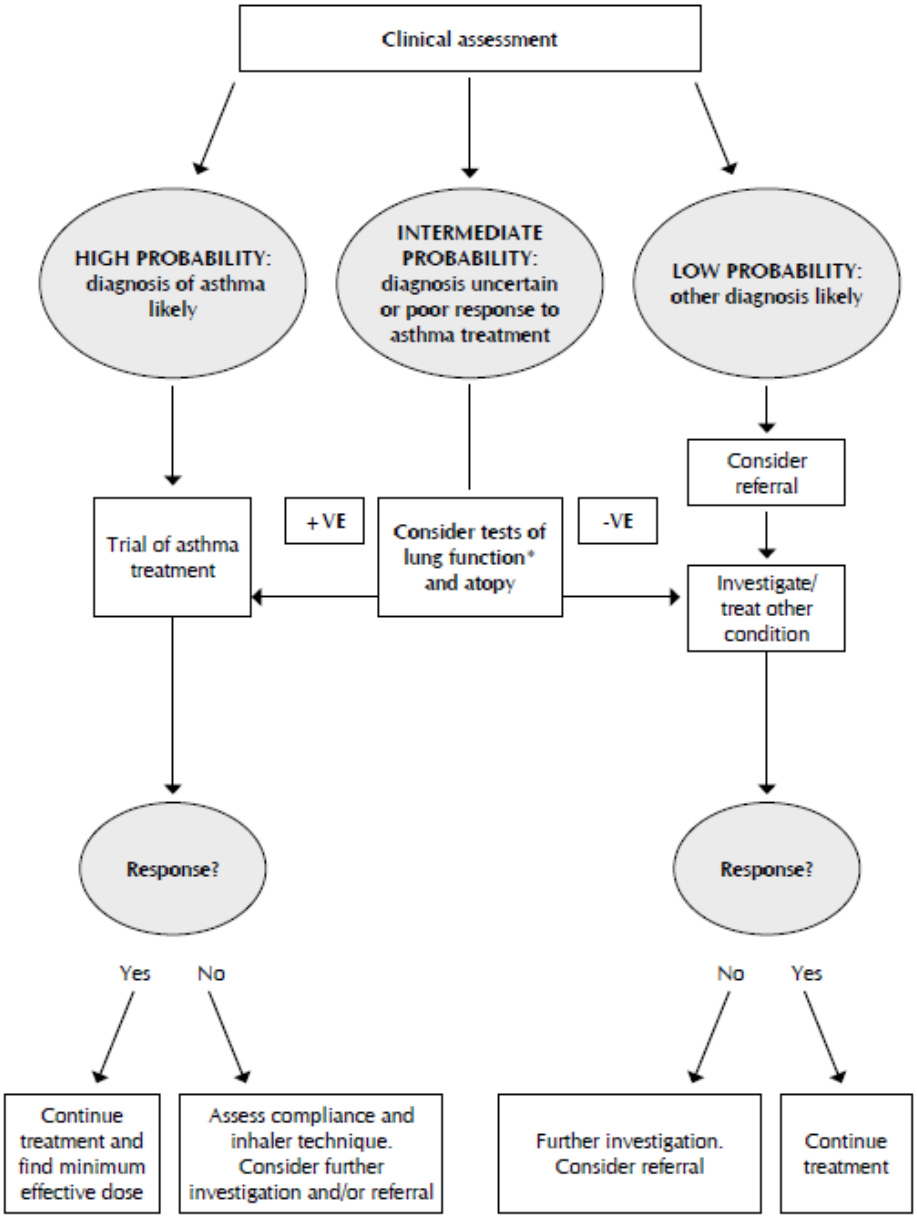
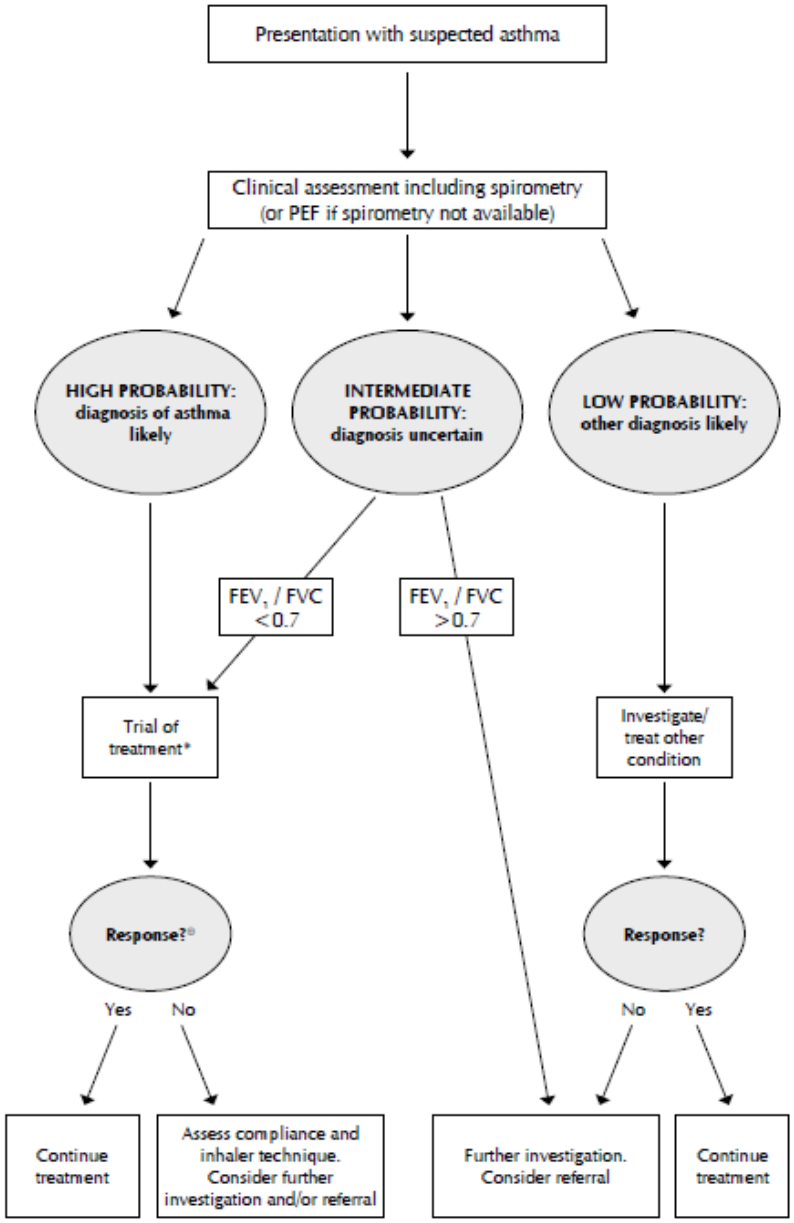


Figure 4: Diagnosis of asthma in adults, BTS/SIGN guidelines[8]



3.5.2 Monitoring and management of diagnosed asthma

Asthma management aims to control symptoms (including nocturnal symptoms and exercise-induced asthma), prevent exacerbations and achieve the best possible lung function, with minimal side effects of treatment. For both children and adults, asthma is monitored and managed in primary care by routine clinical reviews on at least an annual basis. These reviews include (but are not limited to) assessment of patient's symptom score (using a validated questionnaire), exacerbations, oral corticosteroid use, time off school or work, growth, inhaler technique and in adults, lung function assessed by spirometry of peak expiratory flow. Patients are managed in a stepwise manner, with escalation of medication until control is reached. This approach to pharmacological management for children and adults is represented in Figures 5 and 6 respectively (taken from the BTS/SIGN guidelines).[8] Treatment is started at the step most appropriate to the initial severity of the asthma, with the aim of achieving early control of symptoms and optimising respiratory function. Control is maintained by stepping up treatment as necessary and stepping down when control is good (see Figures 5 and 6).

Figure 5: Management of asthma in children (BTS/SIGN guidelines)[8]

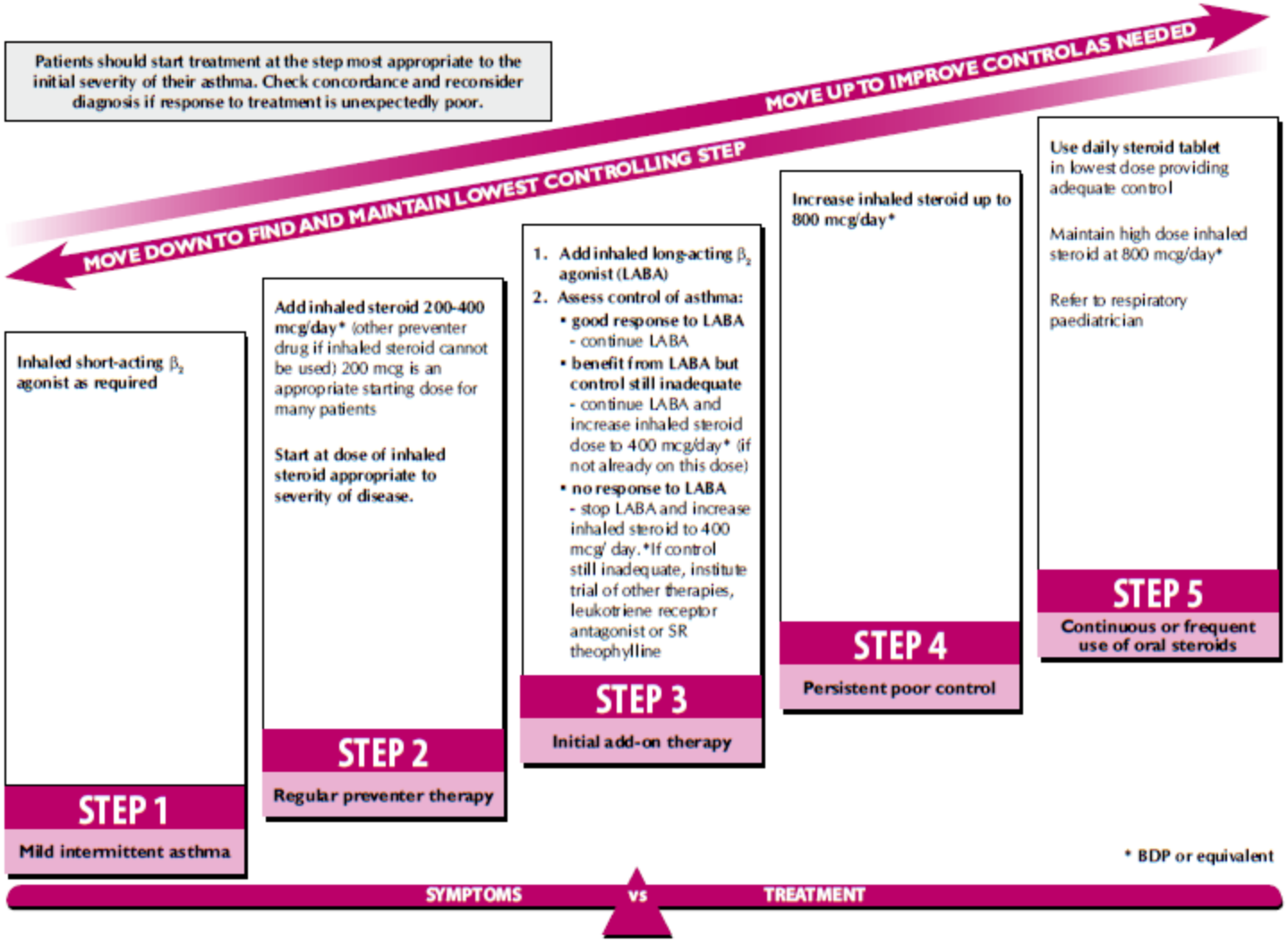
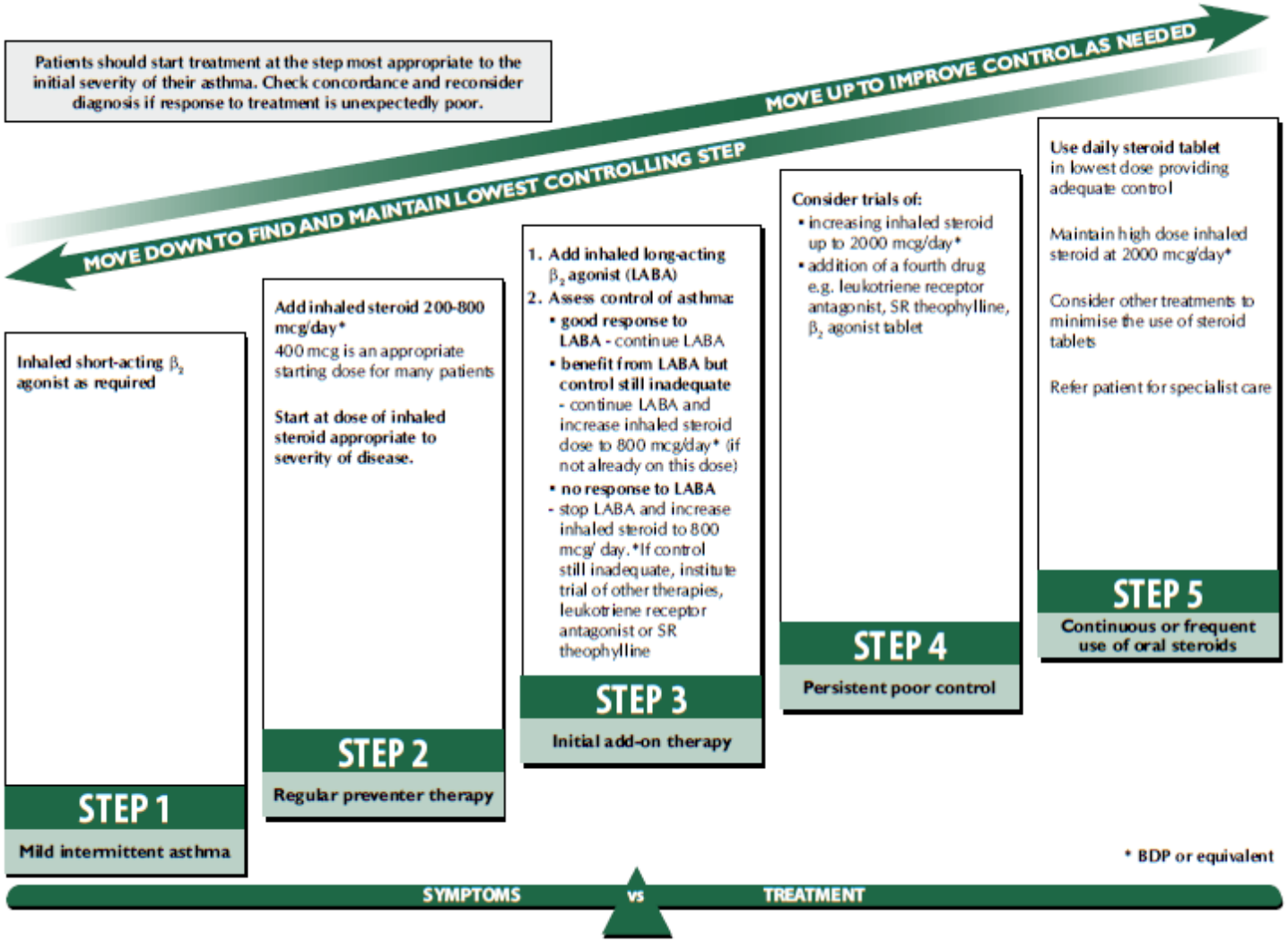


Figure 6: Management of asthma in adults (BTS/SIGN guidelines)[8]



3.5.2.1 Monitoring asthma in children

The BTS/SIGN guidelines[8] on the Management of Asthma state that monitoring of asthma in children should include assessment and recording of:

- symptom score, for instance the Children's Asthma Control test or Asthma Control Questionnaire
- exacerbations, oral corticosteroid use and time off school/nursery due to asthma since last assessment
- inhaler technique
- adherence to treatment, which can be assessed by reviewing prescription refill frequency
- possession of and use of self-management plan/personalised asthma action plan
- exposure to tobacco smoke
- growth (height and weight centile).

The guideline is indistinct with respect to the use of biomarkers such as FeNO in the monitoring of asthma. It states: *"a better understanding of the natural variability of biomarkers independent of asthma is required and studies are needed to establish whether subgroups of patients can be identified in which biomarker guided management is effective."*[8]

3.5.2.2 Monitoring asthma in adults

According to the BTS/SIGN guideline,[8] symptom-based monitoring is adequate in the majority of adults with asthma. Those patients with poor lung function and with a history of exacerbations in the previous year may be at a greater risk of future exacerbations for a given level of symptoms. For adults, the factors that should be assessed and recorded include:

- symptomatic asthma control: best assessed using directive questions such as the Asthma Control Questionnaire (ACQ) or Asthma Control Test (ACT)
- lung function, assessed by spirometry or by PEF
- exacerbations, oral corticosteroid use and time off work or school since last assessment
- inhaler technique
- adherence to treatment, which can be assessed by reviewing prescription refill frequency
- bronchodilator reliance, which can be assessed by prescription refill frequency
- possession of and use of self management plan/personal action plan

3.6 Description of technologies under assessment

3.6.1 The potential role of FeNO devices in the diagnosis and management of asthma

Nitric oxide monitors measure FeNO. High FeNO levels in people with symptoms suggestive of asthma, such as coughing and wheezing, may suggest that the patient has eosinophilic asthma that

could be treated with ICS (see Section 3.1.2). In indivduals already diagnosed with asthma, changes in FeNO levels may indicate how well a patient is responding to ICS-based medication, whether medication is being adhered to, and whether the dosage of medication should be increased or decreased (titrated, or step-up/step-down adjustment). Consequently, FeNO monitors may have a role in the diagnosis, monitoring and management of patients with asthma.

However, current opinion is divided as to the utility of this measurement, in large part due to the potiential for various factors to confound FeNO levels. Amongst these are age, gender, smoking status, exposure to environmental tobacco, pregnancy, height, measurement technique and atopic status and medication.[22,23]

3.6.2 Current service provision

A number of FeNO devices have been developed. Some of these are handheld portable devices (such as the devices that are the focus of this assessment) whilst others are stationary devices which measure FeNO through chemiluminescent techniques. Both types of FeNO monitors have been available for use in the NHS for a number of years. However, they are not available in all secondary care settings, and their use in primary care is extremely rare. There are a number of possible reasons why FeNO devices have not had a more widespread diffusion into care. For example, it may be as a result of the lack of clear guideance in the BTS/SIGN guidelines[8] as to how they should be used, which itself is as a consequence of contradictory research; or due to the previously prohibitive cost and operational requirements of large chemiluminescent devices.

A number of other diagnostic interventions are commonly used in the diagnosis of asthma in England and Wales, as described in Section 3.5.1. Some of these are performed in primary care, such as spirometry, reversibility testing and trials of treatment, whilst other are performed in secondary care, such as airway hyper-responsiveness (MCT) and sputum induction. As noted above, monitoring and management of asthma in diagnosed patients is guided by BTS/SIGN guidelines.[8]

3.6.3 Technologies under assessment

The three handheld FeNO devices included in this assessment are NIOX MINO® (Aerocrine), NIOX VERO® (Aerocrine) and NObreath® (Bedfont scientific).

3.6.3.1 NIOX MINO®

NIOX MINO determines FeNO concentration in a breath sample. The device is small, hand-held and portable, and it can be used by both adults and children. It requires a 10 second exhalation of breath by the patient, at an exhalation pressure of 10-20 cm H_2O to maintain a fixed flow rate of 50±5 mL/s. The last 3 seconds of the 10 second exhalation are analysed by a calibrated electrochemical sensor, to

give a definitive result in parts per billion (ppb). Clinical cut-off values can be applied to the FeNO values to categorise readings as low, intermediate or high according to the reference ranges for ages less than 12 years and 12 years or more, as detailed in the sponsor's submission.[24]

NIOX MINO is pre-calibrated and designed to ensure a service- and calibration-free system. It can be used as a stand-alone device or connected to a PC for monitoring with the NIOX MINO Data Management Program and for use with Electronic Medical Record systems.

NIOX MINO is CE-marked and was launched in the UK in November 2004. According to the scope provided by NICE for this assessment, it is currently available in eight GP surgeries and used in more than 90 hospitals across the UK.[25] The manufacturer claims that NIOX MINO is indicated for use as follows:

- To diagnose the specific type of airway inflammation to guide treatment;
- To predict the onset of asthma symptoms or loss of asthma controls due to eosinophilic airway inflammation;
- To monitor compliance to corticosteroid therapy and effectiveness of treatment (frequency of exacerbations).

3.6.3.2 NIOX VERO®

Since commencing this assessment, Aerocrine have begun launching a new FeNO device which is intended to replace NIOX MINO. The new device is called the NIOX VERO. This is a battery powered device which features a longer operational life and extended test volume life than NIOX MINO.

3.6.3.3 NObreath®

NObreath (Bedfont Scientific Ltd.) is a diagnostic monitoring device that measures FeNO. The reading is presented in ppb and is claimed to be directly related to the severity of inflammatory disease (for example, asthma). NObreath requires 12 seconds of exhalation of breath in adults and 10 seconds in children. NObreath weighs approximately 400g (including batteries). It has a battery life that lasts up to 120 tests. The device is CE marked. The device does not have a finite lifetime although sensor cells require replacement every 2 years.

3.6.4 Anticipated costs associated with intervention

The marginal per-test costs of each of the three technologies considered within this assessment depend on both fixed costs, such as the initial cost of the devices, and variable costs, such as the costs of consumables.

The NIOX MINO device has a unit cost of £2,100 and has an effective unit lifetime of 3-years or 3,000 tests (whichever comes first). The NIOX VERO device has a unit cost of £2,310 and has an effective unit lifetime of 5-years or 5,000 tests (whichever comes first). The NObreath device costs £1,995 and has an unlimited unit lifetime. Maintenance for NObreath is provided free of charge by Bedfont Scientific.

Test kits for NIOX MINO are available in packs of 300 at a price of £1,350, 500 kits at a price of £2,100 or 1,000 kits at a price of £3,950. Test kits for NIOX VERO are available in packs of 300 at a price of £1,500, 500 kits at a price of £2,200 or 1,000 kits at a price of £4,200. Mouthpieces for NObreath are available in packs of 50, 100, 300 or 1,000 at prices of £195, £365, £995 and £2,995 respectively.

The NObreath device requires replacement of the sensor unit every 2-years at a cost of £295. Besides test kits, NIOX MINO and NIOX VERO do not require any further replacement costs.

This information is summarised in Table 2.

Table 2: Cost of equipment and consumables for NIOX MINO, NIOX VERO and NObreath

Item	NIOX MINO	NIOX VERO	NObreath
Lifetime	3 years or 3,000 tests	5 years or 5,000 tests	Unlimited
Equipment cost	£2,100	£2,310	£1,995
Test kits (100)	n/a	n/a	£365
Test kits (300)	£1,350	£1,500	£995
Test kits (500)	£2,100	£2,200	n/a
Test kits (1,000)	£3,950	£4,200	£2,995
Sensor replacement	n/a	n/a	£295
Maintenance	n/a	n/a	Provided free by Bedfont Scientific

4. DEFINITION OF THE DECISION PROBLEM

4.1 Purpose of the decision to be made

The aim of the assessment is to assess the clinical effectiveness and cost-effectiveness of FeNO measurement in people with asthma. This can be separated into two distinct questions:

1. What is the clinical and cost-effectiveness of FeNO testing in the diagnosis of asthma in adults and children?
2. What is the clinical and cost-effectiveness FeNO testing in the management and monitoring of asthma in adults and children?

The cut-off values used in diagnostic technologies affect their sensitivity and specificity and result in different proportions of patients being true-positive (TP), true-negative (TN), false-positive (FP) and false-negative (FN). The consequences of being TP, TN, FP and FN are different in terms of costs and health impacts, hence the highest sum of sensitivity and specificity may not necessarily lead to optimal health outcomes. This is relevant to the use of FeNO in the diagnosis of asthma, and also to its use in guiding asthma management.

4.2 Definition of the scope of the assessment

The scope of this assessment was informed by two scoping workshops attended by specialist committee members, the External Assessment Group (EAG), the manufacturers, NICE and patient stakeholders. The definition of the decision problem reflects the initial NICE scope[25] and the subsequent discussions in the second workshop.

4.2.1 Definition of the interventions

Two monitors were identified at the scoping stage for this appraisal: NIOX MINO® which is manufactured by Aerocrine, and NObreath which is manufactured by Bedfont Scientific. During the latter stages of the assessment, Aerocrine alerted the EAG to a follow-up device to NIOX MINO, NIOX VERO. This device is also considered within this assessment, though the evidence base is limited. All three interventions are evaluated in the context of the diagnosis and management of asthma.

4.2.2 Populations and relevant subgroups

4.2.2.1 Relevant population for assessment of FeNO in the diagnosis of asthma

The population of interest is people with clinical characteristics suggestive of asthma. Relevant subgroups include:

- Any patient 5 years old or older presenting to primary care with symptoms of asthma
- People with clinical characteristics suggestive of asthma who are difficult to diagnose

- Patients who may experience different outcomes from the use of FeNO when compared to the main population under assessment defined as smokers, the elderly and pregnant women.

4.2.2.2 Relevant population for assessment of FeNO in the management of asthma

The population of interest is patients 5 years old or older diagnosed with asthma. There are two subgroups of particular interest:

- Those with good asthma control who are being considered for a dose reduction
- Those with uncontrolled asthma who are experiencing exacerbations or worsening of symptoms, and are being considered for a dose increase of ICS or are being checked for compliance to treatment.

4.2.3 Comparators

The relevant comparators are diagnosis or management according to the current UK guidelines as described in Chapter 3. In the diagnostic setting, the relevant comparator is comprised of the current diagnostic pathway without the use of FeNO measurements; these are different for children and adults (see Section 3.5). In the management setting, the relevant comparator is management according to current guidelines without the use of FeNO.

4.2.4 Relevant outcomes for the assessment

The assessment includes consideration of available evidence across a wide range of clinical and economic outcomes.

Clinical considerations

The intermediate measures for consideration include:

- Diagnostic test accuracy
- Test failure rate

The clinical outcomes for consideration include:

- Asthma control which includes asthma symptoms
- Exacerbation rates – this includes the frequency of exacerbations requiring unscheduled contact with healthcare professionals, visits to accident and emergency departments or hospitalisations.
- Clinical complications associated with acute exacerbations
- Levels of inhaled corticosteroids
- Use of oral corticosteroids
- Adverse effects of treatments (including bronchodilators and steroids)
- HRQoL

- Mortality

Cost considerations

- Cost of equipment, reagents and consumables
- Maintenance and renewal of equipment
- Costs associated with asthma medication
- Cost associated with acute exacerbations
- Cost of further investigations avoided

4.2.5 Place of the intervention in the diagnostic/treatment pathways

During the scoping phase of this appraisal, workshop attendees considered that the interventions should be assessed when added to current practice. There are a number of potential places within the pathways where FeNO may be of clinical use, and each is likely to have different consequences for clinical and cost-effectiveness.

4.2.5.1 Position of FeNO in the diagnostic pathway - children

During the scoping workshop, it was agreed that FeNO is likely to be of most use in Positions 1, 2 and 3 as shown in Figure 7. This figure was based on the BTS/SIGN clinical guidelines, with input from a clinician about how the tests are used in practice (personal communication with Dr John White, 17th July 2013). This equates to patients who are difficult to diagnose. Depending on whether FeNO is used as a direct replacement for a test, or as a rule-in or rule-out test at these positions in the pathway, it may have the ability to prevent expensive secondary care visits if used in primary care. In secondary care, it may have additional value alone or in conjunction with existing secondary care tests. FeNO could also be considered to replace the whole pathway, or be inserted at other points along the pathway. Tables 3, 4 and 5 detail the actions and consequences associated with some different replacement and rule-in/rule-out scenarios. In rule-in scenarios, patients testing positive are assumed to have asthma, and those testing negative go on to have further tests for asthma. In rule-out scenarios, those who test negative are assumed not to have asthma, and those who test positive go on to have further tests for asthma.

Figure 7: Potential positions for FeNO in the diagnostic pathway - children

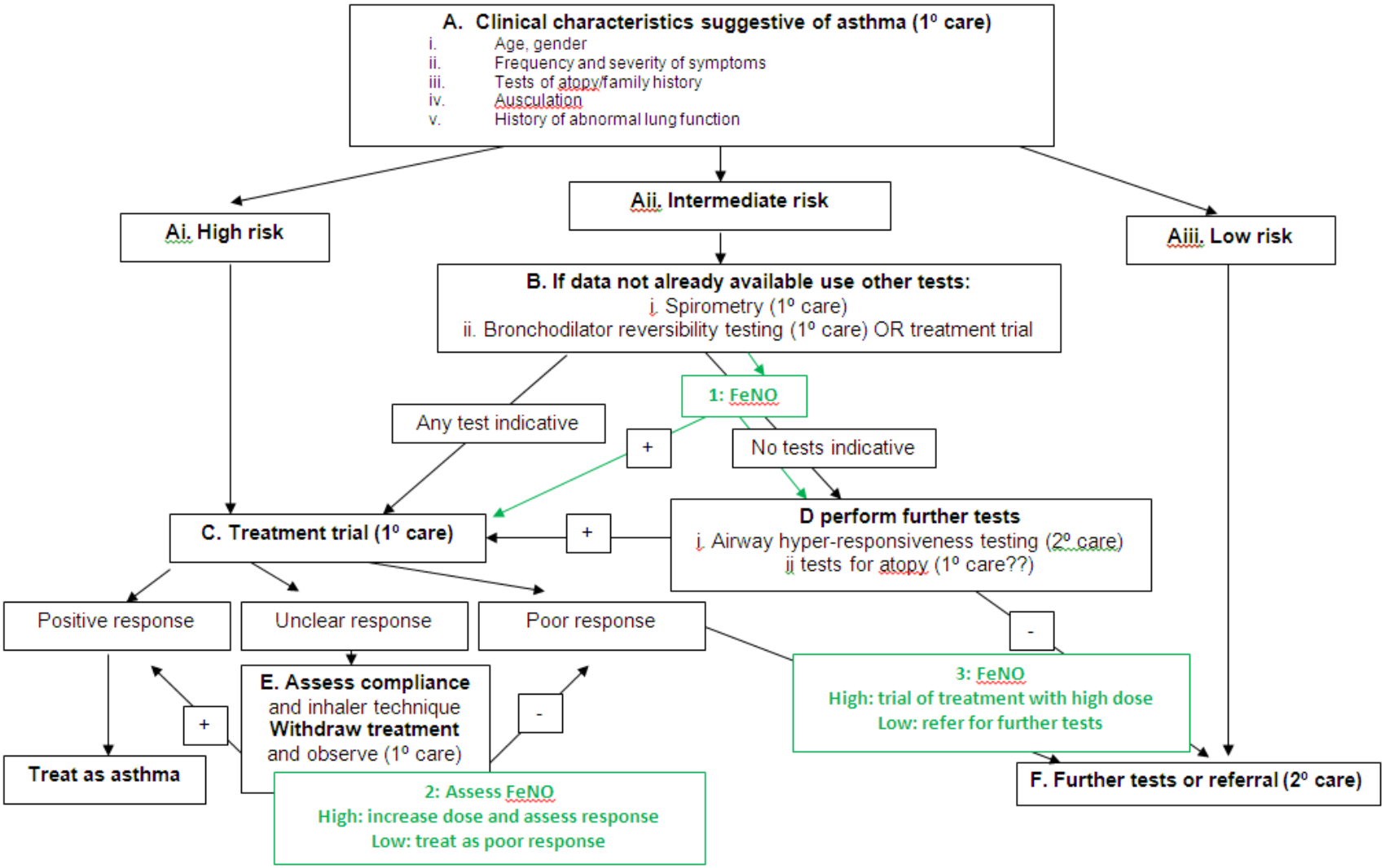


Table 3: Details of consequences of using as a direct replacement for the whole pathway or for airway hyper-responsiveness in patients indicated for this test within the pathway in adults and children

Replacement scenario	FeNO measurement	Action taken	Consequence 1	Consequence 2
TP	High FeNO measurements	Treat as asthma	Correct diagnosis of asthma reached	None
FP			Patients misdiagnosis goes undetected until worsening, routine review or continues lifelong	None
TN	Low FeNO measurements	Treat as not asthma	Further tests for other conditions	Correct diagnosis reached
FN				Further tests negative, re-enter asthma pathway, or remain misdiagnosed until exacerbation or return to GP with ongoing symptoms.

Table 4: Details of consequences of using as a rule-out test before airway hyper-responsivness in adults and children

Rule-out scenario	FeNO measurement	Action taken	Consequence 1	Consequence 2
TP	High FeNO measurements	Treat as may be asthma and give further confirmatory tests	Further tests confirm asthma diagnosis	Treat as asthma
FP			Further tests reject asthma diagnosis	Further tests for other conditions or diagnose as non-specific symptoms
TN	Low FeNO measurements	Treat as not asthma	Further tests for other conditions	Correct diagnosis reached
FN				Further tests negative, re-enter asthma pathway, or remain misdiagnosed until exacerbation or return to GP with ongoing symptoms.

Table 5: Details of consequences of using as a rule-in test before airway hyper-responsiveness in adults and children

Rule-in scenario	FeNO measurement	Action taken	Consequence 1	Consequence 2
TP	High FeNO measurements	Treat as asthma	Correct diagnosis of asthma reached	None
FP			Patients misdiagnosis goes undetected until worsening, routine review or continues lifelong	None
TN	Low FeNO measurements	Further tests for asthma	Tests for asthma negative	Further tests for other conditions or diagnose as non-specific symptoms
FN			Correct diagnosis of asthma reached	None

4.2.5.2 Position of FeNO in the diagnostic pathway - adults

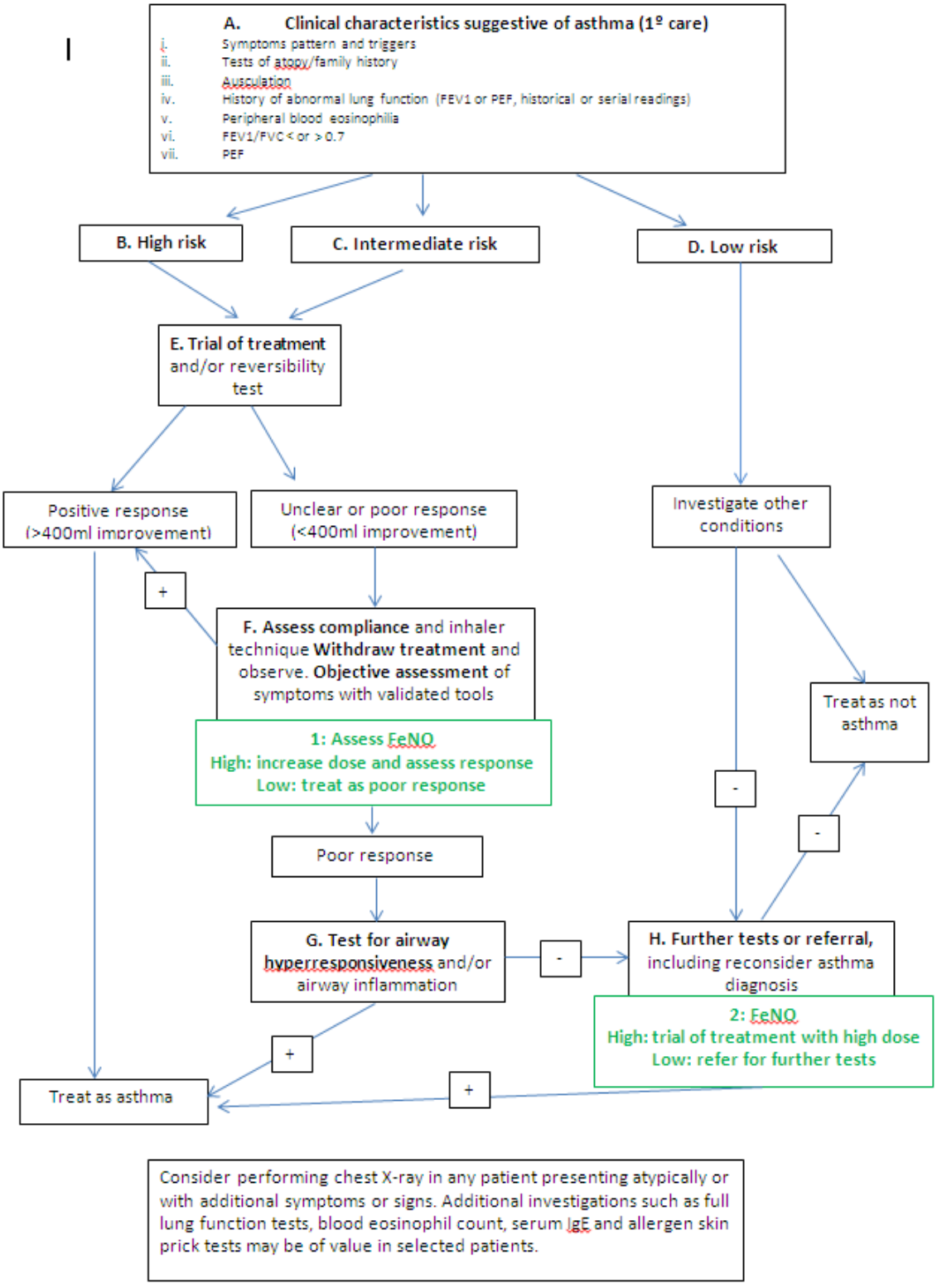
FeNO is thought to be of most use in Positions 1 and 2 as shown in Figure 8. This equates to patients who are difficult to diagnose. This figure was based on the BTS/SIGN clinical guidelines, with input from a clinician about how the tests are used in practice (personal communication with Dr John White, 17th July 2013). This lead to the understanding that in nearly all or at least most cases, patients would undergo a trial of treatment or airway reversibility testing before being referred to secondary care, regardless of their FEV_1/FVC ratio. This is slightly different to our initial reading of the BTS/SIGN guidelines, where only patients with FEV_1/FVC<0.7 would undergo these tests, whilst those with FEV_1/FVC>0.7 go on to secondary care for airway hyper-responsiveness testing. Our initial diagrammatic representation of the adult pathway is included in the protocol (Appendix 1) for comparison.

Depending on whether FeNO is used as direct replacement for an exisiting test, or as a rule-in or rule-out test at these positions in the pathway, it may have the ability to prevent expensive secondary care visits if used in primary care. In secondary care, it may have additional value alone or in conjunction with existing secondary care tests. FeNO could also be considered to replace the whole pathway, or be inserted at other points along the pathway.

4.2.5.3 Position of FeNO in the management pathway

FeNO measurement may be helpful in individuals diagnosed with asthma to facilitate titration of corticosteroid therapy, to check for compliance with medication, and ultimately to lead to better asthma control. It is likely that management decisions would be based on a combination of the monitoring information collected at review and FeNO measurements. In these scenarios, high levels of FeNO could indicate that a patient's asthma is not fully controlled and may be interpreted in combination with symptoms and medication use. A lack of control could be due to worsening of the disease, or could be due to failure to comply with medication. The latter could be ascertained through additional checks on collection of scripts or number of doses used as measured by a dose-counter inhaler. Low FeNO could indicate that asthma is well controlled and may be interpreted in combination with symptoms and medication use, and could guide step-down of medication and subsequent monitoring of control.

Figure 8: **Potential positions for FeNO in the diagnostic pathway – adults.** Diagram based on BTS/SIGN guidelines[8] with clinical input from Dr John White (personal communication, 17th July 2013).



4.3 Structure of the assessment report

The assessment report is comprised of two main parts: (1) an assessment of the clinical evidence relating to FeNO in the diagnosis and management of asthma, and; (2) an assessment of the cost-effectiveness of FeNO versus standard care in the diagnosis and management of asthma.

4.3.1 Clinical evidence review

Two systematic reviews and one rapid review were conducted concurrently to identify clinical evidence relevant to the decision problem.

- *Rapid review of equivalence of FeNO devices.* It was not clear at the outset if there would be sufficient primary research evidence relating to the three devices to inform the appraisal. As such, a review of the equivalence of these devices to other FeNO measurement devices was anticipated, and appropriate searches conducted. The review of equivalence was conducted in full when it became apparent that sufficient evidence was not available from the diagnostic accuracy review and management efficacy review. The equivalence review aimed to establish whether measurements from different FeNO measurement devices could be considered to be equivalent to one another, and so whether studies that used other devices could helpfully inform this appraisal. This review was thought to be the least critical in terms of informing key model inputs and a rapid review using systematic methods was therefore conducted due to time and resource constraints, this represents a change to the published assessment protocol (Appendix 1).
- *Systematic review of diagnostic accuracy of FeNO for asthma.* As described in the protocol (Appendix 1, Section 3.1.5), the ideal study design would recruit patients with symptoms of asthma, have a cohort design or randomise them to diagnosis using FeNO or diagnosis using other methods and follow them to clinical outcomes. Such studies are known as end-to-end studies and demonstrate the ability of the test to improve patient outcomes. In the absence of such studies, diagnostic cohort studies represent the next best level of evidence, with modelling of clinical outcomes based on the numbers of patients classed as true-positive, true-negative, false-positive and false-negative. Below this are correlation studies. All levels of evidence were searched for in this review; lower levels of evidence were consulted where the higher levels of evidence were not identified. Where available, three pairs of sensitivity and specificity were selected: those that produced the highest sum of sensitivity and specificity; those that had the highest sensitivity for rule-in scenarios; those that had the highest specificity for rule-out scenarios. In rule-in scenarios, patients testing positive are assumed to have asthma, and those testing negative go on to have further tests for asthma. In rule-out scenarios, those who test negative are assumed not to have asthma, and those who test positive go on to have further tests for asthma.

- *Systematic review of the efficacy of FeNO-guided management of asthma*. Existing systematic reviews of RCT evidence in adults (Petsky 2010)[26] and in children[27,28] meant that only RCT evidence was searched for in this review, with additional interrogation of the database for data on subgroups where RCT evidence was not found.

4.3.2 Cost-effectiveness assessment

The cost-effectiveness assessment of FeNO includes two components: a systematic review of existing economic analyses and the development of two *de novo* health economic models.

- *Systematic review of the cost-effectiveness of FeNO for the diagnosis and/or management of asthma*. A systematic review was undertaken to identify all existing economic analyses of FeNO testing for asthma; this includes published studies as well as evidence submitted by the manufacturers of NIOX MINO, NIOX VERO and NObreath. The section includes a critical appraisal of the available evidence and a summary of methodological problems and concerns relating to these analyses.
- *Development of two de novo models.* Independent health economic models were developed to assess the incremental cost-effectiveness of FeNO versus standard care in the diagnosis and management of asthma.

5. CLINICAL REVIEW

5.1 Methods

As described in Section 4.3.1 above, two systematic reviews and one rapid review were conducted concurrently to identify clinical evidence relevant to the decision problem.

1. Rapid review of equivalence of FeNO devices.
2. Systematic review of diagnostic accuracy of FeNO for asthma.
3. Systematic review of the efficacy of FeNO-guided management of asthma.

5.1.1 Clinical reviews search methodology

Systematic searches were carried out between March 2013 and April 2013. For the review of device equivalence, and for both diagnostic and management reviews, the following databases were searched:

- MEDLINE and MEDLINE In Process: Ovid. 1948-present
- EMBASE: Ovid. 1974-present
- Cochrane Library
    - Cochrane Database of Systematic Reviews (CDSR)1996-present
    - Database of Abstracts of Reviews of Effects (DARE) 1995-present
    - Cochrane Central Register of Controlled Trials (CCRCT) 1898-present
    - Health Technology Assessment Database (HTA) 1995-present
    - NHS Economic Evaluation Database (NHS EED) 1995-present
- Science Citation Index Expanded (SCIE): Web of Science 1899-present
- Conference Proceedings Citation Index – Science (CPCI-S): Web of Science. 1990-present

The following trial registers and websites were searched in March 2013 for all three reviews (search terms used are provided in Appendix 2):

- ClinicalTrials.gov http://clinicaltrials.gov/
- metaRegister of Controlled Trials http://www.controlled-trials.com/mrct/
- FDA Manufacturer and User Facility Device (MAUDE) http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm
- EuroScan International Network http://euroscan.org.uk/

5.1.1.1 Management review searches

Searches for the management review were developed following the identification of a 2009 Cochrane review.[26] Study design filters were not applied to the strategy in case lower levels of evidence were needed for the subgroups defined *a priori* in the protocol. The strategy is made up of free-text terms for NIOX MINO and NObreath including manufacturer names, subject headings and free-text terms

for asthma (e.g. respiratory hypersensitivity, bronchoconstriction) and lower respiratory tract symptoms (e.g. coughing, wheezing, chest pain) that were combined with terms for exhaled nitric oxide (e.g. feno, eno). Searches were limited to publications since 2009.

A summary of the search records retrieved from the searches are provided in Table 6.

Table 6: Search records retrieved by database: management review

Database	**Management review searches**
MEDLINE and MEDLINE In Process	991
EMBASE	2269
CDSR	44
DARE	1
CCRCT	117
HTA	8
SCIE	1387
CPI-S	70
Total unique references	2747

CDSR, Cochrane Database of Systematic Reviews; DARE, Database of Abstracts of Reviews of Effects; CCRCT, Cochrane Central Register of Controlled Trials; HTA, Health Technology Assessment;SCI, Science Citation Index Expanded; CPI-S, Conference Proceedings Citation Index – Science.

5.1.1.2 Diagnostic review searches

Similar to the management review search strategy, the diagnostic search comprises terms for NIOX MINO and NObreath including manufacturer names, subject heading and free-text terms for asthma and lower respiratory tract symptoms combined with terms for exhaled nitric oxide (see Figure 9 below). The strategy was combined with three filters 1) systematic reviews filter 2) an RCT filter 3) a diagnostic filter. No date limits were applied to the searches.

A summary of the search records retrieved from the searches is given in Table 7.

Table 7: Search records retrieved by database: diagnostic review

	Search by study design			Equivalence review
Database	**Systematic reviews**	**RCTs**	**Diagnostic studies**	
Medline	26	958	377	97
Embase	114	1386	452	282
CDSR	44	-	-	0
DARE	1	-	-	0
CCRCT	-	509	-	10
HTA	8	-	-	4
NHS EED	2	-	-	1
SCIE	76	637	284	92
CPI-S	3	17	10	8
Total unique references	**227**	**1635**	**680**	**309**

CDSR, Cochrane Database of Systematic Reviews; DARE, Database of Abstracts of Reviews of Effects; CCRCT, Cochrane Central Register of Controlled Trials; HTA, Health Technology Assessment;SCI, Science Citation Index Expanded; CPI-S, Conference Proceedings Citation Index – Science.

5.1.1.3 Equivalence of devices review searches

The analytic validity studies search for NIOX MINO and NObreath was carried out using terms for NIOX MINO, NObreath and the manufacturer names without any application of filters and limits in the database listed. The number of records retrieved are included in Table 7 (final column).

Figure 9: Management and clinical utility studies search of FeNO

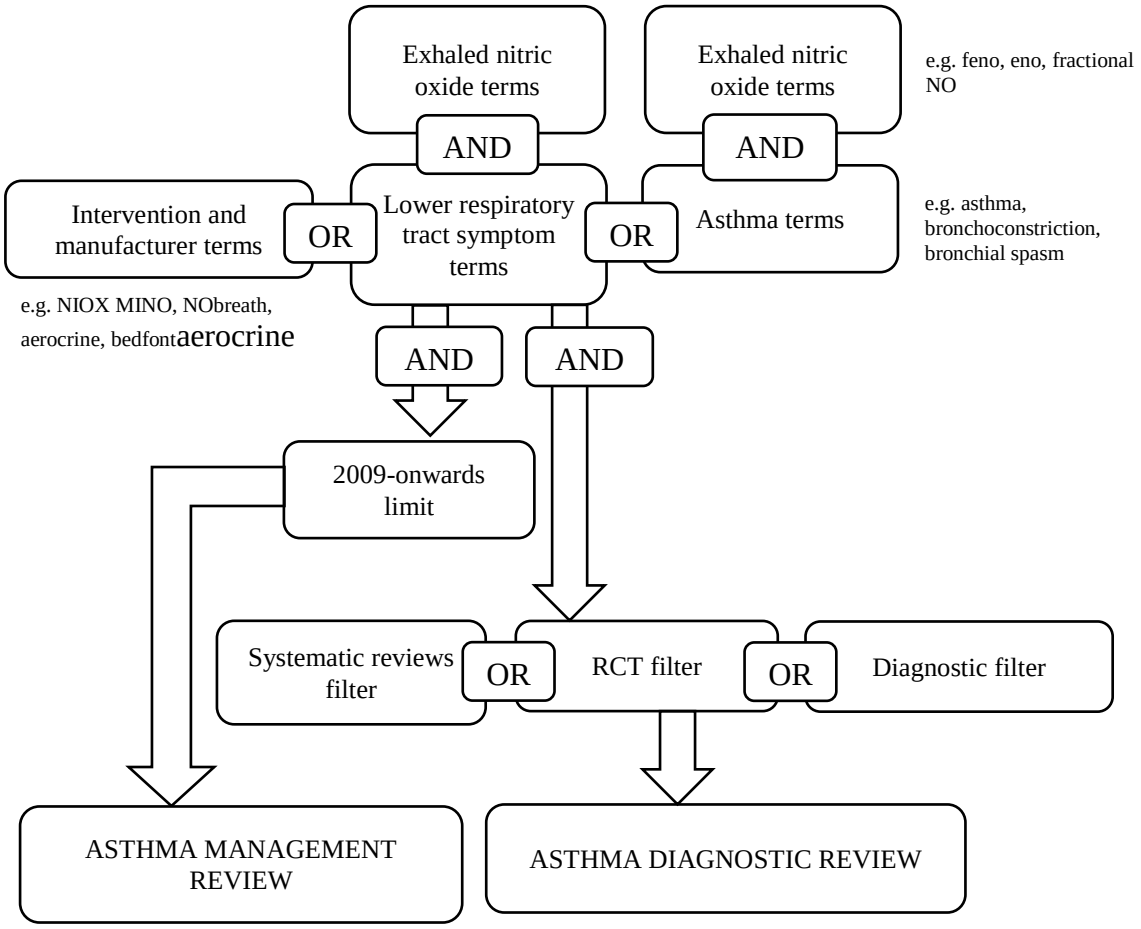
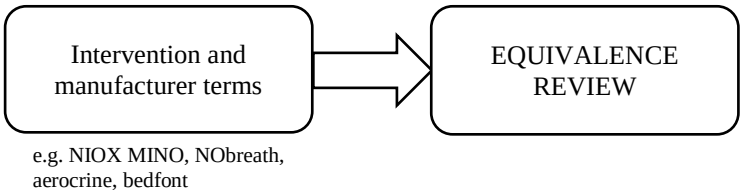


Figure 10: Equivalence studies of FeNO devices



5.1.1.4 Additional search for NIOX VERO

Aerocrine's new device, NIOX VERO, will be launched shortly and was brought to the attention of the ERG in July 2013. An additional search was conducted on 13th August 2013 to check for any publications relating to this device that would have been missed by the original search. This search comprised simply the term "Niox Vero" across all previously listed databases. A summary of the search records retrieved from the searches is given in Table 8.

Table 8: Search records retrieved by database: NIOX VERO review

Database	Niox Vero searches
MEDLINE and MEDLINE In Process	0
EMBASE	0
CDSR	0
DARE	0
CCRCT	0
HTA	0
NHS EED	0
SCIE	0
CPI-S	0
ClinicalTrials.gov	0
metaRegister of Controlled Trials	0
MAUDE	0
EuroScan International Network	2

5.1.1.5 Reference management

All retrieved citations were downloaded into Reference Manager® bibliographic software[29] and deduplicated to include only unique citations.

5.1.2 Study selection

Retrieved citations were considered for inclusion in several stages. Firstly, titles were considered and any studies obviously not relevant were excluded. Secondly, abstracts were consulted. At this stage, tags were applied to studies in Reference Manager® to identify the device used, the age group of the participants and the study design. In instances whereby it was obvious which review the study was likely to inform, this tag was also applied. In the third stage, articles tagged as the highest levels of evidence for each review were retrieved and the full text was obtained for comparison against the inclusion and exclusion criteria.

Once the full text selection process was complete, a decision was made as to whether there were gaps in the evidence that would require lower levels of evidence to be consulted. This was the case for the diagnostic review, where no end-to-end studies were identified; for the management review, where only limited evidence was identified using NIOX MINO and no evidence was identified relating to NObreath; and for some of the subgroups of interest to the review. For the diagnostic review, studies including any device were included rather than just those using NIOX MINO, NIOX VERO or NObreath (see Section 5.1.2.2); for the management review studies using any FeNO device were included (see Section 5.1.2.3); and the rapid review of the equivalence of devices was conducted in full (see Section 5.1.2.1). In order to retrieve relevant titles from the database for the subgroups of interest to the review, the search facility in Reference Manager® was used to search all indexed fields for the following terms:

Elderly asthmatics: elderly, old, older and elderly care.

Smokers: smoke, smoking, smoking .adverse effects, smoking .epidemiology, smoking cessation, smoking cessation programme, smoking habit, smoking/ae [Adverse Drug Reaction] and smoking: epidemiology.

Pregnant women: pregnant, pregnancy, expectant, pregnancy complication/co [Complication], pregnancy complication/si [Side Effect], pregnancy complications, pregnancy diabetes mellitus, pregnancy diabetes mellitus/dt [Drug Therapy], pregnancy outcome, pregnancy test and pregnant women.

These titles were then sifted by title, abstract and full text for inclusion in the review with relation to criteria for population, intervention and comparator. Criteria on study design and specific outcomes were relaxed, and studies of the next best level of evidence which provided data evaluating the use of FeNO measurements in appropriate subgroups were included. The hierarchy of evidence used was as described in the NICE guidelines methods guide.[30]

5.1.2.1 Review of equivalence of devices

Table 9 describes the inclusion and exclusion criteria for this review.

Table 9: Inclusion and exclusion criteria for the review of equivalence of devices

	Inclusion	Exclusion	Change from protocol
Population	Studies conducted in humans only, regardless of asthmatic status or recruitment methods	Studies performed *in vitro* on gas samples unless no test evidence was found in humans.	None
Primary device	NIOX MINO, NIOX VERO or NObreath operated in accordance with American Thoracic Society (ATS) 2005 guidelines:[31] • Expiratory flow rate of 50mL per second (0.05L/sec) • An exhalation time of ≥ 10 seconds for adults, ≥ 6 seconds for children		None
Comparator	Other chemiluminescent devices operated in accordance with ATS 2005 guidelines: • Expiratory flow rate of 50mL per second (0.05L/sec) • An exhalation time of ≥ 10 seconds for adults, ≥ 6 seconds for children If no studies at this flow rate and exhalation time were found, any flow rate or exhalation time was be included.		None
Outcomes	Studies of analytic validity were included if they report the ability of the test to measure FeNO accurately, by any statistical method, as compared to chemiluminescent devices.	Studies of inter-rater reliability or inter-subject repeatability were excluded.	None
Study design	Any		None

5.1.2.2 Review of diagnostic accuracy of FeNO for asthma

Table 10 describes the inclusion and exclusion criteria for this review and any differences from the published protocol (See Appendix 1). At full text sift stage, some unforeseen questions about the scope were sent to specialist committee member (SCM) clinicians for clairification. This is documented in Appendix 3.

Table 10: Inclusion and exclusion criteria for the review of diagnostic accuracy

	Inclusion	Exclusion	Change from protocol
Population	Primary population are patients presenting with clinical characteristics suggestive of asthma. The main relevant subgroups within this population are: • Those presenting with clinical characteristics suggestive of asthma and who are difficult to diagnose. • Women during pregnancy • Older people • Smokers Studies were included if they recruited a wider population but report *a priori* subgroup analyses for the populations of relevance to this review.	Children <5 years old Studies that recruited a wider population and did not report *a priori* subgroup analyses for the populations of relevance to this review Animal models Unselected specific population (e.g fire fighters, obese, athletes)	None
Intervention	Use of NIOX MINO or NObreath in the diagnosis of asthma, either with or without another test. NIOX MINO and NObreath devices are set to record according to ATS 2005 criteria: • Expiratory flow rate of 50mL per second	Expiratory flow rate not 50mL per second (0.05L/sec) An exhalation time of < 10 seconds in adults, < 6 seconds in children	The protocol stated that studies using the following cut-off would be included: • FeNO less than 25 ppb (< 20 ppb in children) indicates that eosinophilic inflammation and responsiveness to

	Inclusion	Exclusion	Change from protocol
	(0.05L/sec) • An exhalation time of ≥ 10 seconds for adults, ≥ 6 seconds for children If data were not available for the above interventions, studies were included if they reported clinical validity of FeNO measured by any chemiluminescent device with appropriate measurement methods: • Expiratory flow rate of 50mL per second (0.05L/sec) • An exhalation time of ≥ 10 seconds for adults, ≥ 6 seconds for children • Online measurement Studies that did not report any of these details were included and discussed in the narrative review. Studies using any cut-off value or combination of cut-off values were included	Offline measurements Alveolar nitrogen oxide or nasal nitrogen oxide measurements	corticosteroids are less likely • FeNO greater than 50 ppb (>35 ppb in children) indicates that eosinophilic inflammation and, in symptomatic patients, responsiveness to corticosteroids are likely • FeNO values between 25 ppb and 50 ppb (20–35 ppb in children) should be interpreted cautiously and with reference to the clinical context However, no studies using these exact cut-off values and protocol were found, so all cut-off values were included.
Comparator	Comprises any combination or selection of the tests and clinical characteristics described in the UK guideline for	Uses tests to diagnose asthma that are not included in the BTS/SIGN	Studies using tests not in routine use in the UK but mentioned in the BTS/SIGN

	Inclusion	Exclusion	Change from protocol
	the diagnosis of asthma.	guidelines or if the comparator includes the use of FeNO measurements.	guidelines were included in the review
Outcome	End-to-end studies – include studies with relevant clinical outcomes (see management study outcomes) Clinical validity studies - Include studies that report data that allow the extraction of the numbers of patients who are true-positive, true-negative, false-positive and false-negative against the reference standard. Studies which report test failure rates were also included.	Does not report useable diagnostic validity data (i.e. extraction of the numbers of patients who are true-positive, true-negative, false-positive and false-negative against the reference standard).	None
Study design	End-to-end studies – (which follow patients from diagnostic test to clinical outcomes) If no evidence was found at this level, clinical validity studies (which compare the diagnosis of patients by the intervention with a reference standard) were included. These should be prospective cohort studies, cross sectional studies or retrospective cohort studies. If studies of these designs were not located, other study	Preclinical and biological studies Editorials and opinion pieces Studies only published in languages other than English	Studies published as abstracts not reporting sufficient methodological details to allow critical appraisal of study quality were NOT excluded.

	Inclusion	Exclusion	Change from protocol
	designs were considered (e.g. case control studies). Both studies deriving cut-off values for diagnosis and studies validating existing cut-off values for diagnosis will be included. Abstracts with comparable data that do not exist in full published studies		
Setting	Primary care, secondary care, out-patient clinic or specialist clinic	Emergency care diagnosis of exacerbation	None
FeNO, fractional exhaled nitric oxide; ppb, parts per billion; ATS, American Thoracic Society; BTS, British Thoracic Society; SIGN, Scottish Intercollegiate Guidelines Network.			

5.1.2.3 Review of the efficacy of FeNO-guided management of asthma

Table 11 describes the inclusion and exclusion criteria for this review and any differences from the published protocol (Appendix 1).

Table 11: Inclusion and exclusion criteria for the review of FeNO guided management of asthma

	Inclusion	Exclusion	Change from protocol
Population	Patients diagnosed with asthma. The two subgroups of particular interestwere: • Those with good asthma control who are being considered for a dose reduction • Those with uncontrolled asthma who are experiencing exacerbations or worsening of symptoms, and are being considered for a dose increase of ICS or are being checked for compliance to treatment And further subgroups within each of these categories: • Women during pregnancy • Older people • Smokers Recruit whole asthma populations or if they recruit patients exclusively from any of the subgroups.	Children <5 years old Recruited patients not diagnosed with asthma Animal models Unselected specific population (e.g. firefighters, obese, athletes)	None

	Inclusion	Exclusion	Change from protocol
Intervention	Use of NIOX MINO or NObreath in the diagnosis of asthma, either with or without another test. NIOX MINO and NObreath devices are set to record according to ATS 2005 criteria: • Expiratory flow rate of 50mL per second (0.05L/sec) • An exhalation time of ≥ 10 seconds for adults, ≥ 6 seconds for children If data were not available for the above interventions, studies were included if they reported clinical validity of FeNO measured by any chemiluminescent device with appropriate measurement methods: • Expiratory flow rate of 50mL per second (0.05L/sec) • An exhalation time of ≥ 10 seconds for adults, ≥ 6 seconds for children • Online measurement Studies that did not report any of these details were included and discussed in the narrative review . Studies monitoring at intervals of greater than 2 weeks were included.	Device which is not validated for measuring FeNO Expiratory flow rate not 50mL per second (0.05L/sec) An exhalation time of < 10 seconds. Offline measurements Studies where FeNO is measured on a more regular basis (i.e. not during routine annual review) were excluded.	Studies that did not report any details about device or measurement methods were included and discussed in the narrative review . The protocol stated "Only studies using FeNO measurements in: • routine annual monitoring • dose titration indicated during routine monitoring • assessment of compliance will be included in the review." However, no such studies were located, so studies monitoring at intervals of greater than 2 weeks were included.

	Inclusion	Exclusion	Change from protocol
	Any protocols and cut-off values for management decisions or compliance monitoring will be included.		
Comparator	Studies comparing the interventions to any other management strategy that does not utilise FeNO measurements were included. Studies using management strategies that closely match all or part of UK practice as described in the UK guidelines will be included. Where no studies which closely match UK practice werefound, studies using other management strategies wereincluded.	Includes the use of FeNO measurement as part of the management strategy	None
Outcome	Incidence of acute exacerbations, including those requiring unscheduled contact with healthcare professionals, visits to accident and emergency departments or hospitalisations. Other measures (time-to-event data; numbers of patients experiencing an exacerbation) were only considered if insufficient data were available for the rate of exacerbations. Any definition of exacerbation was acceptable.	Does not report data on FeNO guided step up step down therapy Measure of alveolar nitrogen oxide or nasal nitrogen oxide	None

	Inclusion	Exclusion	Change from protocol
	Asthma control which includes asthma symptoms, either reported individually or by use of a standardised patient outcome measure or symptom score. Clinical complications associated with acute exacerbations Levels of inhaled corticosteroids Use of oral corticosteroids Adverse effects of treatments (including bronchodilators and steroids) Health-related quality of life Mortality Compliance		
Study design	Randomised controlled trials. If insufficient RCT evidence is identified, other study designs will be included according to the hierarchy of evidence for efficacy trials Abstracts with comparable data that do not exist in full published studies and sufficient methodological details were included.	Preclinical and biological studies Editorials and opinion pieces Studies only published in languages other than English	Studies published as abstracts not reporting sufficient methodological details to allow critical appraisal of study quality were NOT excluded.
Setting	Primary care, secondary care, out-patient clinic or specialist clinic	Emergency care	None

FeNO, fractional exhaled nitric oxide; ATS, American Thoracic Society.

5.1.3 Data extraction

A different standardised data extraction form was developed for each review following the guidelines given in the CRD handbook for systematic reviews and the Cochrane Handbook[32,33] and piloted using two studies per review. Missing fields were added as appropriate and back–filled where necessary. Appendix 4 lists the fields that were data extracted for each review. Data were extracted from the studies by one of three reviewers and was then checked by a second reviewer (SH, ME, TG), except in the rapid review of equivalence of devices where a sole reviewer (SH) extracted all relevant data. Any discrepancies were resolved by discussion, with involvement of a third reviewer when necessary. Where appropriate, authors were contacted for missing or unclear data. Data from multiple publications of the same study were extracted and quality assessed as a single study. In a change from the protocol, data were not extracted from existing systematic reviews, but directly from the primary research journal articles and conference abstracts.

5.1.4 Quality assessment

As a rapid review, quality assessment was not conducted for the review of equivalence of devices.
Diagnostic cohort studies were assessed using Quality Assessment of Diagnostic Accuracy Studies II (QUADAS II[34]). The tool was adapted to the specifics of this appraisal and the scoring scheme can be found in Appendix 5.

Management RCT studies were assessed using domains listed in the Cochrane risk of bias tool. The scoring scheme can also be found in Appendix 5.

Studies of lower quality were not formally quality assessed but were considered on their individual merits.

Quality assessment was conducted by one reviewer and checked by a second. A third reviewer was consulted in cases of disagreement.

5.1.5 Analysis and synthesis

A narrative synthesis was conducted for the rapid review of the equivalence of devices and no meta-analysis was planned or attempted.

A narrative synthesis was conducted for the review of diagnostic studies. A meta-analysis was planned where sufficient studies of acceptable clinical heterogeneity in terms of patient populations, devices, cut-off points and reference standards were available. A meta-regression to allow the use of multiple cut-off points in the modelling was planned, again, should the necessary data be available

with appropriate levels of heterogeneity between studies. However, data were not suitable for meta-analysis or meta-regression.

A narrative synthesis was conducted for the review of management studies. A meta-analysis was planned where enough studies of acceptable clinical heterogeneity in terms of patient populations, devices, cut-off points, interventions, comparators and outcomes were available. Clinical heterogeneity indicated that such an analysis was unlikely to produce meaningful results, but exploratory analyses and sensitivity analyses to elements of study design were conducted in the review of adults, even though clinical heterogeneity was high. For rate outcomes, the generic inverse variance method was used in Review Manager®[35] to meta analyse rate ratios. For continuous outcomes, a standardised mean difference analysis was conducted as metrics for ICS use were different.

In all cases, fixed effects were used first, and random effects applied if the I^2 statistic indicated that heterogeneity was moderate or high. This was judged to be the case at >40%.

5.2 Results

A total of 4861 citations were retrieved and considered for inclusion in the review. After scrutiny of the titles and abstract, 4436 studies were excluded and the full text of 425 citations were obtained and consulted. Of these, 347 were excluded (see Appendix 6) and 62 studies (78 citations) were included in the review.

For the review of subgroups, a total of 162 citations were identified of which 15 studies were included. Appendix 7 summarises the process of identifying and selecting relevant literature.

No end-to-end studies were identified within the review. As previously described, a review of the equivalence between FeNO devices was undertaken, alongside a review of diagnostic validity (cohort study design) and management (RCT study design), with data for subgroups of interest to the review from lower levels of evidence where necessary. This report considers each review separately in the following order:

- Rapid review of equivalence of devices (analytic validity, Section 5.2.1)
- Systematic review of diagnostic studies (diagnostic validity, Section 5.2.2)
- Systematic review of management studies (Section 5.2.3)

5.2.1 Equivalence of devices (analytic validity)

A total of 27 studies (32 citations) comparing the intervention devices (NIOX MINO, NIOX VERO and NObreath) to other devices were included in the review. One additional study, Olaguibel 2011[36] was excluded as it compared NIOX MINO to another handheld device (NoVario) not in the scope of this appraisal. The studies have been categorised for presentation and discussion according to the devices compared and population age ranges as follows:

- NIOX MINO versus Niox in adults
- NIOX MINO versus Niox in children
- NIOX MINO versus other stationary chemiluminescent devices in adults and/or children
- NIOX VERO versus NIOX MINO
- NObreath versus other stationary chemiluminescent devices in adults and/or children
- NIOX MINO versus NObreath in adults and/or children
- AUCs, cut-off points and correction equations
- Test failure rates
- Conclusions

Three main comparisons were considered in this review:

- **Comparison of means:** comparison between reported mean FeNO values as measured by each device in the same cohort. This comparison may be confounded by natural within-patient variance between measurements by the two devices.
- **Correlation coefficients:** these show whether measurements by the two devices are correlated, but not whether the actual values produced are the same (agreement). Highly correlated devices might produce slopes on a graph (plotting FeNO measurement against a known FeNO concentration) of the same gradient, but at different heights, indicating that one device measures consistently higher or lower than another. Correlation co-efficients can be confounded by the fact that comparison over wider ranges of values can lead to higher correlation values.[37]
- **Bland-Altman analysis:** produces a number of useful comparison statistics which assess agreement between devices rather than just correlation. Bland-Altman plots[37] plot the mean of two measurements by two devices (x axis) against the difference between the measurements (y axis). If the devices agreed perfectly across the whole range of measurements, all points would be at point zero on the y axis across the range of measurements. However, if agreement is not perfect, the points will fall above and below zero. If there is a systematic bias in results, such as one device consistently reading higher than the other, the mean of the points will be clustered either above or below zero on the y axis, and this would be evident both visually

and by the mean difference value produced. If this deviation is consistent and can be relied upon, readings between devices could be corrected by subtracting or adding the mean difference. However, if there is also variance in the difference between devices, points will be more scattered, and will produce a wider "limit of agreement", which is calculated as ±2SD. If this limit of agreement is wide by clinical standards, it may be concluded that devices are not clinically interchangeable, even if the mean difference is relatively small.

5.2.1.1 NIOX MINO versus Niox in adults

Eight studies compared NIOX MINO to Niox (Table 12); five studies were exclusively in adults,[38-42] and three studies were undertaken in a mix of adults and other age groups.[43-45] When considering the mean values recorded in each study, NIOX MINO was found to give largely similar results to Niox in five studies,[38,39,42,44,45] but was found to provide higher FeNO readings in three other studies (range 0.5 to 9ppb).[39,41,43] One study[39] tested two NIOX MINO devices side by side and found that the mean FeNO recorded was higher in one device compared to Niox than in the second NIOX MINO device. Another study[44] tested three devices, and found excellent correlation between the devices, and no statistically significant difference between them. This may indicate that there is some variation between NIOX MINO devices themselves, which may account for some of the heterogeneity in estimates of equivalence in other studies versus other devices. In summary:

- Where cohort mean FeNO values were below 30ppb as measured by Niox, studies showed small differences between the cohort means of devices,[38,39,45] whereas where the mean FeNO was above 35ppb as measured by Niox, larger and statistically significant differences in cohort means were seen.[41,43]
- Correlation coefficients ranged from 0.73 to 0.998
- Bland-Altman analyses were not reported in a consistent way, with some studies using proportions, some using absolute values and some using log values. It is not clear whether log transformation is appropriate as results varied across studies and were apparently conflicting on this point. Where the relationship between devices was multiplicative, differences between devices became greater at higher values. Studies saw limits of agreement (where reported on the absolute scale) of around 10ppb in both directions.[44,45] These large limits of agreement may be due to an assumption that the relationship is additive rather than multiplicative. The difference in % reported by Korn 2010[40] is large, with limits of agreement of -46% to 73% and it is assumed that a log transformation was performed. However, the log values reported by Menzies 2007[38] indicate tighter limits of agreement, but the Bland-Altman[37] plot did not suggest a multiplicative relationship on the absolute scale. It is therefore unclear if the upper and lower limits of agreement between devices is of clinical importance, and whether this is a

multiplicative or additive relationship. It seems likely that a range of 20ppb could be important even at high FeNO values.

5.2.1.2 NIOX MINO versus Niox in Children

Three studies compared NIOX MINO to Niox in children (Table 13). All cohorts included children with asthma.

- Two studies[46,47] reported statistically significantly higher mean FeNO values with NIOX MINO, whilst one study[48] reported statistically significantly lower values. This study had low mean values (below 10ppb)
- All studies reported good correlation between the devices
- Bland-Altman statistics were reported in two studies[47,48] and indicated that NIOX MINO gave higher readings in both cases, by 1.1ppb, (limits of agreement -4.4 to 6.7) and 3.9 ppb, (limits of agreement -1.1 to 8.9) respectively.

Table 12: Equivalence review: NIOX MINO versus Niox in adults, adolescents and adults or all ages

Author name, year	Spons-ored?	Popul-ation	ATS 2005?	N[a]	Mean FeNO NIOX MINO	Mean FeNO chemiluminescence	Comparison data	Correlation coefficient	Bland-Altman	Interpretation
Adults – Asthmatics or mixed including some asthmatics										
Menzies 2007[38]	Yes	As	Yes	101	GM (SE): 26.6 (24.5–28.9)ppb	GM (SE): 26.9 (24.8–29.6)ppb	NR	Pearson: r = 0.94, p<0.001 Spearman: lack of bias at either end of values	Log scale: -0.0 (0.2 to -0.2)[b] Bigger differences at higher values[c, d]	Similar (SSNR)
Grob 2008[39]	NR	As	NR	1	32.5ppb	26.9ppb	NIOX MINO device A: no stat difference NIOX MINO device B: 7.2ppb greater (p=0.26)	NIOX MINO device A: r^2=0.73, p<0.0001 NIOX MINO device B: r^2 = 0.74, p<0.0001		NIOX MINO device A: similar (NSD) NIOX MINO device B: trend to read higher (NSD)
Korn 2010[40]	NR	As COPD He	Yes	85	Median (95% CI):16.3ppb (5.0 to 208.3)	Median (95% CI): 14.5ppb (0 to 196.6)	NIOX MINO 9%(SEM 3%) higher (range -32% to 38%)	Spearman r= 0.860	2% (73% to -46%)	Similar, but wide upper and lower limits of BA plot.
Adults – Non-asthmatics										
Chen 2007[41]	NR	Al.	Yes	27	43.1ppb	36.9ppb	NIOX MINO 7.2ppb higher (range 6.5 to 8.0), p<0.0001	r=0.94 to 0.99	nr	NIOX MINO reads higher (SS)
Adults – Healthy participants										
Menzies 2007[38]	Yes	He	Yes	50	GM (SE): 19.3 (17.6–21.1)ppb	GM (SE): 17.7 (16.1–19.4)ppb	NR	Pearson: r = 0.96, p<0.001	Log scale: -0.0 (0.1 to -0.2) [b]	Similar (SSNR)

Author name, year	Spons-ored?	Popul-ation	ATS 2005?	N[a]	Mean FeNO NIOX MINO	Mean FeNO chemiluminescence	Comparison data	Correlation coefficient	Bland-Altman	Interpretation
								Spearman: lack of bias at either end of values	Bigger differences at higher values[c, d]	
Hemmingson 2004[42]	Yes	He	NR	19	NR	NR	0.5ppb (SD 3.8ppb)	NR	NR	Similar (SSNR)
Adolescents and adults - Asthmatics or mixed including some asthmatics										
Pizzimenti 2008[43]	No	As HE	Yes	32	47.1 ppb (95% CI 35.2 to 59.1)	36.9 ppb (95% CI 25.0 to 49.0)	NIOX MINO higher, p<0.05	r = 0.998, p<0.001	NR	NIOX MINO reads higher (SS)
No age restrictions - Asthmatics or mixed including some asthmatics										
Khalili 2007[44]	yes	As	Yes	110	NR	NR	NIOX MINO higher, -0.5ppb (p=0.21)	Spearman: r=0.98, p<0.0001	0.5ppb (8.3 to -9.4) Bigger differences at higher values[c]	Similar, but with wide range in BA analysis (NSD)
Alving 2006[45]	Yes	As He	Yes	71	27.5 (SD23.2) n=62	26.5 (SD24.2) n=63		r=0.97	1.5ppb (10.2 to -13.2) Bigger differences at higher values[c]	Similar (SSNR) NIOX MINO slighly higher than Niox.

a, number analysed; b, estimated from graph; c, visual interpretation of Bland-Altman plot; d, values were logged – therefore the apparently random distribution on the graph represents a multiplicative association between values in the direction of greater differences at higher values.

ATS, America Thoracic Society; ppb, parts per billion; SSNR, statistical significance not reported; SS; statistically significant; NR, not reported; NSD, not statistically significantly different AS, asthmatic; COPD, chronic obstructive pulmonary disease; HE, healthy; Al, allergy; n, number; r, correlation coefficient; GM, geometric mean; SE, standard error; CI, confidence interval; BA, Bland-Altman

Table 13: Equivalence review: NIOX MINO versus Niox, children

Author name, year	Spons-ored?	Popul-ation	ATS 2005?	N	Mean FeNO NIOX MINO	Mean FeNO chemiluminescence	Comparison data	Correlation coefficient	Bland-Altman	Interpretation
Children – Asthmatics or mixed including some asthmatics										
Vahlkvist 2006[46]	Yes	AS with AL	Y	11	30ppb	26ppb	NIOX MINO Higher, p=0.004	r=0.977; unclear which method	Correlation independent of level	NIOX MINO higher (SS)
Kalliola 2011[48]	No	As & He	y	40	GM: 7.8	GM: 9.9	NIOX MINO lower, p = 0.002	Pearson's: r = 0.972, p<0.001	1.1ppb, (-4.4 to 6.7) No obvious difference at higher values[a]	NIOX MINO slightly lower (SS)
McGill 2006[47]	Yes	AS & others	Y	34	NR	NR	Niox higher, greater difference at higher values (p<0.001)	0.986 [95% CI 0.972, 0.993]	3.9 ppb, (-1.1 to 8.9) Mean difference greater with higher FeNO values[a]	NIOX MINO higher (SS)

a, visual interpretation of Bland-Altman plot
ATS, America Thoracic Society; ppb, parts per billion; SS; statistically significant; NR, not reported; AS, asthmatic; HE, healthy; Al, allergy; r, correlation coefficient; GM, geometric mean; CI, confidence interval

5.2.1.3 NIOX MINO versus other chemiluminescent devices in adults and children

Twelve studies compared NIOX MINO to chemiluminescent devices other than Niox and were included in the review (Table 14). Six studies (reported across eight publications) were in adults,[49-56] three in an unspecified age group,[57-59] (two of which had potentially largely overlapping cohorts and will be considered as one study)[57,58] and three in children.[60-62] All studies included at least some asthmatic patients, except de Laurentiis 2008.[56] Mean FeNO values varied across studies (range 7ppb in a healthy cohort of patients,[51,52] to 64.3ppb in an asthmatic cohort[53,59]) Devices studied in adults/unspecified age group were EndoNO, N-6008, NA623N, NOA280i, Ecomedics CLD88sp, NOA and LR2000, whilst in children only Ecomedics CLD88 and CLD77 were tested. No chemiluminescence device was tested in more than one study. In summary:

- Correlation coefficients (r) in adults/unspecified age group ranged from 0.876 to 0.96, indicating a good level of correlation between devices.

However, mean FeNO levels between devices and Bland-Altman statistics show a more variable picture.

- NIOX MINO appeared to give higher readings than the comparator device according to the mean FeNO values in four studies (counting Michils 2008[58] and Peche 2007[57] as one study),[49-52,57-59] and lower readings in a further two studies.[53-55] Devices appeared to be comparable in only two studies.[56,63] Absolute differences in mean FeNO values on the natural scale were not always reported, but where they were, they ranged from 9ppb[49,50] to 47ppb,[51,52] which could represent a clinically meaningful difference.
- Bland-Altman statistics were reported in only four studies,[49,50,55,56,63,64] and were not reported consistently Mean values were reported as relative values, log transformed data and absolute data. Interpretation would suggest mean differences were small, 0 to 5ppb, but that limits of agreement were much larger, with ranges of around 10ppb above and below the mean. The studies with the largest mean differences in absolute FeNO values did not report Bland-Altman statistics.

In children:

- Correlation coefficients (r) in children ranged from 0.69 to 0.98, indicating variable correlation. The study with the poorer correlation[60] also had higher mean FeNO levels, and it would be tempting to suggest that the poorer correlation is due to the greater variability at higher FeNO values. However, the study authors state that correlation improved at higher values. One study[62] noted that the direction of disagreement was different in children aged over and under 12 years.

- The back-transformed Bland-Altman statistics[61] and range of ratios reported[62] indicate a wide range of agreement, and suggest the devices are not interchangeable.

Table 14: Equivalence review: NIOX MINO versus other chemiluminescent devices

Author name, year	Spons-ored?	Popul-ation	Comparator device	ATS 2005	N	Mean FeNO NIOX MINO	Mean FeNO chemiluminescence	Comparison data	Correlation coefficient	Bland-Altman	interpretation
Adults - Asthma or mixed including some asthmatics											
Ozier 2010[49] Ozier 2011[50]	Yes	As	EndoNO	Yes	NIOX MINO: 78 Endo-NO: 89	Controlled asthmatics: approx. 23ppb[a] Uncontrolled asthmatics: approx. 38ppb[a]	Controlled asthmatics: approx. 19ppb[a] Uncontrolled asthmatics: approx. 30ppb[a]	9ppb difference, p<0.0001	Pearson, log transformed data r = 0.96 p=<0.001	Log scale: 0.12 (0.3 to -0.6) [a] No obvious difference at higher values[b]	NIOX MINO higher (SS)
Fortuna 2006,[51] Fortuna 2007[52]	NR	As	N-6008	Yes	11	79 (SD55) ppb	40 (SD 30) ppb	47 ppb (SD30)	r=0.9	NR	NIOX MINO higher (SS NR)
Fukuhara[65] Fukuhara 2011[53]	Yes	As-treated	NA623N	NR	13	Treated asthmatics: GM50.0 (26.5 to 73.4)	Treated asthmatics: GM64.5 (33.4 to 95.6)	p =0.009	NR	NR	NIOX MINO lower (SS)
Fukuhara[65] Fukuhara 2011[53]	Yes	As - untreated	NA623N	NR	14	Non-treated asthmatics:GM52.3 (28.8 to 75.7)	Non-treated asthmatics: GM64.3 (39.6 to 89.0)	p =0.0167	NR	NR	NIOX MINO lower (SS)
Fukuhara[65] Fukuhara 2011[53]	Yes	As He	NA623N	NR	32	NR	NR	p<0.001	NR	NR	Niox mIno lower (SS)
Kim 2012,[54] Yoon 2011[55]	No	As He	NOA280i	Yes	100	18.8 (SEM 0.9, 95% CI 17.0 to 20.6)	22.1 (SEM 1.2, 95% CI 19.8 to 24.5)	14.5% (2.5%).[c] Range -61.7 to 111.1%	spearman: r=0.876, p<0.001.	3.3 (13.6 to -7.0) Mean difference greater with higher FeNO values[b]	NIOX MINO lower (SS)
Boot 2008[63]	No	As He	Ecomedics CLD88sp	Yes	37	Healthy: 20.3 (8.0 to 39.0) Healthy smokers:12.2 (5 to 23)	Healthy: 18 (7.4 to 35.5) Healthy smokers: 11.1 (4.7 to 20.5) Asthmatics: 60.8	Non-significant difference	R = 0.975, p<0.0001	-10% (-36%, to 28%) Bigger differences at higher values[b, d]	Similar (NSD), but with wide range in BA analysis

Author name, year	Spons-ored?	Popul-ation	Comparator device	ATS 2005	N	Mean FeNO NIOX MINO	Mean FeNO chemiluminescence	Comparison data	Correlation coefficient	Bland-Altman	interpretation
						Asthmatics: 63.8 (13 to 172)	(10.9 to 184.6)				
Adults – Non-asthmatics											
de Laurentiis 2008[56]	Yes	COPD He	NOA, Sensor-medics	Yes	20	14.8 (SD5.7)	14.2 (SD5.9)	NR	r = 0.96, p<0.0001	- 0.4 ppb, (-2.7 to 1.9) Bigger differences at higher values[b]	Similar (SSNR)
Adults – Healthy participants											
Fortuna 2006,[51] Fortuna 2007[52]	NR	He	N-6008	Yes	28	20 (SD 8) ppb	7 (SD 5) ppb	13 (SD14)	r=0.92, p=0.001[52]	NR	NIOX MINO higher (SSNR)
Fukuhara[65] Fukuhara 2011[53]	Yes	He	NA623N	NR	5	GM23.4 (7.18 to 39.6)	GM29.3 (7.77 to 50.8)	p =0.073	NR	NR	NIOX MINO lower (trend, NSD)
Population unclear – mix including some asthmatics											
Peche 2007[57 e]	NR	As He lung trans.	LR-2000	Yes	118	NR	NR	NIOX MINO averge 35% higher	NR highly correlated (p<0.001)	NR	NIOX MINO higher (SSNR)
Michils 2008[58 e]	NR	As He	LR-2000	Yes	102	NR	NR	mean log feno difference 0.144 NIOX MINO averge 39% higher	r = 0.957, p<0.001 Bigger differences at higher values[f, d]	NR	NIOX MINO higher (SSNR)
Logan Research Ltd 2009[59]	Yes	As He	LR2000	NR	16	NR	NR	+13.1ppb		NR	NIOX MINO higher (SSNR)
Children – mix including some asthmatics											
Park 2011[60]	NR	As He	EcoMedics CLD88	NR	188	30.8 (SD23.4)	42.8 (SD 30.1)	NR	0.690 (p<0.001)	NR	NIOX MINO lower (SSNR)
Schiller	NR	As	EcoMedics	Yes	66	3 measures: 23.7	3 measures: 20.1	NR	0.98,	Ratio: 0.79	NIOX MINO

Author name, year	Spons-ored?	Popul-ation	Comparator device	ATS 2005	N	Mean FeNO NIOX MINO	Mean FeNO chemiluminescence	Comparison data	Correlation coefficient	Bland-Altman	interpretation
2009[61]		He	CLD77	.		ppb 1st measure: 23.6 ppb	ppb 1st measure: 20.3 ppb		p<0.001	(0.44 to 1.42)	higher (SSNR) Larger difference at higher values. (SSNR)
Chladkova 2008[62]	No	As Al Rh	EcoMedics CLD88sp	Yes	82	>12 years: GM17.4 (7.05 to 43.4) [g] <12 years: 11.9 (6.87 to 21.9) [g]	>12 years: GM19.6 (7.43 to 51.6) [g] <12 years: 9.59 (4.74 to 19.4) [g]	>12 years: Ecomedics 11% higher <12 years: Ecomedics 11% lower	NR	Median ratio: >12 years: 1.11 (0.59 to 2.08) <12 years: 0.89 (0.52 to 1.52, bigger difference at higher values[b]	>12 years: NIOX MINO lower <12 years: NIOX MINO higher

[a] Data estimated fromgraph
[b] Visual interpretation of Bland-Altman plot
[c] Mean relative difference (SEM)
[d] values were logged – therefore the apparently random distribution on the graph represents a multiplicative association between values in the direction of greater differences at higher values.
[e] likely that Peche 2007[57] and Michils 2008[58] include some of the same patients.
[f] assumed from mean values
[g] Geometric mean (exp(mean±SD))

ATS, America Thoracic Society; ppb, parts per billion; SSNR, statistical significance not reported; SS; statistically significant; NR, not reported; NSD, not statistically significantly different AS, asthmatic; HE, healthy; Al, allergy; Rh, patients with rhinitis; n, number; r, correlation coefficient; GM, geometric mean; SE, standard error; CI, confidence interval

5.2.1.4 NIOX VERO [redacted]

One recently completed and as yet unpublished study was submitted by the sponsors, Aerocrine.[24] [redacted] The results are summarised in Table 15.

Table 15: [redacted]

Author name, year	Spons-ored?	Popul-ation	Comparator device	N	Correlation coefficient	Bland-Altman	Interpretation
[redacted]							

[redacted]

5.2.1.5 NObreath versus chemiluminescent devices in adults and children

Four studies compared NObreath to chemiluminescent devices (Table 16).[53,59,65,67,68] All studies were in adults, or an unspecified age group likely to be adults, and all included some asthmatic patients, and reported good correlation coefficients. Only one study compared NObreath to Niox, and reported but a small statistically significant difference (geometric mean 22.6 ppb (geometric standard error of the mean 1.075) in the intervention arm and 24.6 ppb (1.073) respectively in the control arm, p = 0.0002), and a good level of agreement with Bland-Altman analysis. However, the cut-off points with the best combination of sensitivity and specificity derived in this study for each device differed by 10 ppb (25ppb for NObreath, 15ppb for Niox, see Section 5.2.1.7), indicating that even small differences in agreement may have potentially large effects on derived sensitivity and specificity.

NObreath was compared to three other chemiluminescent devices: NA623N, Logan LR2500 and Logan LR2000. Agreement by Bland-Altman analysis was only reported in one study[68] and showed a

mean difference of -3.95 ppb in comparison to the LR2500 in a healthy cohort with low FeNO values, and limits of agreement were wide (-10.98 to 4.08). Similarly, another study[59] using a Logan device (LR2000) reported a absolute mean difference in FeNO measurements of -3.81 ppb. Comparison with the NA623N[65,65,69] showed small differences between mean FeNO values for the cohort, with NObreath giving lower values in some cohorts.

Table 16: Equivalence review: NObreath versus chemiluminescent devices

Author name, year	Spons-ored?	Popu l-ation	Comparat or device	Method	N	Mean FeNO NioxMino, ppb (95% CI)	Mean FeNO chemiluminescence , ppb (95% CI)	Comparison data	Correlation coefficient	Bland-Altman	Interpretation
Adults – Asthmatics or mix including some asthmatics											
Pisi 2010[67]	No	As	Niox	Yes	154	GM: 22.6 (GSEM 1.075)	GM: 24.6 (GSEM 1.073)	p=0.0002	Pearson's r = 0.95, p<0.001 Spearman's r =−.088, p =0.275).	Log scale: +0.0 (0.5 to -0.5)[a] No obvious difference at higher values[b]	NObreath lower (SS), but small difference
Fukuhara[65] Fukuhara 2011[53]	Yes	As – treated	NA623N	NR	13	55.6 (28.1 to 83.2)	64.5 (33.4 to 95.6)	p=0.015	NR	NR	NObreath lower (SS)
Fukuhara[65] Fukuhara 2011[53]	Yes	As – untre-ated	NA623N	NR	14	66.8 (38.3 to 95.4)	64.3 (39.6 to 89.0)	p=0.413	NR	NR	Similar (NSD)
Fukuhara[65] Fukuhara 2011[53]	Yes	As He	NA623N	NR	32	NR	NR	p=0.138	r = 0.969 (p<0.001)	NR	Similar (NSD)
Adults – Healthy participants											
Fukuhara[65] Fukuhara 2011[53]	Yes	He	NA623N	NR	5	25.0 (7.87 to 42.1)	29.3 (7.77 to 50.8)	p=0.233	NR	NR	NObreath lower (trend, NSD)
Antus 2010[68 c]	Yes	He	Logan LR2500	Yes	18	15.7 (11.7 to 21.9) 14.8 (10.4 to 21.3)	13.0 (10.1 to 16.7) 13.5 (10.4 to 17.4)	p value 0.299 0.351 0.179	r = 0.897, p<0.001 r = 0.913, p<0.001 r = 0.938,	-3.95ppb (-10.98 to 4.08) Plot suggestive	Similar (NSD)

						16.4 (12.3 to 21.9)	12.9 (9.9 to 16.6)		p<0.001	of systematic bias[b]	
Population unclear – mix including some asthmatics											
Logan Research Ltd 2009[59]	Yes	As He	LR2000	NR	16			NObreath to LR2000: -3.81ppb			NObreath lower (SSNR)

[a]Estimated from graph
[b]Visual interpretation of Bland-Altman plot
[c] FeNO measured at three time points

ATS, America Thoracic Society; ppb, parts per billion; SSNR, statistical significance not reported; SS; statistically significant; NR, not reported; NSD, not statistically significantly different AS, asthmatic; HE, healthy; n, number; r, correlation coefficient; GM, geometric mean; GSEM, geometric standard error of the mean.

5.2.1.6 NObreath versus NIOX MINO in adults and children

Table 17 details the two studies which compared NObreath to NIOX MINO in adults.[53,65,68] Both studies found that in most analyses NIOX MINO provided lower mean FeNO values than NObreath. This contradicts the available evidence for comparisons of NIOX MINO to Niox and NObreath to Niox, where NIOX MINO>Niox>NObreath. This would predict that NIOX MINO should provide higher readings than NObreath. However, there is only one study comparing NObreath to NIOX MINO, and the difference observed was small. The two direct comparisons of NObreath and NIOX MINO are in small numbers of patients, and only one includes asthmatic patients,[53,65] but does not provide a Bland-Altman analysis to assess agreement. As such, it is unclear whether the two devices are interchangeable, and if not in which direction the difference may be.

5.2.1.7 AUCs, cut-off points and correction equations

The mean FeNO, correlation and Bland-Altman data for each of these trials have already been considered in the previous narrative synthesis (see Sections 5.2.1.1 to 5.2.1.6); this section considers the impact that differences between devices can have on cut-off points, and reports the attempts researchers have made to provide correction equations for measurements between devices.

Six studies reported other comparative data between devices (Table 18). One study[38] demonstrated that the area under the curve and cut-off points derived to diagnose asthma using Niox or NIOX MINO were very similar (Note: this study has a case-control design so data was not includable in the diagnostic review), supporting a conclusion that Niox and NIOX MINO are roughly interchangeable. However, another study[43] reports a correction factor that should be used to convert NIOX MINO values to Niox values. Three[49,50,53,65,67] of the remaining four studies[49-53,65,67] demonstrate how cut-off points derived using measurements from different devices can be very different, with 7ppb, 9ppb and 10ppb differences. One of these studies, Pisi 2010[67] compares NObreath to Niox directly, and finds a 10ppb difference between cut-off points that provide the highest AUC, (15ppb and 25ppb respectively) and very different sensitivity and specificity values at these cut-off points. Another, Fukuhara[65] and Fukuhara 2011[53] compares NObreath to NIOX MINO and finds a 7ppb difference in derived cut-off points. In this case, the cut-off for NObreath is higher at 36ppb. Two studies[51-53,65] also report correction equations between various devices, indicating measurements from these devices are not directly interchangeable.

Table 17: Equivalence review: NIOX MINO and NObreath compared to each other

Author name, year	Sponsored?	Population	Method	N	Mean FeNO NIOX MINO ppb (95% CI)	Mean FeNO NObreath ppb (95% CI)	Comparison data	Correlation coefficient	Bland-Altman	Interpretation
Adults – Asthmatics or mix including some asthmatics										
Fukuhara[65] Fukuhara 2011[53]	Yes	As-treated	NR	13	50.0 (26.5 to 73.4)	55.6 (28.1 to 83.2)	p=0.135	NR	NR	NIOX MINO lower (trend, NSD)
Fukuhara[65] Fukuhara 2011[53]	Yes	As - untreated	NR	14	52.3 (28.8 to 75.7)	66.8 (38.3 to 95.4)	p=<0.001	NR	NR	NIOX MINO lower (SS)
Fukuhara[65] Fukuhara 2011[53]	Yes	As He	NR	32	NR	NR	p<0.001	r = 0.973 (p<0.001)	NR	NIOX MINO lower (SS)
Adults – Healthy participants										
Fukuhara[65] Fukuhara 2011[53]	Yes	He	NR	5	23.4 (7.18 to 39.6)	25.0 (7.87 to 42.1)	p=0.669	NR	NR	Similar (NSD)
Antus 2010[68a]	Yes	He	Yes	18	3 time points GM (95% CI): 12.6 (9.3 to 17.1) 10.3 (7.8 to 13.4) 10.9 (8.4 to 14.2)	3 time points GM (95% CI): 14.9 (11.9 to 18.8) 16.1 (13.3 to 19.6) 16.3 (12.8 to 20.6)	p value: 0.409 0.010 0.043	r = 0.661, p=0.004 r= 0.750, p<0.001 r=0.654, p= 0.003	4.36ppb (-7.38 to 16.1)	NIOX MINO lower (SS in two analyses) Correlation relatively poor
Children – asthmatics										
Kapande 2012[70,71]	No	As	Yes [b]	109	NR	NR	7.8ppb (-11.5 to 27.52 ppb). NIOX MINO higher	Lin's CC (rho) = 0.65 with reduced major axis slope of 1.32 and intercept of 5.03	NR	NIOX MINO higher (SSNR)

Lin's CC, Lin's concordance correlation; ATS, America Thoracic Society; ppb, parts per billion; SSNR, statistical significance not reported; SS; statistically significant; NR, not reported; NSD, not statistically significantly different AS, asthmatic; HE, healthy; GM, geometric mean

[a]FeNO measured at three time points

[b] At least partially

Table 18: Equivalence review: Area under the Curve, sensitivity, specificity and cut-off points using different devices and correction equations derived to convert FeNO values between devices

Author, year	Popu-lation	Comparator device	Auc	Sensitivity, specificity and cut-off values	Correction equation
NIOX MINO versus Niox: Adults/adolescents - mix including some asthmatics					
Menzies 2007[38]	As He	Niox	Niox: 0.654 (95% CI 0.565 to 0.744, p=0.002) NIOX MINO: 0.619 (95% CI 0.527 to 0.711, p=0.018) Pairwise comparison difference in AUC of 0.036 (95% CI -0.002 to 0.073, p=0.061)	Sens 83.2%, spec 27%: Niox: 13ppb NIOX MINO: 12.5ppb	NR
Pizzimenti 2008[43]	As He	Niox	NR	NR	FeNO (niox) = -1.656 (SE=0.61) + 0.808 (SE=0.009) x FeNO (NIOX MINO) Correction factor = 0.81 approximately to convert NIOX MINO into Niox.
NIOX MINO versus other chemiluminescent devices: Adults - mix including some asthmatics					
Ozier 2010[49] Ozier 2011[50]	As	EndoNO	NR	To identify patients who will lose control: NIOX MINO: 40ppb (sens 85.7%, spec 87.8%) EndoNO: 31ppb (sens 80.0%, spec 91.1%)	NR
Fukuhara[65] Fukuhara 2011[53]	As He	NObreath NA623N	NR	NObreath: >36ppb NIOX MINO: >29ppb	NA623N = NIOX MINO x 1.278 + 3.065 NA623N = NObreath x 0.953 + 5.779
Fortuna 2006,[51] Fortuna 2007[52]	As He	N-6008 (SIR, Spain)	NR	NR	Correction factor = 3 For healthy: FeNO NIOX MINO = 10+(1.5 x FeNO N6008)
NObreath versus Niox: Adults, asthmatics					
Pisi 2010[67]	As	Niox	To identifiy patients who have ACT≥20	To identifiy patients who	NR

			(uncontrolled asthma): NObreath: 0.607 (95% CI: 0.525–0.684) Niox: 0.644 (95% CI: 0.562–0.719) Pairwise comparisons of difference in AUC of 0.0369 (95% CI: 0.004–0.0697; $p = 0.028$)	have ACT≥20 (uncontrolled asthma): NObreath: 15ppb (sens 84%, spec 42%) NIOX: 25ppb (sens 53%, spec 69%)	
ATS, America Thoracic Society; ppb, parts per billion; NR, not reported; AS, asthmatic; HE, healthy; AUC, area under the curve; CI, confidence interval; sens, sensitivity; spec, specificity; ACT, asthma control test.					

5.2.1.8 Test failure rates

The review of test failure rates intended to draw evidence from studies included in the review of equivalence of devices, the diagnostic accuracy review and the efficacy of FeNO guided management for asthma review. However, all nine of the studies that provided some relevant information with respect to test failure rates,[38,44,45,47,48,62,63,67,70,72] came from the review of equivalence of devices. Eight studies[38,44,47,48,62,63,70,72,73] examined NIOX MINO, and two[67,70,72] used NObreath. The definition of a test failure was reasonably consistent across the body of literature. Boot 2008,[63] Kalliola 2011,[48] Kapande 2012,[70,72] Khalili 2007,[44] Pisi 2010,[67] and Menzies 2007,[38] all defined test failure rates in terms of the number of patients who could not perform acceptable measurements. However, McGill 2006[47] specified a failure as inability to provide a successful reading from six attempts; Chladovka 2008[62] defined a failure as three unsuccessful attempts, and Alving 2006[45] defined test failure as 3 invalid readings out of 6, or one failed single first attempt, depending on the device used.

All studies included patients with confirmed asthma or symptoms suggestive of asthma; however, the criteria for establishing this diagnosis varied across the literature. For instance, Pisi 2010[67] included those who met the criteria for a diagnosis of asthma according to GINA guidelines;[74] while Menzies 2007[38] stated that they included those with mild to moderate persistent asthma, and McGill 2006[47] included children attending a respiratory clinic. Kalliola 2011[48] included children who had been referred to a specialist clinic due to asthma-like symptoms. In terms of the age range of study samples, Kalliola 2011,[48] Kapande 2012[70], Chladovka 2008,[62] and McGill 2006[47] were all conducted with children. The study of adults was Boot 2008.[63] In addition, although Menzies 2007[38] did not report any cut-off ages for inclusion, the mean age of the study sample suggests it was conducted with adults only. Alving 2006[45] included all ages, and provided separate data for children and adults, while Khalili 2007[44] included all ages and reported test failures for the whole study cohort. Pisi 2010[67] included adolescents and adults (the cut-off age for inclusion was $\geq$14 years of age).

NIOX MINO: Eight studies[38,44,45,47,48,62,63,70,72] reported test failure rates with NIOX MINO. Kalliola 2011,[48] Kapande 2012,[70] Chladovka[62] and McGill 2006[47] were conducted with children, while Boot 2008,[63] and Menzies 2007[38] were conducted with adults only. Alving 2006[45] and Khalili 2007[44] included all age ranges, however, only Alving provided separate data for adults and children. Although the data sets were limited in both age cohorts, the test failure rates for NIOX MINO were consistently higher in the studies of children. In the adult only studies of Boot 2008[63] and Menzies 2007,[38] no test failures were observed in cohorts of 50 and 151 participants, respectively, and similarly, the adult cohort in Alving 2006[45] showed a test failure rate of 0% (0/34 participants). The overall test failure rate in adults is therefore likely to be close to 0%. However, data were unavailable as to how many attempts were required, on average, to obtain a successful reading.

In the children's cohorts, however, there were test failures in each study. The rate ranged from 2 out of 36 (5.5%) in Chladovka 2008[62] to 27% (15/55 participants) in Kalliola 2011.[48] Alving 2006[45] reported a failure rate of 16% (6/37 participants). McGill 2006[47] classified failures as those who were unable to provide a successful reading from six attempts. They reported 11 patients who fell into this category (20%). In terms of overall FeNO measurement attempts, this would indicate at least 66 failed tests out of 330, i.e. also a 20% test failure rate. It was also notable that, in the study with the highest incidence of failure (Kalliola 2011),[48] the age of the children who failed was significantly lower than those who successfully provided a measurement (p=0.004). In Khalili 2007,[44] a failure rate of ~0.9% was observed (one failure out of 115 participants); however, as data were not presented separately for adults and children, the age of this participant was not clear.

NObreath: Only two studies, Pisi 2010,[67] and Kapande 2012[70,72] reported test failure rates with the NObreath device. In the Pisi 2010 study of adolescents and adults,[67] there were five failures in a cohort of 154 patients (3.2%), and a single patient who failed on both NObreath and Niox. Two patients were said to have provided 'unapproved values'; however, it was unclear whether this was on Niox chemiluminescence, NObreath, or both. The Kapande 2012 study of children only[70,72] reported no test failures in a cohort of 109.

[redacted]

Based on the available data, FeNO test failure rates appeared to generally be low. Most studies reported test failure rates in terms of the number of patients who were unable to provide a satisfactory reading; however, the data also appeared to indicate that multiple readings would be needed for some patients. As this data was not quantified, and usually not reported at all, it is likely that the review underestimates the number of test failures. Moreover, variations in failure rates may be a result of individual differences in operator skills and techniques. Notably, the highest rate of test failure for both NIOX MINO and chemiluminescence was observed in the same study.[48]

There may also be important variations in test failure rates depending on age, particularly when using NIOX MINO. Three of the four NIOX MINO studies of adults reported failure rates of 0%, while one study reported a withdrawal rate of 13.3%; however, it was unclear whether withdrawals could be treated as synonymous with test failures. By contrast, the failure rate in children's studies ranged from 5.5%% to 27%. Indeed, in the study that reported the latter figure,[48] the children who failed the test were significantly younger than those who provided a successful reading. Although the Khalili 2007 study[44] included all age groups and reported a much lower failure rate (~0.9%), the mean age of the

study cohort (41.9 range: 6 to 86) may indicate that few children took part. Hence, although the data are too limited to make any definitive conclusions, it seems likely that higher test failure rates may be encountered when using NIOX MINO with children. Finally, with respect to NObreath, the data were particularly sparse, although a low failure rate was apparent. Pisi 2010,[67] reported six failures in a cohort of 154 adults and adolescents (3.9%), while Kapande 2012[70,72] saw no failures in a cohort of 109 children.

Table 19: Test failure rates

Author, year	Patient sample and n participants	Relevant device(s) used	Definition of test failure	Test failure rates
Alving 2006[45]	Asthmatic and healthy, all ages (n=75)	NIOX MINO	Successful test defined as 3 valid readings out of 6, or one single first attempt	NIOX MINO All: 92% (65/71) patients successful Children: 84% (31/37) successful adults: 100% (34/34) successful
Boot 2008[63]	Asthmatic and healthy adults (n=50)	NIOX MINO	No. patients could not perform acceptable measurement	NIOX MINO 0/50
Chladovka 2008[62]	Children with asthma (n=82)	NIOX MINO	No. patients could not perform acceptable measurement within 3 attempts	NIOX MINO 2/36
Kalliola 2011[48]	Children referred due to asthma symptoms Healthy age-matched children (total n=55)	NIOX MINO	No. patients could not perform acceptable measurement	NIOX MINO: 15/55 (younger than successful measurement group, P=0.004)
Kapande 2012[70,72]	Children (age 4 to 14 years) with asthma (n=109)	NIOX MINO	No. patients could not perform acceptable measurement	NIOX MINO: 7/109 Nobreath: 0/109

		NObreath		
Khalili 2007[44]	Patients (all ages) with asthma (n=115)	NIOX MINO	No. patients could not perform acceptable measurement	NIOX MINO: 1/115 A few subjects needed to perform a test 4 to 7 times
McGill 2006[47]	Children attending respiratory clinic over 5 years of age (n=55)	NIOX MINO	Number of children unable to provide a single measurement out of 6 attempts	NIOX MINO: 11/55 (therefore at least 66 failed tests out of 330 = 20% test failure rate)
Menzies, 2007[38]	Patients known to have mild-moderate asthma (n=101); healthy volunteers (n=50)	NIOX MINO	No. patients could not perform acceptable measurement	NIOX MINO: 0
Pisi 2010[67]	Patients age $\geq$ 14 years Diagnosed with asthma according to GINA guidelines Only patients able to perform at least 2 acceptable measurements with both devices.	NObreath	No. patients could not perform acceptable measurement	Nobreath: 5 /154 Both (Niox and NObreath): 1 Unapproved values: 2

5.2.1.9 Discussion

It is worth noting that there was data available within some of these studies as to which device was used first. However, owing to time constraints, this data was not formally analysed. In some cases the order was random, in others the order was fixed, and in yet others this information was not provided. We cannot therefore rule out the possibility that the order of device use may have confounded results.

5.2.1.9.1 NIOX MINO

The comparability of NIOX MINO to chemiluminescent devices appears to be influenced by several factors. There may be some variability between NIOX MINO devices themselves,[39] though the extent of this is unclear and may be small.[44] There seems to be a generally consistent observation of poorer equivalence between FeNO devices at higher FeNO levels. There also appears to be a lack of comparability between other chemiluminescent devices themselves, as concluded by one study,[40] which leads to heterogeneity in estimates of the comparability of NIOX MINO to chemiluminescent devices. Comparability studies gave different estimates of equivalence between NIOX MINO and other devices, and it is therefore unclear if equivalence can be assumed.

5.2.1.8.2 NIOX VERO

Only one study provides data on this device. [redacted]

5.2.1.9.3 NObreath

There is not enough data and too much apparently conflicting data on the comparability of NObreath to other devices to draw any specific conclusions as to its comparability with other devices in asthma populations. However, based on the available evidence, it would seem likely that any differences in absolute values between NObreath and other devices are relatively small, though derived cut-offs and maximum sensitivity and specificity may be quite different.

5.2.1.9.4 Test failure rates

Due to the small number of studies using NObreath and NIOX VERO, it is not possible to state definitively whether any FeNO measurement device has advantages over any other in terms of test failure rates. In adults using NIOX MINO, the test failure rate was 0% in all three studies,[38,45,63] whilst none of the studies conducted in children using NIOX MINO reported 0% test failure, with the lowest being 5.5%[62] and the highest 27%.[48] As such, there may be some problems in using NIOX MINO with children, although further research would be needed to confirm this pattern. Conversely, in NObreath, the study in adults reported a 3.3%[67] failure rate, whilst in children it was 0%.[70,72] [redacted]

[redacted]

In summary, the overall test failure rate for FeNO measurement in adults was generally low across all devices, and most patients appear to be able to provide FeNO readings, provided they are permitted sufficient measurement attempts. There may be a higher test failure rate in children using NIOX MINO.

5.2.1.10 Conclusions

Overall, it cannot be concluded that any two devices are equivalent in all situations. Whilst there may be situations where they are similar, it appears to depend on the characteristics of the studies and cannot be generalised to all situations. Further research is required to identify what is driving the variability between studies and devices. However, as there is mostly a high degree of correlation between measurements across all devices, estimates of sensitivity and specificity are likely to be a reasonable indication of potential diagnostic accuracy of using FeNO to guide diagnosis and management, but the derived cut-off points are not likely to be interchangeable between devices. As such, for the purpose of this assessment, sensitivities and specificities will be assumed to be interchangeable, but it cannot be assumed that the cut-off points that should be used to achieve them will be the same in each device, and there is still some doubt as to whether the same diagnostic accuracy would be achievable in all devices. The committee will need to consider this in their recommendations.

5.2.2 Diagnostic review

In the absence of an end to end study, the next best study design is a cohort study. The ideal cohort study would have recruited patients presenting to their GP with symptoms of asthma, and would have assessed the standard UK diagnostic pathway[8] as well as this pathway with the addition of FeNO against a reference standard of long term follow-up. No studies of this design were found. Instead, studies which compared FeNO with or without another test against a reference standard of any test or combination of tests in the UK guidelines (Figures 3 and 4) were included. UK guideline tests include:

- Spirometry and lung function tests (mostly FEV_1%, FEV_1/FVC, PEF)
- Airway reversibility: airway obstruction which shows reversibility when a bronchodilator is taken
- ICS responsiveness: response to a trial of treatment with inhaled corticosteroids
- Airway hyper-responsiveness: to methacholine, histamine, exercise or mannitol
- Tests for airway inflammation (FeNO or sputum eosinophil counts), though these are currently restricted to a few specialist centres. Studies which use sputum eosinophils within

the reference standard are not considered to be similar to UK practice as this test is not widely available.

Twenty four cohort studies (across 25 publications) which reported sensitivity and specificity of FeNO with or without another test versus some or all of the comparators within the UK diagnostic pathway were identified and included in the review. Four of these studies included data for FeNO in conjunction with another test.

Studies were not similar enough to warrant a meta-analysis, with substantial heterogeneity in populations, cut-off values, devices used, and reference standards. We decided to instead focus on key studies which most closely resembled UK practice, and resifted the included studies to separate out these studies. We did not, however, want to exclude completely the other studies in case they might prove useful to the committee in their decision making, especially as some were studies that the SCM had indicated might be of use when consulted during the clarification of the inclusion and exclusion criteria (see Appendix 3).

The next section is subdivided into a number of other sections and to aid reading a summary is given here:

- 5.2.2.1 – Studies including adults, adults plus adolescents, all age groups and unspecified age groups
    - **5.2.2.1.1 - Studies meeting the inclusion criteria** – this section tabulates all included studies and assess their relevance to the decision problem
    - **5.2.2.1.2 – Studies relevant to the decision problem using only FeNO as the index test** – this section provides an appraisal of study quality, a narrative synthesis and greater detail relating to these studies, along with estimates of sensitivity and specificity.
    - **5.2.2.1.3 - Studies using FeNO in conjunction with another test as the index test**– this section provides a narrative synthesis and greater detail relating to these studies, along with estimates of sensitivity and specificity.
- 5.2.2.2 - Studies including children or children and adolescents
    - **5.2.2.2.1 - Studies using only FeNO as the index test** - owing to the smaller number of studies relating to children, all studies are included without selection based on their relevance to the decision problem. This section provides a narrative synthesis and greater detail relating to these studies, along with estimates of sensitivity and specificity.
    - **5.2.2.2.2 - Studies using FeNO in conjunction with another test as the index test** - this section provides greater detail relating to one study, along with estimates of sensitivity and specificity.

- 5.2.2.3 – Studies providing data on subgroups of interest to the review
    - **5.2.2.3.1 – Adult smokers**
    - **5.2.2.3.2 – Children exposed to tobacco smoke**
    - **5.2.2.3.2 – Pregnant women**
    - **5.2.2.3.3 – The elderly**

More detailed descriptions of the study characteristics for studies which were judged to be less relevant, along with sensitivity and specificity data are included in Appendix 8 and Appendix 9, for reference.

5.2.2.1 Adults, adults plus adolescents, all age groups and unspecified age groups

Of the 25 studies included in the review, 20 were conducted in adults (11 studies, 12 citations),[53,75-85] adults plus adolescents (three studies, four citations),[86-89] all age groups (three studies, three citations)[90-92] or an unspecified age range (three studies, three citations).[93-95] Table 20 summarises the characteristics of the studies and provides a brief description explaining their relevance to the decision problem (note: Schneider 2009[77,78] appears twice in this table as it reported two differently defined populations). This table should be read alongside Figure 8, which ascribes a letter to positions which already exist in the current diagnostic pathway, and a number to the positions where FeNO could be added to the pathway, as agreed during the scoping workshop.

5.2.2.1.1 Studies meeting the inclusion criteria

Of the 20 studies conducted in adults that met the inclusion criteria for the review

- Nine studies recruited patients with symptoms of asthma who were broadly equivalent to patients entering the UK pathway at Position A (Figure 8). Of these, six were considered to be of most relevance to the decision problem, though the studies had not necessarily been conducted in the UK.
- Nine studies recruited patients who could be considered "difficult to diagnose" and are located at other points along the diagnostic pathway. These patients had already undergone some of the tests in the UK pathway, and had tested negative for asthma thus far. One further study (Schneider 2009)[77,78] reported a subset of difficult to diagnose patients from a larger cohort of patients at Position A (Figure 8). Of these, six were considered to be of most relevance to the decision problem, though the studies had not necessarily been conducted in the UK.
- One study recruited patients with suspected exercise induced bronchoconstriction. This study was considered to be relevant to the decision problem, though it had not been conducted in the UK.

- One study recruited army recruits, amongst whom a high proportion are thought to have lied about their asthmatic status. This study was not considered to be relevant to the decision problem, though it had not been conducted in the UK.

Reasons for considering a study not relevant to the decision problem are given in Table 20. A total of thirteen studies were considered relevant and are considered in greater detail in Section 5.2.2.1.2. Full study details and results for the eight studies considered not relevant are provided in Appendices 8 and 9.

Table 20: Diagnostic review: Key study characteristics of diagnostic cohort studies

Study author, year	Population	Age group	Device	Cut-off values	Reference standard	Position of FeNO in the pathway	Relevance to decision problem
Position A							
Schneider 2013[75]	Position A	Adults	NIOX MINO	9, 12,16, 20, 25, 35, 41, 42, 43, 44, 45, 46, 71 ppb	FEV_1, FEV_1/FVC, airway reversibility, hyper-responsiveness (MCT)	FeNO replaces whole pathway (excluding trial of treatment)	Relevant
Schneider 2009[77,78]	Position A	Adults	NIOX MINO	12, 16, 20, 35, 46, 76 ppb	FEV_1, FEV_1/FVC, airway reversibility, airway hyper-responsiveness MCT in sequence similar to UK guidelines	FeNO replaces whole pathway (excluding trial of treatment)	Relevant
Smith 2005[87]	Position A	Adults and adolescents	Niox	≥15, >47, <15 ppb	ATS 1987 symptoms plus one of: Airway reversibility, positive response to ICS, airway hyper-responsiveness (MCT).	FeNO replaces whole pathway	Relevant
De La Barra 2011[88]				25, 40, 50, 70, 90, 110, 130, 150 ppb	Airway reversibility	FeNO replaces airway reversibility	Relevant
Smith 2004[90]	Position A	All ages	NR	20 ppb	ATS 1987 symptoms plus one of: Airway reversibility, positive response to ICS, airway hyper-responsiveness (MCT).	FeNO replaces whole pathway	Relevant though uses unknown device of unknown equivalence
Fortuna, 2007[76]	Position A	Adults	SIR N-6008	20 ppb	FEV_1, FEV_1/FVC, airway responsiveness, WBP, airway hyper-responsiveness (MCT)	No equivalent position in UK pathway	Not relevant – uses WBP and device of unknown equivalence.
Fukuhara	Position A	Adults	NA623 (Chest	40 ppb	At least 2 of the 3	No equivalent	Not relevant– reference

Study author, year	Population	Age group	Device	Cut-off values	Reference standard	Position of FeNO in the pathway	Relevance to decision problem
2011[53,65]			MI, Tokyo, Japan)		criteria of induced sputum eosinophilia, airway hyper-responsiveness, and reversibility. Exclusion of other lung diseases.	position in UK pathway (sputum eosinophilia)	standard not similar enough to common UK practice (sputum eosinophilia)
Subset of Position A							
Cordeiro 2011[91]	Position A with high prevalence of atopy	All ages	Niox-Flex	27 ppb	Airway reversibility, airway hyper-responsiveness (histamine)	FeNO replaces whole pathway	Relevant – though population not whole spectrum
			Niox-Flex Airway reversibility	FeNO 27ppb AND/OR >12% and 200ml improvement in FEV_1 with bronchodilator	Airway reversibility, airway hyper-responsiveness (histamine)	Combination replacing FEV_1/FVC and airway hyper-responsiveness	Relevant – though population not whole spectrum
Heffler 2006[86]	Position A WITH rhinitis	Adults and adolescents	Niox	10, 15, 20, 25, 30, 34, 36, 40, 454, 50, 55, 60, 65, 75, 80, 85, 100 ppb	Airway hyper-responsiveness (MCT) or airway reversibility	FeNO replaces pathway (not including ICS responsiveness, but including airway reversibility)	Relevant – rhinitis population generalizable to whole asthma population
Pizzimenti 2009[94]	Position A Chronic cough	Unclear age group	NIOX MINO	55 ppb	Airway hyper-responsiveness (MCT)	No equivalent position in UK practice	Not relevant – not equivalent to UK practice
Difficult to diagnose							
Schleich, 2012[83]	Position E	Adults	Niox	34 ppb	Airway hyper-responsiveness (MCT)	FeNO at Position 1 and 2 (in place of MTC)	Relevant

Study author, year	Population	Age group	Device	Cut-off values	Reference standard	Position of FeNO in the pathway	Relevance to decision problem
			Niox FEV_1 101%	FeNO >34ppb AND FEV_1 ≤101%	Airway hyper-responsiveness (MCT)	Combination replaces FEV_1/FVC and airway hyper-responsiveness (MCT)	Relevant
Pedrosa 2010[89]	Position D via Bii and negative airway reversibility test	Adults and adolescents	NIOX MINO	40 ppb	Airway hyper-responsiveness (MCT)	FeNO at Position 1	Relevant
Bobolea 2012[92]	Position F	All ages	NIOX MINO	30 ppb	Adenosine challenge test	FeNO at Position 2	Relevant
Schneider 2009[77,78]	Position Bii	Adults	NIOX MINO	46, 16 ppb	Airway hyper-responsiveness (MCT)	NA	Not relevant – in UK these patients more likely to receive trial of treatment than MCT
Mathew 2011[95]	Patients at E or F	Unclear age group	NR	NR	Airway hyper-responsiveness (MCT)	Equivalent to FeNO at 1 or 2	Not relevant – device and cut-off values not reported and uses unknown device of unknown equivalence
Difficult to diagnose with chronic cough							
Hsu 2013[79]	Position F Chronic cough	Adults	280i Sievers; GE; Boulder, CO	33.9, 30 ppb	ICS responsiveness	FeNO replaces trial of treatment (whole treatment pathway)	Relevant, though diagnosis is of ICS responsiveness not asthma and uses device of unknown equivalence

Study author, year	Population	Age group	Device	Cut-off values	Reference standard	Position of FeNO in the pathway	Relevance to decision problem
Hahn 2007[80]	Position F Chronic cough	Adults	280i Sievers; GE; Boulder, CO	35, 38 ppb	ICS responsiveness	FeNO replaces trial of treatment (whole treatment pathway)	Relevant, though diagnosis is of ICS responsiveness not asthma and uses device of unknown equivalence
Prieto 2009[82]	Position F Chronic cough	Adults	Niox	20 ppb	ICS responsiveness	FeNO replaces trial of treatment	Relevant, though diagnosis is of ICS responsiveness not asthma
Sato 2008[81]	Position F Chronic cough	Adults	Chemi-luminescence analyzer (Kimoto, Osaka, Japan)	38.8 ppb	Cough with/without wheeze, sputum eosinophilia, airway reversibility, airway hyper-responsiveness (MCT)	No equivalent position in UK pathway (sputum eosinophilia)	Not relevant – reference standard not similar to common UK practice (sputum eosinophilia) and uses device of unknown equivalence
Zhang 2011[93]	Position F Chronic cough	Unclear age group	NIOX MINO	40, 31 ppb	Sputum eosinophilia, pulmonary function test, airway hyper-responsiveness, 24-h oesophageal pH monitoring, SPT and serum IgE	No equivalent position in UK practice	Not relevant – reference standard uses tests not used in UK standard practice
EIB							
El Halawani 2003[84]	Suspected EIB	Adults	Sievers 280A (Sievers Instruments, Boulder, CO)	12 ppb	Exercise challenge	NA	Relevant, but unclear if patient selection similar to patients who would be referred to exercise challenge test in UK and uses device of unknown

Study author, year	Population	Age group	Device	Cut-off values	Reference standard	Position of FeNO in the pathway	Relevance to decision problem
							equivalence
Other							
Arora 2006[85]	Mix of undiagnosed and diagnosed[a]	Adults	Niox	6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 46 ppb	Airway hyper-responsiveness (Histamine)	? equivalent to MCT, or later?	Not relevant - Too different to UK population, reference standard would not be applied to all UK patients.

ppb, parts per billion; FEV1, forced expiratory volume in first second; FVC, forced vital capacity; MCT, methacholine challenge test; ATS, American Thoracic Society; ICS, inhaled corticosteroids; WBP, whole body plethysmography; h, hour; EIB exercise induced bronchoconstriction

[a] Army recruits, some of whom are thought to have lied about existing asthma diagnosis

5.2.2.1.2 Studies relevant to the decision problem using only FeNO as the index test

From the initial 20 studies conducted in adults, adults plus adolescents, all age groups and unspecified age groups, 13 studies were considered to be of most or some relevance to the decision problem. These studies are Schneider 2013,[75] Schneider 2009,[77,78] Schleich 2012,[83] Prieto 2009,[82] El Halawani 2003,[84] Hsu 2013,[79] Hahn 2007,[80] Pedrosa 2010,[89] Heffler 2006,[86] Smith 2005,[87] De La Barra 2011,[88] Smith 2004,[90] Bobolea 2012[92] and Cordeiro 2011.[91]

Table 21 groups studies according to position on the pathway and reference standard, and tabulates the other key variables, study and patient characteristics. Appendix 10 provides more detail about the specifics of the reference standards used and Appendix 11 provides more detail about the patient inclusion and exclusion criteria. There are several main sources of heterogeneity amongst these studies that preclude meta-analysis of the results. These include: the age groups recruited; the spectrum of patients in terms of their position in the pathway and other restrictions in recruitment such as having rhinitis or chronic cough; the device used to measure FeNO; the reference standards used; and the cut-offs reported.

For each study, we have selected and presented in tables three sets of sensitivity and specificity estimates. These are:

- The highest sum of sensitivity and specificity as reported by the authors of the study
- The highest sensitivity – in this scenario a negative test result rules out a diagnosis (see Tables 3, 4 and 5 for details). This was selected as the cut-off that provided the highest sensitivity. Where 100% sensitivity was reported for more than one cut-off, the cut-off that maintained the highest specificity was selected. It should be noted that some studies did not report 100% sensitivity, though this may have been achievable at lower cut-off points. Where the cut-off with the highest sensitivity was not also the cut-off with the highest positive predictive value (PPV), this latter cut-off was also presented.
- The highest specificity – in this scenario, a positive test result rules in a diagnosis of asthma. Selected as for the highest sensitivity, but for specificity. Where the cut-off with the highest specificity was not also the cut-off with the highest negative predictive value (NPV), this latter cut-off was also presented.

It should be noted that superior sets of sensitivity and specificity may have in fact been achieved, but selection was limited to the range of cut-off points reported within studies.

Table 21: Diagnostic review: Study and patient characteristics of the thirteen studies considered of relevance to the decision problem

Study author, year	Study design, Funding	Country, setting, recruitment dates	Population	Device	Reference standard	n analysed/N recruited, reasons for withdrawals	Mean age in years (SD)	Gender, n male (%)	Severity FEV_1% (SD)	Mean FeNO	Smokers	Atopic
Position A versus whole pathway												
Schneider 2013[75]	Prospective, consecutive cohort study Funding NR: authors no conflicts	Germany Private practice run by 5 pneumologists June 2010 to October 2011	Adults Position A	NIOX MINO	FEV_1, FEV_1/FVC, airway reversibility, hyper-responsiveness (MCT)	393/400 (98%) 7 - lack of data	Asthma: 40.5 (15.4) COPD: 60.8 (17.0) No OAD 44.6 (16.5)	158/393 (40.2%)	Asthma: 101.3 (17.0) COPD: 74.1 (12.3) No OAD 107.7 (16.3)	Asthma:42.4 (46.4) COPD: 16.6 (6.8) No OAD 22.0 (16.5)	Current: 39/393 (9.9%) Ex: 139/393 (35.4%)	NR
Schneider 2009[77,78]	Prospective, consecutive cohort study Funding: Government	Germany Primary care - 14 GPs across 10 practices Feb 2006 to June 2007	Adults Position A	NIOX MINO	FEV_1, FEV_1/FVC, airway reversibility, airway hyper-responsiveness MCT	160/160 (100%)	43.9	72/160 (45%)	Asthmatics (n=75) 100 (12.2) COPD (n=25) 67.8 (18.5) Overlap (n=8) 68.8 (18.4) No OAD (n=52)107.4 (12.8)	Asthmatics (n=75) 42.6 (47.9) COPD (n=25) 16.2 (11.1) Overlap (n=8) 20.4 (18.6) No OAD (n=52)24.7 (16.0)	Current and ex 86/160 (54%)	NR
Smith 2005[87]	Prospective, consecutive cohort study Funding: Mix of industry and non-industry but not device manufacturer	New Zealand Secondary care, 1 centre Dates NR	Adults and adolescents Position A	Niox	Airway reversibility, positive response to ICS, airway hyper-responsiveness (MCT).	52/60 (87%) 8 withdrew: 3 withdrew consent, 1 respiratory	40.5 (range 14 to 71)	20/52 (38.5%)	97.8 (14.2)	Range 6.3 to 242.0ppb (De la Barra 2011)[88]	Current 3/52 (5.8%); Ex 10/52 (19.2%);	40/52 (77%) (De la Barra 2011)[88]

Study author, year	Study design, Funding	Country, setting, recruitment dates	Population	Device	Reference standard	n analysed/N recruited, reasons for withdrawals	Mean age in years (SD)	Gender, n male (%)	Severity FEV_1% (SD)	Mean FeNO	Smokers	Atopic
						tract infection, 1 acute rhinitis, 3 LTFU						
Smith 2004[90]	Prospective, consecutive cohort study Funding: Mix of industry and non-industry but not device manufacturer	New Zealand Secondary care, 1 centre Dates NR	All patients	NR	Airway reversibility, positive response to ICS, airway hyper-responsiveness (MCT).	44/51 (86%) 7 withdrew: 4 withdrew consent (time) 3 technical difficulties	Asthmatics (n=17) 41.6 (range 9 to 72) Non-asthmatics (n=30) 31.8 (range 9 to 64)	20 (42.6%)	Asthmatics (n=17) 90.5 (18.4) Non-asthmatics (n=30) 110.0 (13.5)	Asthmatics (n=17) 52 (34.0) Non-asthmatics (n= 30) 15.7 (12.9)	Ex 5/47 (10.6%)	NR
Position A versus airway reversibility												
De La Barra 2011[88]	Prospective, consecutive cohort study Funding: Aerocrine AB, Solna, Sweden and from Lottery Grants New Zealand	New Zealand Secondary care, 1 centre Dates NR	Adults and adolescents Position A	Niox	Airway reversibility	52/60 (87%) 8 withdrew: 3 withdrew consent, 1 respiratory tract infection, 1 acute rhinitis, 3 LTFU	40.5 (14 to 71) (Smith 2005)[87]	20 (38.5%) (Smith 2005)[87]	97.8 (14.2) (Smith 2005)[87]	Ranged from 6.3 to 242.0 ppb	Current 3/52 (5.8%); Ex 10/52 (19.2%) ;	40/52 (77%)
Subset of Position A versus airway reversibility or airway hyper-responsiveness												

Study author, year	Study design, Funding	Country, setting, recruitment dates	Population	Device	Reference standard	n analysed/N recruited, reasons for withdrawals	Mean age in years (SD)	Gender, n male (%)	Severity FEV_1% (SD)	Mean FeNO	Smokers	Atopic
Heffler 2006[86]	Prospective, consecutive cohort study Funding: Government/non-industry	Italy Allergy and Immunity Clinic Dates NR	Adults and adolescents WITH rhinitis Position A	Niox	Airway hyper-responsiveness (MCT) or airway reversibility	48/48 (100%)	40.08 (SD: NR)	21 (43.75%)	89.2 (95% CI 80.1 to 98.4)	59.7 (95% CI 50.2 to 89)	0	35/38 (92.1%)
Cordeiro 2011[91]	Retrospective (analysis of prospective data base) Funding: NR, authors no conflicts	Netherlands Secondary care Jan 2007 to Sept 2007	All ages with high prevalence of atopy Position A	Niox-Flex	Airway reversibility, airway hyper-responsiveness (histamine)	114/114 (100%)	median (range) Asthmatic n=42: 39 (7 to 83); non Asthmatic n=72: 38 (7 to 87)	43/114 (37.7%)	FEV_1/FVC % median (range) Asthmatic 70 (42 to 95); Non asthmatic 77(69 to 95)	Asthmatics n=42; median 44ppb range (6 to 290) Non asthmatics n=72; median 17ppb range (5 to 45)	11/114 (9.6%)	81/114 (71.1%)
Difficult to diagnose versus airway hyper-responsiveness												
Schleich 2012[83]	Prospective cohort study Funding: Non-industry	Belgium, secondary care Dates March 2009 to Dec 2009	Adults with chronic cough Position E	Niox	Airway hyper-responsiveness (MCT)	174/237 (73%) 63 did not meet inclusion criteria	41 (16)	72/174 (41%)	97% (13)	Median 17ppb range (4 to 271)	59/174 (33.9%)	84/174 (48%)
Pedrosa 2010[89]	Prospective, consecutive cohort study Funding: NR, authors no conflicts	Spain Secondary care Dates NR	Adults and adolescents Position E	NIOX MINO	Airway hyper-responsiveness (MCT)	114/115 (99%) 1 withdrawal, reason NR	34 (13)	N=115: 72/115 (62.6%)	N=115: 104.29 (14.95)	N=115: 34ppb	N=115: current: 17 (14.8%) ex: 11 (9.6%)	N=115: Atopy: 100 (87%)

Study author, year	Study design, Funding	Country, setting, recruitment dates	Population	Device	Reference standard	n analysed/N recruited, reasons for withdrawals	Mean age in years (SD)	Gender, n male (%)	Severity FEV_1% (SD)	Mean FeNO	Smokers	Atopic
Bobolea 2012[92]	Prospective, consecutive cohort study Funding: NR	Spain Assuming secondary care Dates NR	All ages Position F	NIOX MINO	Adenosine challenge test	30/30 (100%)	37.3 (13-69)	13/30 (43.3%)	NR	NR	NR	NR
Suspected EIB versus Exercise challenge test												
El Halawani 2003[84]	Prospective, consecutive cohort study Funding: NR	USA Naval Medical Centre Dates NR	Adults Suspected EIB	Sievers 280A	Exercise challenge	49/50 (98%) 1 - inability to complete spirometry	27.9 (SD: NR)	35 (71.4%)	NR	EIB group, (n=7): 41ppb; Non-EIB (n=42): 25.6 ppb	0	0
Position F with chronic cough versus ICS responsiveness												
Prieto 2009[82]	Prospective cohort study, unclear if consecutive Funding: None	Spain Allergy or respiratory clinics Dates NR	Adults with chronic cough Position F	Niox	ICS responsiveness	43/43 (100%)	48 (95% CI 43 to 52)	18/43 (41.9%)	113.2 (95% CI 108.0 to 118.3)	GM Mean (95% CI): responders to ICS 23.2 (17.5 to 30.7) non responders 18.6 (14.7 to 24.0)	0/43 (0%)	43/43 (100%)
Hsu 2013[79]	Retrospective cohort study Funding: NR	Taiwan Asthma and cough-specific clinic June 2007 to May 2008	Adults with chronic cough Position F	280i Sievers	ICS responsiveness	81/114 (71%) 33 (26 lost after first visit, 7 stopped coughing after 1-2	49 (14)	33/81 (40.7%)	91.8 (15.3)	mean rank FeNO by Kruskal-Wallis test: 47ppb	0/81 (0%)	NR

Study author, year	Study design, Funding	Country, setting, recruitment dates	Population	Device	Reference standard	n analysed/N recruited, reasons for withdrawals	Mean age in years (SD)	Gender, n male (%)	Severity FEV_1% (SD)	Mean FeNO	Smokers	Atopic
						week treatment for UACS and GORD)						
Hahn 2007[80]	Retrospective cohort study Funding: NR, authors no conflicts	Secondary care. Mayo Clinic, Rochester, USA December 2004 to November 2005	Adults with chronic cough Position F	280i Sievers	ICS responsiveness	64/64 (100%)	Pooled weighted mean: 46.8 (SD: NR)	26/64 (40.6%)	ICS unresponsive group: 98; responsive group: 94	ICS unresponsive group: 26.0 ±16.5; responsive group: 51.25 ±20.1	Current: 0/64 Ex: 10/64 (15.6%)	NR

FeNO, fractional exhaled nitric oxide; FEV_1, forced expiratory volume in first second; FVC, forced vital capacity; COPD, chronic obstructive pulmonary disease; OAD, obstructive airway disease; GP, general practitioner; MCT, methacholine challenge test; ICS, inhaled corticosteroids; LTFU, lost to follow-up; n, number; N total number; NR, not reported; CI, confidence interval; Dec, December; EIB, exercise induced bronchoconstriction; GM, geometric mean; SD, standard deviation; UACS, upper airway cough syndrome; GORD, gastro-oesophageal reflux disease.

a. Quality assessment of studies relevant to the decision problem

Thirteen studies (fourteen references) exploring FeNO measurement for the diagnosis of asthma in adults were assessed for quality according to QUADAS-2 criteria[34] for diagnostic accuracy studies. Although based on the same data as Smith 2005,[87] de la Barra 2011[88] was assessed separately as the analysis was different.

The overall quality was variable, with the Smith 2005 study[87] scoring well on all the domains, and thus being at least risk of bias. The studies at highest risk appeared to be Hsu[79] and Cordeiro,[91] neither of which provided sufficient information on the nature of blinding for index and reference standard tests. There were also some issues in terms of patient flow in both studies: In Cordeiro[91] patients did not all receive the same reference test (MCT was only provided if asthma was suspected from other tests). Similarly, in Hsu,[79] the reference standard was allocated based on an algorithm, rather than an *a priori* set of tests (Figure 11). The risk of bias from the conduct of the index test scored worst overall with only one study scoring positively for this domain. Studies scored poorly for risk of bias from the conduct of the reference standard, with nine scoring unclear for this domain.

Figure 11: Risk of bias summary: review authors' judgements about each risk of bias item for each included study

Green circles with + sign, low risk of bias; red circles with – sign, high risk of bias; yellow circles with ?, unclear risk of bias.

Risk of bias from patient selection

As far as we could ascertain, patient selection did not appear to be a source of bias in the body of literature. All studies avoided case control design, and recruited appropriately (i.e. those patients presenting with clinical signs of asthma, a subset thereof, or patients at a definable point in the UK pathway). However, it was unclear in five cases (Schleich;[83] Prieto,[82] Cordeiro,[91] Hahn;[80] Hsu;[79]) whether a consecutive sample was recruited, and in three cases this was because the study was retrospective (Hahn;[80] Hsu;[79] Cordeiro[91]).

Risk of bias from the conduct of the index test

The conduct of the index test was a potentially important source of bias, with only Smith2005[87] and Smith 2004[90] being free from bias in this domain as a whole. There were two component questions for this domain, one relating to blinding, and one to whether the study was a derivation study or a

validation study. Seven studies were unclear as to whether the index test was interpreted blind to the reference standard: Schneider 2009;[77] Schleich;[83] Pedrosa;[89] Hsu;[79] Heffler;[86] Hahn;[80] and Bobolea.[92] The De la Barra[88] and Cordeiro[91] studies were not explicit with regard to the blinding of reference standard results; however, as the index test was performed prior to the reference standard, it would not have been possible for the investigator to be aware of the reference results at the time of the index test unless interpretation was not performed at the time of the test. This would seem unlikely as FeNO measurement done according to standardised protocols is objective and interpretation is not required.

Several further studies were at potential risk of bias in that they were derivation studies which fitted cut-off points to the data *post-hoc*, and were thus likely to over-estimate accuracy. These studies were: Cordeiro,[91] De la Barra,[88] El Halawani,[84] Hahn,[80] Heffler,[86] Pedrosa,[89] Prieto,[82] Schleich,[83] and Schneider 2013.[75]

Risk of bias from the conduct of the reference standard

The reference standard and its interpretation was a further source of bias among much of the literature, with only Prieto,[82] the two Schneider studies,[75,77] and Smith 2005[87] being free of bias. It was not possible to ascertain in any of the remaining literature whether the operator conducting the reference standard had been blinded to the results of the index test.

Risk of bias from patient flow and timing of the study

For the most part, there was little concern about the patient flow and study timing. However, there were at least two studies which did not provide an identical reference standard for all patients: In Cordeiro,[91] patients only received MCT if asthma was suspected based on other tests, while in Hsu,[79] reference standard provision was algorithm-based, with some patients not receiving ICS treatment. It was not necessarily clear in all other studies whether patients received all reference standard tests or a sequence. In other respects, the patient flow and study timing was satisfactory. Drop-out rates were low, and, where drop-outs occurred, these were adequately accounted for in the study reports.

Summary

The corpus of included literature was of variable quality, with Smith 2005[87] being at the least risk of bias, and the two Schneider studies[75,77] and the earlier Smith 2004 study[90] also performing well. The conduct of the index test was identified as a potentially serious source of bias among the literature, with few studies providing adequate information on how blinding to the reference test results was achieved. Nine of the 13 studies were derivation studies and, in fitting cut-off points to the data *post-hoc*, are likely to over-estimate the accuracy of FeNO as a diagnostic test. With the reference standard conduct too, few studies provided satisfactory information on how operators were blinded to the results of the index test. However, it is important to stress that this may reflect lack of clarity in the

study reports, rather than in the conduct of the reference test itself. The likelihood of unblinding biasing the results is therefore unclear.

b. Studies recruiting patients at Position A

Position A is the start of the UK pathway. Patients will have undergone no other tests. The reference standard used in studies which recruit patients at this position will determine whether the results relate to a scenario where FeNO is replacing the whole pathway or a scenario where it is replacing just one test within the pathway. Where replacing just one test, it could be used as a rule-in scenario; patients testing positive would go on to be treated as asthmatic and patients testing negative would go on to have further tests for asthma. Where a rule-out scenario is used, patients testing positive would go on to have further tests and patients testing negative would go on to be treated as not asthmatic.

i. Position A versus whole pathway

Population: Four studies (Schneider 2013,[75] Schneider 2009,[77,78] Smith 2005[87] and Smith 2004[90]) recruited patients with symptoms of asthma who had not undergone any other tests. These studies are unlikely to have recruited the full spectrum of patients at this point in the pathway due to common exclusions such as those who experienced a respiratory infection in the last month, and those taking ICS (see Appendix 11). As in many cases a GP may provide a patient with ICS before confirmation of asthma, these exclusions may result in a patient spectrum that does not reflect UK practice.

Of the four studies, Schneider 2013[75] and Schneider 2009[77,78] recruited adults, Smith 2005[87] recruited adults and adolescents and Smith 2004[90] recruited patients of any age. The largest study was Schneider 2013[75] with 393 participants, and the smallest was Smith 2004[90] with 44 participants. Mean ages, FEV_1% and FeNO values were not always reported for the whole cohort making it difficult to compare across studies. All recruited more females than males with proportion of males ranging from 38.5 to 45%. Schneider 2013,[75] Schneider 2009[77,78] and Smith 2005[87] recruited a mix of smokers, ex-smokers and non-smokers whilst Smith 2004[90] did not recruit any smokers, though this was not listed as an exclusion criteria and may be due to the small sample size. Only Smith 2005[87] reported how many participants were atopic, with a high prevalence of 77%, though the other three studies did not list atopy as an exclusion criterion making it likely they included a proportion of atopic patients. Schneider 2013[75] excluded pregnant women.

Intervention: Schneider 2013[75] and Schneider 2009[77,78] both used NIOX MINO. Smith 2005[87] used Niox and Smith 2004[90] did not report the device used.

Reference standard: The reference standard for Schneider 2013[75] and Schneider 2009[77,78] was airway reversibility or airway hyper-responsiveness (depending on spirometric test results), whilst in both

Smith 2005[87] and Smith 2004[90] the reference standard also incorporated ICS responsiveness. Whilst these reference standards do differ, bronchodilator reversibility and ICS responsiveness appear to be used interchangeably in the UK pathway, so both reference standards are equivalent to the whole pathway. However, it is likely that these reference standards may differentially influence estimates of FeNO diagnostic accuracy as FeNO would be expected to correlate better with ICS responsiveness than airway reversibility testing with a bronchodilator.

Study design and setting: All studies were prospective, consecutive cohort studies and none of the studies were funded by the manufacturers of a FeNO device. Schneider 2013[75] and Schneider 2009[77,78] were conducted in Germany in primary care or a private practice whilst Smith 2005[87] and Smith 2004[90] were conducted in New Zealand in secondary care.

Estimates of diagnostic accuracy: Table 22 details the estimates of sensitivity and specificity for these studies. Results did not appear to be similar between studies. The cut-off for the highest sum of sensitivity and specificity varied from 20 ppb to 47 ppb and this did not appear to be dependent on any variable. Schneider 2013,[75] Schneider 2009[77,78] and Smith 2005[87] all reported higher specificity values than sensitivity values, whilst Smith 2004[90] reported the opposite. This study recruited a mixed population of adults and children, and also did not report the device used to measure FeNO. Sensitivities varied greatly across studies, and ranged from 32% to 88%. Specificities were more consistent across studies and ranged from 75% to 93%.

Rule-out cut-off points varied from 9 ppb to 16 ppb with sensitivities between 69% and 96%, specificities between 13% and 53%, PPV between 29.4% and 56.5%, and NPV between 37.1% and 83.8%. Rule-in cut-off points varied from 47ppb to 76 ppb, with specificities between 92% and 100%, sensitivities between 13% and 55.6%, PPV between 79.5% and 100% and NPV between 56.7% and 65.7%. Schneider 2013[75] and Schneider 2009[77,78] reported very similar rule-in (71 ppb and 76 ppb respectively) and rule-out (9 ppb and 12 ppb respectively) cut-off points, but the cut-off providing the highest sum of sensitivity and specificity was not similar between these two studies (25 ppb and 46 ppb respectively). Smith 2005[87] reported a similar rule-out cut-off point (15 ppb) to these studies, but a quite different rule-out cut-off point (47 ppb). Only Schneider 2009[77,78] reported a 100% PPV that would reliably rule patients in and no studies report a 100% NPV.

ii. Position A versus airway reversibility

De La Barra 2011[88] performed a secondary analysis of the data from Smith 2005[87] against a reference standard of airway reversibility only. This is equivalent to replacing airway reversibility with FeNO, or placing FeNO before airway reversibility as a rule-in test or rule-out test, with patients going on to receive this and further tests as appropriate. The cut-off point with the best sum of sensitivity and

specificity seemed fairly similar to that reported in Smith 2005[87] at 41.7ppb compared to 47ppb. The rule-out cut-off point was somewhat higher at 25ppb in De La Barra 2011[88] compared to 15ppb in Smith 2005[87] and the rule-in cut-off point was higher at 110ppb (or 90ppb if selecting the cut-off with the highest NPV) compared to 47ppb respectively. Sensitivity and specificity values were also different (see Table 22).

iii. Subset of patients at Position A versus airway reversibility or airway hyper-responsiveness

Population: These studies recruited patients who may represent a narrower selection of the full spectrum of patients who present with symptoms of asthma than described in the previous studies. Heffler 2006[86] recruited 48 adults and adolescents with rhinitis and symptoms of asthma whilst Cordeiro 2011[91] recruited 114 patients with a "high prevalence of atopy" as reported by the authors. However, it would appear that these two studies are in fact reasonably comparable to the studies that recruited a full spectrum of patients at Position A. The prevalence of atopy in Heffler 2006[86] is higher than the prevalence in Cordeiro 2011[91] at 92% compared to 71%. Smith 2005,[87] was the only study to report prevalence of atopy from the studies that recruited the fuller spectrum of patients at Position A and this study reported a similar prevalence of 77%. Similar to the previous studies, Heffler 2006[86] did not recruit any smokers, whilst Cordeiro 2011[91] recruited 9.6%. Mean ages were similar at around 40 years, and severity measures and FeNO values were not reported in a way that allowed comparison between studies.

Intervention: Heffler 2006[86] used Niox and Cordeiro 2011[91] used Niox-Flex, which may be equivalent to NIOX MINO.[40]

Reference standard: Both studies used a combination of airway reversibility and airway hyper-responsiveness as the reference standard, which is equivalent to the whole UK pathway.

Study design and study setting: Heffler 2006[86] is a prospective consecutive cohort study conducted in Italy in an allergy and immunity clinic. Cordeiro 2011[91] is a retrospective analysis of a prospective database conducted in the Netherlands in secondary care. Neither study was funded by manufacturers of FeNO devices.

Estimates of diagnostic accuracy: Table 22 details the estimates of sensitivity and specificity for these studies. In Heffler 2006,[86] the highest sum of sensitivity and specificity do not seem noticeably different to studies recruiting the fuller spectrum of patients with symptoms of asthma (77.8% and 60% respectively), though both rule-in and rule-out scenarios achieved 100% in their respective specificity and sensitivity, which was not achieved by the studies with a full spectrum of patients at

Position A. This study also reported higher values for the paired sensitivity and specificity in the rule-in, rule-out scenarios compared to the studies with a fuller spectrum of patients at Position A. In Cordeiro 2011,[91] sensitivity and specificity were 78% and 92%, which has a similar sum to the highest pair of sensitivity and specificity reported for studies recruiting the full spectrum of patients, reported by Smith 2004[90] at 88% and 79% respectively, but with the balance between sensitivity and specificity inverted.

c. Studies recruiting patients who are difficult to diagnose of relevance to the decision problem

The most appropriate reference standard in the difficult to diagnose population in relation to UK guidelines varies according to where in the pathway the group is recruited from, in other words, which tests they have already undergone, and which test they would get next. As previously described, where a rule-in scenario is used, patients testing positive would go on to be treated as asthmatic and patients testing negative would go on to have further tests for asthma. Where a rule-out scenario is used, patients testing positive would go on to have further tests and patients testing negative would go on to be treated as not asthmatic.

Population: Three studies recruited patients who fall into the difficult to diagnose category were of relevance to the decision problem, but none of them recruited patients at exactly the same point in the pathway (see Appendix 11 for more details of study inclusion criteria). Schleich 2012[83] recruited adults with chronic cough who had a negative test for airway reversibility and normal spirometry (Position E in UK pathway). Pedrosa 2010[89] also recruited patients at Position E, but this was not restricted to those with chronic cough, and included adolescents as well as adults. Bobolea 2012[92] recruited a somewhat different spectrum of patients who were of all ages and who had a negative test for airway reversibility, normal spirometry and a negative MCT. Mean ages were between 34 and 41 years of age, FEV_1% were similar at 97% and 104% where reported, though FeNO values were not reported in a way that allowed comparison between studies. Both Schleich 2012[83] and Pedrosa 2010[89] recruited smokers and atopic patients, though the prevalence of atopy was higher in Pedrosa 2010[89] at 87% compared to 48%.

Intervention: Pedrosa 2010[89] and Bobolea 2012[92] used NIOX MINO, whilst Schleich 2012[83] used Niox.

Reference standard: Schleich 2012,[83] Pedrosa 2010[89] and Bobolea 2012[92] all used airway hyper-responsiveness as the reference standard, which was appropriate to the UK pathway for the patients they selected. Schleich 2012[83] and Pedrosa 2010[89] used MCT as the reference standard. Bobolea 2012[92] used an adenosine challenge test as patients had already had a negative MCT test.

Estimates of diagnostic accuracy: Table 22 details the estimates of sensitivity and specificity for these studies. Schleich 2012[83] and Pedrosa 2010[89] reported quite similar cut-offs for the highest sum of sensitivity and specificity at 34 and 40ppb respectively, but the paired estimates of sensitivity and specificity were different with Schleich 2012[83] reporting sensitivity of 35% and specificity of 95% and Pedrosa 2010[89] reporting 74.3% and 72.5% respectively. Based on only two studies, it is unclear whether the difference in estimates is due to the selection of chronic cough only patients by Schleich 2012,[83] or due to some other factor such as natural variation. Bobolea 2012[92] reported 100% sensitivity, but only 29.2% specificity, indicating that FeNO in this Position would be most likely to be useful as a rule-out test. No data were available for other cut-off points.

In comparison to studies which recruited patients at Position A in the pathway, patient populations are perhaps somewhat younger. Other patient spectrum characteristics look comparable. The range of estimates of sensitivity and specificity also look largely comparable. 100% sensitivity was achieved by Bobolea 2012,[92] though it is not clear if this was for the highest sum of sensitivity and specificity, or if the cut-off was selected so that FeNO could perform as a rule-out test with high sensitivity.

d. Studies recruiting patients with chronic cough at Position F

Population: Prieto 2009,[82] Hsu 2013[79] and Hahn 2007[80] all recruited adults with chronic cough who were negative for some other causes of cough (see Appendix 11 for more detail of inclusion and exclusion criteria). Prieto 2009[82] recruited patients with FEV_1 of at least 80% predicted with chronic cough and no signs of other lung disease. Hsu 2013[79] recruited patients who were negative for upper airway cough syndrome and gastro-oesophagela reflux disease (GORD), and had no obvious chest x-ray abnormalities. Hahn 2007[80] recruited patients with normal chest radiographs. All three appear equivalent to patients at Position F in the UK pathway. All three recruited no current smokers and only Prieto 2009[82] reported the prevalence of atopy, and this was 100%. Cohort were perhaps somewhat older than in other studies with all averaging in the mid to high 40's.

Intervention: The device used was Sievers 280A in Hsu 2013[79] and Hahn 2007[80] NIOX MINONObreath, whilst Prieto 2009[82] used Niox.

Reference standard: The reference standard was ICS responsiveness, which would be the next test in UK practice for some or all of these patients.

Estimates of diagnostic accuracy: Table 22 details the estimates of sensitivity and specificity for these studies. In Prieto 2009[82] sensitivity and specificity were poor at 53% and 63% respectively, though Hsu 2013[79] and Hahn 2007[80] both report high sensitivities (94.7% and 90% respectively) and fairly high specificities (76.3% and 85% respectively), indicating that FeNO could be a useful rule-out

test. It is not clear why the estimates reported by Prieto 2009[82] differ from the other two similar studies, though this may be due to differences in patient selection. Whilst the device used by Prieto 2009[82] is Niox, it is not thought that the device would alter estimates of diagnostic accuracy, but rather the cut-off points derived.

e. Other studies of some interest to the assessment

Table 22 details the estimates of sensitivity and specificity for these studies. El Halawani 2003[84] recruited adults with suspected exercise induced bronchoconstriction. As with Arora 2006,[85] which was not considered relevant to the review due to the reference standard, this group of patients were army recruits. It is not clear what previous tests these patients had undergone, if any. None of the patients were smokers or atopic. The reference standard was exercise challenge test, which will only identify patients with exercise induced bronchoconstriction rather than other forms of asthma. The device used (Sievers 280A) is of unknown equivalence to NIOX MINO and NObreath. The study reports 100% sensitivity and 31% specificity, indicating that this test could be used as a rule-out test.

f. **Diagnostic accuracy meta-analysis**

From Table 21 it can be seen that only two sets of two studies are similar enough to each other to warrant meta-analysis:

- Schneider 2009[77,78] and Schneider 2013,[75] two studies conducted by the same research group with populations recruited in 2006 to 2007, and 2010 to 2011 respectively
- Hsu 2013[79] and Hahn 2007,[80] which recruited in 2009 to 2010 and 2004 to 2005 respectively, and were conducted in different countries (China and USA respectively).

However, the value of such a meta-analysis is limited given that these studies are no more or less relevant to the decision problem than any of the other studies found.

Table 22: Diagnostic review: Diagnostic accuracy of FeNO tests in adults, adults and adolescents and all ages. A table of all results reported by these studies is given in Appendix 12

					Highest sum of sens and spec					Rule-out					Rule-in				
Study author, year	**Population**	**Device**	**N**	**Reference standard**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**
Position A versus whole pathway																			
Schneider 2013[75]	Adults Position A	NIOX MINO	393	FEV_1, FEV_1/FVC, airway reversibility, hyper-responsiveness (MCT)	25	49	75	56.0	69.5	9	96	13	41.6	83.8	71	18	97	79.5	64.6
Schneider 2009[77,78]	Adults Position A	NIOX MINO	160	FEV_1, FEV_1/FVC, airway reversibility, airway hyper-responsiveness (MCT).	46	32	93	77.3	59.5	12 16	85 69	24 53	49.6 56.5	64.5 66.2	76	13	100	100	56.7
Smith 2005[87]	Adults and adolescents Position A	Niox	52	Airway reversibility, positive response to ICS, airway hyper-responsiveness (MCT).	47	55.6	92	88.2	65.7	15	81.5	48	29.4	37.1	As highest sum				
Smith 2004[90]	All patients	NR	44	Airway reversibility, positive	20	88	79	70	91.7	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

					Highest sum of sens and spec					Rule-out					Rule-in				
Study author, year	**Population**	**Device**	**N**	**Reference standard**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**
				response to ICS, airway hyper-responsiveness (MCT).															
Position A versus airway reversibility																			
De La Barra 2011[88]	Adults and adolescents Position A	Niox	52	Airway reversibility	41.7	NR	NR			25	83.3	57.5	37.0	92	110 90	25 41.7	95 92.5	60 62.5	80.9 84.1
Subset of Position A versus airway reversibility or airway hyper-responsiveness																			
Heffler 2006[86]	Adults and adolescents WITH rhinitis Position A	Niox	48	Airway hyper-responsiveness (MCT) or airway reversibility	36	77.8	60	53.8	81.8	25	100	46.7	52.9	100	100	27.8	100	100	69.8
Cordeiro 2011[91]	All ages with high prevalence of atopy Position A	Niox-Flex	114	Airway reversibility, airway hyper-responsiveness (histamine)	27	78	92	84.6	88	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
		Niox-Flex Airway		Airway reversibility, airway hyper-responsive	27	87	90			NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

					Highest sum of sens and spec					Rule-out					Rule-in				
Study author, year	**Population**	**Device**	**N**	**Reference standard**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**
		reversibility		ss (histamine)															
Difficult to diagnose versus airway hyper-responsiveness																			
Schleich, 2012[83]	Adults with chronic cough Position E	Niox	174	Airway hyper-responsiveness (MCT)	34	35	95	87.8	62.4	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
		Niox FEV_1 ≤101%		Airway hyper-responsiveness (MCT)	34	24.4	98.9			NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Pedrosa 2010[89]	Adults and adolescents Position E	NIOX MINO	114	Airway hyper-responsiveness (MCT)	40	74.3	72.5	54.1	86.3	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Bobolea 2012[92]	All ages Position F	NIOX MINO	30	Adenosine challenge test	30*	100	29.2	26	100	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Suspected EIB versus Exercise challenge test																			
El Halawani 2003[84]	Adults Suspected EIB	Sievers 280A	49	Exercise challenge	12*	100	31	19.4	100	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Position F with chronic cough versus ICS responsiveness																			
Prieto 2009[82]	Adults with chronic	Niox	43	ICS responsiveness	20	53	63	52.6	62.5	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

					Highest sum of sens and spec					Rule-out					Rule-in				
Study author, year	**Population**	**Device**	**N**	**Reference standard**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**
	cough Position F																		
Hsu 2013[79]	Adults with chronic cough Position F	280i Sievers	81	ICS responsiveness	33.9	94.7	76.3	80	94	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Hahn 2007[80]	Adults with chronic cough Position F	280i Sievers	64	ICS responsiveness	38	90	85	89.5	84.6	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

N, number analysed; Sens, sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value; FeNO, fractional exhaled nitric oxide; FEV_1, forced expiratory volume in first second; FVC, forced vital capacity; MCT, methacholine challenge test; ICS, inhaled corticosteroidsNR, not reported; EIB, exercise induced bronchoconstriction;

5.2.2.1.3 Studies using FeNO in conjunction with another test as the index test

Studies which report estimates of the diagnostic accuracy of FeNO in conjunction with other tests are

From the initial 20 studies conducted in adults, adults plus adolescents, all age groups and unspecified age groups, two (Schleich 2012[83] and Cordeiro 2011[91]) reported diagnostic accuracy data for FeNO in conjunction with another test as the index test. One further sudy (Fortuna 2007[76]) did not report actual data, but did state that addition of certain tests (see next paragraph) did not increase accuracy.

The study characteristics for Schleich 2012,[83] Cordeiro 2011[91] and Fortuna 2007[76] are presented in Table 21 and the diagnostic accuracies for Schleich 2012,[83] Cordeiro 2011[91] are presented in Table 22. Neither study reported a change in the optimum cut-off for FeNO when using it in conjunction with another test, but sensitivities and specificities did change. Fortuna 2007[76] was a study not judged to be of high relevance to the decision problem as it used sputum eosinophilia as part of the reference standard. This test is not widely available in the UK and so this study has low generalisability. The study reported that the addition of sputum eosinophilia to FeNO measurements increased specificity from 64% to 76%; sensitivity was not reported for the two tests together. The authors also stated that the addition of lung function tests and bronchodilator tests did not increase accuracy, but actual data were not provided.

Cordeiro 2011[91] used FeNO with a cut-off of 27ppb in conjunction with airway reversibility in a population of patients at Position A in the UK pathway. If patients were positive by either test they were considered to have tested positive. When compared to using FeNO at a cut-off of 27ppb alone, sensitivity increased from 78% to 87%, whilst specificity decreased from 92% to 90%. However, it should be noted that the reference standard for this study was airway reversibility or airway hyper-responsiveness to histamine. As such, the study results are at high risk of incorporation bias as the reference standard incorporates some of the same results as the index test. This is likely to overestimate the actual diagnostic accuracy of this combination of tests.[96]

Schleich 2012[83] used FeNO with a cut-off of 34ppb in conjunction with an FEV_1% predicted ≤101% in a population of patients with chronic cough and at Position E (difficult to diagnose) in the diagnostic pathway. Patients were required to have both a FeNO >34ppb and an FEV_1% predicted ≤101% to be judged positive by this combination of tests. This resulted in an increase in specificity from 95% to 98.9%, but a decrease in sensitivity from 35% to 24.4%. In this case the reference standard was airway hyper-responsiveness to MCT so incorporation bias was avoided.

Conclusions: In both cases the improvements in diagnostic accuracy are modest (or negative when considering the sum of sensitivity and specificity), and necessitate the usual trade off between

sensitivity and specificity. As both studies are derivation studies rather than validation studies (where the cut-off points are pre-set) it is possible that the gains seen are an overestimate of increases in diagnostic accuracy. However, it would seem that using a combination of tests may have additional benefit to using FeNO on its own, and these studies equate more accurately to adding FeNO in to the pathway than studies that do not use FeNO in conjunction with other tests.

5.2.2.2 Studies including children or children and adolescents

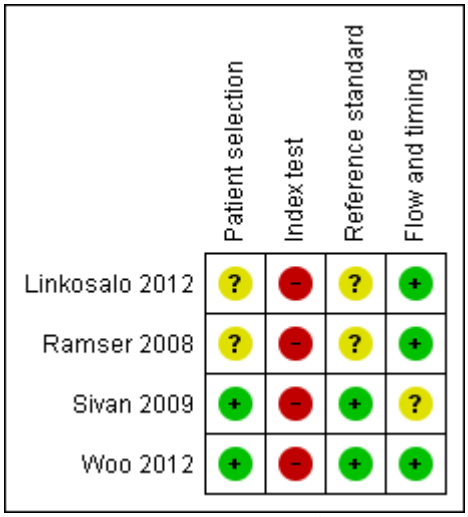
5.2.2.2.1 Studies using only FeNO as the index test

Four studies which recruited children (plus adolescents and/or young adults) and compared FeNO-guided diagnosis to non-FeNO-guided diagnosis were identified. All the studies were based in secondary care, and each study was undertaken in a different country: Finland (Linkosalo 2012[97]), Switzerland (Ramser 2008[98]), Israel (Sivan 2009[99]), and Korea (Woo 2008[100]). Funding sources were reported only in the articles by Linkosalo 2012[97] and Woo 2008[100]: these sources were the Tampere Tuberculosis Foundation /Medical Research Fund of Tampere University Hospital; and the National Research Foundation of Korea, respectively. Sivan 2009[99] declared that there were no conflicts of interest in their research.

a. Quality assessment

Four studies exploring FeNO measurement for the diagnosis of asthma in children were assessed for quality according to QUADAS-2[34] criteria for diagnostic accuracy studies. The overall quality was variable, with no one study being free from potential bias in all domains, and no single domain being free from bias in all studies. The Woo study[100] appeared to be at the lowest risk of bias, while Ramser[98] displayed the highest risk (Figure 12).

Figure 12: Risk of bias summary: review authors' judgements about each risk of bias item for each included study



Green circles with + sign, low risk of bias; red circles with – sign, high risk of bias; yellow circles with ?, unclear risk of bias.

Risk of bias from patient selection

Both Sivan[99] and Woo[100] appeared to be free of bias in terms of patient selection. Both studies enrolled consecutive samples, avoided case control design, and recruited appropriately (i.e. those patients presenting with clinical signs of asthma). However, there were potential sources of bias in Linkosalo[97] and Ramser,[98] in that neither study explicitly clarified whether they had enrolled patients consecutively.

Risk of bias from the conduct of the index test

There was potential bias in the conduct of the index test throughout the corpus of literature. All four studies[97-100] were derivation studies; hence, in fitting the cut-off points to the data *post-hoc*, they are likely to provide liberal estimates of diagnostic accuracy. There was an additional source of possible bias in the Ramser study,[98] in that it was not clear whether the index test was interpreted blind to the results of the reference standard.

Risk of bias from the conduct of the reference standard

Sivan[99] and Woo[100] both appeared to provide a satisfactory reference standard, in that both adhered to all or part of the UK guidelines and clearly stated that the results were interpreted by a blinded investigator. Neither Linkosalo[97] nor Ramser[98] provided sufficient information to confirm whether the result was interpreted blind to index test results.

Risk of bias from patient flow and timing of the study

The patient flow and test timing appeared to be broadly satisfactory. Linkosalo,[97] Ramser[98] and Woo[100] each conducted tests consecutively, provided the same reference standard to all patients, and included all enrolled patients in the final analysis. The one study that did display a potential source of bias in this domain was Sivan.[99] These investigators provided a list of criteria that may have been used to confirm a diagnosis of asthma, but it was not clear precisely which of these tests in which combination(s) were given to which patients.

Summary

The small body of research was of variable quality, with Woo[100] displaying the least risk of bias, and Ramser[98] being at the highest risk. The most important source of potential bias in this literature is with respect to the conduct and interpretation of the index test. All studies fitted FeNO cut-off points to the data *post-hoc*, and are thus likely to over-estimate diagnostic accuracy. In addition, Ramser[98] did not provide sufficient information to judge whether the index test results had been interpreted blind to the reference standard. Study flow and timing was the least likely domain to contain sources of bias, in that only Sivan[99] did not provide sufficient clarity on whether all patients received the same reference standard.

b. All studies included in the review of diagnostic accuracy of FeNO in children

Study design and timeline of studies: Study characteristics and timelines are given in Tables 22 and 24. All four studies had a prospective cohort design, and with the exception of Linkosalo 2012[97] (in which the study design was not clear), and they each enrolled consecutive patients. The timing of diagnostic procedures among the studies also appeared broadly comparable. Linkosalo 2012[97] performed FeNO prior to an exercise challenge test, after which spirometric testing was used at 4, 10, and 15 minutes. Final spirometry occurred at 20 minutes, after salbutamol inhalation had been given. Ramser 2008[98] likewise performed FeNO prior to pulmonary function assessment and spirometric testing. In addition, patients who did not react to the exercise testing were provided with an additional methacholine test at 1 hour. Sivan 2009[99] also assessed FeNO first, and followed up with spirometry and sputum induction one to two hours later, though sputum induction did not contribute to the diagnosis of asthma, which was based on assessment by a certified paediatric pulmonologist after at least 18 months' follow-up and treatment. Finally, Woo 2008[100] asked all participants to fill in an ISAAC (International Study of Asthma and Allergies in Childhood) questionnaire and undergo clinical assessment. FeNO measurements were then taken, followed by spirometry and methacholine challenge test.

Population: The study populations were broadly similar in terms of their position on the diagnostic pathway, and all recruited children and adolescents, though upper and lower age cut-offs varied a little, with the most inclusive being Sivan 2009[99] at 5 to 18 years and Linkosalo 2012[97] at 6 to 19 years of age, and the least inclusive being Ramser 2008[98] at 6 to 16 years. There were some further differences in inclusion criteria. Linkosalo 2012[97] only included children and adolescents with confirmed atopy, whilst Ramser 2008[98] and Woo 2008[100] included a mix of atopic and non-atopic patients. Sivan 2009[99] did not report number of patients with atopy, but did not specifically include on this basis and the study is therefore likely to have included a mix of atopic and non-atopic patients. All studies recruited patients at Position A in the UK pathway; Linkosalo 2012[97] recruited patients who had been referred to an allergist with asthma-like symptoms (Position A); Ramser 2008[98] included children in Position A who had been referred to an outpatient clinic for diagnostic assessment of possible reactive airway disease; Woo 2008[100] included children presenting with nonspecific respiratory symptoms suggestive of asthma, and who had not been receiving controller medications for at least three months prior to FeNO testing (Position A); and Sivan 2009[99] also recruited from Position A; in this case, those with nonspecific respiratory symptoms suggestive of asthma for at least three months. This study also excluded patients with any other conditions that may have interfered with FeNO or sputum eosinophil count, especially unresolved respiratory tract infection, or underlying systemic or inflammatory disease.

Sample size ranged from 30 (Linkosalo 2012[97]) to 245 (Woo 2008[100]), and the mean age ranged from 10.26 (Linkosalo 2012[97]) to 12.4 years (Sivan 2009[99]), although mean age was not given by Ramser 2008[98]. Unlike adult studies where males were in the minority, there was a preponderance toward male participants in all four studies, with the lowest percentage being observed in the Sivan 2009[99] study (55.3%).

Interventions: Three of the four studies measured FeNO via chemiluminescence, although each used a different device: Linkosalo 2012[97] used the Sievers NOA 280; Ramser 2008[98] used the CLD 77 AM (Eco Physics, Durnten, Switzerland); and Sivan 2009[99] used the CLD 88 (EcoPhysics, Durnten, Switzerland). Woo 2008[100] was the only study to use NIOX MINO for FeNO evaluation (Aerocrine AS, Solna, Sweden). In terms of FeNO cut-off points, Linkosalo 2012[97] and Ramser 2008[98] both used the same pre-specified cut-off points of 10, 20, 30, 40, and 50 ppb. Sivan 2009[99] used cut-offs of 15, 18, 19, 25, and >20 / <15. Woo 2008[100] reported a large number of cut-off values; ranging from >50 to >5 ppb.

Reference standard: None of the studies fully replicated the UK guidelines, and one (Ramser 2008)[98] used procedures which are not employed in the UK. Linkosalo 2012[97] used an exercise challenge test (free running test) with spirometric tests before and after exercise, and after salbutamol inhalation. Sivan 2009[99] based the diagnosis of asthma on history of two or more exacerbations, evidence of airway reversibility in response to ICS or bronchodilators, or airway hyper-responsiveness at any time during a period of 18 months follow-up. Woo 2008[100] performed a battery of tests similar to those of the UK treatment pathway (spirometry, MTC, and atopy assessment), with FeNO being measured prior to these other tests. Ramser 2008[98] used spirometric testing and MTC, but also bodyplethysmography, which is not currently included in the UK guidelines.

Table 23: Diagnostic review: Study and patient characteristics of studies in children and adolescents

Author Year	Study details	Age group	Inclusion / exclusion criteria	N analysed / N recruited	Age (years) Gender	FEV_1% predicted	FeNO	Atopic
Linkosalo 2012[97]	**Setting:** Paediatric allergist Finland **Funding:** Non-industry **Design:** Prospective cohort study unclear if consecutive	Children and adolescents	Children and adolescents, 6 to 19 years Those with confirmed atopy referred to an allergist with asthma-like symptoms	30/30	**Mean age (±SD):** EIB +ve: 10.7 (range 8 to 19) EIB -ve: 9.6 (range 6 to 13) **Code for population:** Children and adolescents **Male n (%):** 20 (66.7%)	EIB +ve: 97 ± 2 EIB -ve: 96 ± 3, p = 0.723)	EIB +ve: 31.3 (SD 4.1) EIB -ve: 15.6 (SD 3.6)	30/30 (100%)
Ramser 2008[98]	**Setting:** Switzerland secondary care **Funding:** NR **Design:** Prospective, consecutive cohort study	Children and adolescents	Children 6 to 16 years of age Referred to outpatient clinic for diagnostic work on possible reactive airway disease. Short acting beta agonists must have been held on day of testing, long acting beta agonists witheld for at least 24 h before testing.	169/169	**Mean age:** NR **Code for population:** Young children **Male n (%):** 96/169 (57%)	97 +12 (atopic, n=104); 102 +13 (non-atopic, n=57)	Atopic: 35 ± 36 (n=104); Non-atopic: 13 ±16 (n=57	104/169 (61.5%)
Sivan 2009[99]	**Setting:** Israel, secondary care,	Children and adolescents	Children and adolescents Inclusion criteria: (1) nonspecific respiratory	150/156 (n=6 unable to produce	**Mean age (±SD):** Steriod naïve asthmatics: 12.6 (range 5 to 18 years,	Steriod naïve asthmatics: 79.3 ±44.4; Asthma	Steriod naïve asthmatics: 69 ±17; Asthma treated	NR

Author Year	Study details	Age group	Inclusion / exclusion criteria	N analysed / N recruited	Age (years) Gender	FEV_1% predicted	FeNO	Atopic
	outpatient clinic **Funding:** authors declared no conflict of interest **Design:** Prospective, consecutive patients		symptoms suggestive of asthma for ≥ 3 months' duration, including cough, wheezing, and shortness of breath with or without trials of treatment with bronchodilators and inhaled corticosteroids; (2) children were cooperative and successfully completed all 3 tests (3) follow-up at our clinic for at least 1 year. Exclusion criteria: patients with other conditions that could affect FeNO or sputum eosinophil count, including subjects with symptoms of unresolved respiratory tract infection, with systemic clinical manifestations of atopy such as anaphylaxis, angioedema, food allergy, urticaria, or with an underlying systemic or inflammatory disease.	sputum)	n=69); Asthma treated with ICS: 12.3 (range 6 to 18 years, n=37) Non-asthmatics: 12.0 (7 to 18years, n=44) **Code for population:** Children and adolescents **Male n (%):** Steriod naïve asthmatics: 40/69 (58%); Asthma treated with ICS: 19/37 (52%); Non-asthmatics: 24/44 (55%)	treated with ICS: 75.0 ±16.0; Non-asthmatics: 86.1 ±17.1	with ICS: 36 ±57; Non-asthmatics: 12.6 ±9	
Woo 2012[100]	**Setting:** Korea, Secondary care (out patient clinic) **Funding:**	Children and adolescents	Children and adolescents 8 to 16 years Included: children presented with nonspecific respiratory symptoms suggestive of asthma including cough, wheezing, and	245/245	**Mean age (±SD):** 11.7±2.2 Non-atopy asthmatic: 11.6±2.7 (n=38) non-asthmatic: 11.4±2.0 (n=18)	87.6±11.6	GM in asthmatic = 23.4 ppb (95% Cl, 20.9-26.2ppb); non-asthmatic = 12.6 ppb (95%	189/245 (77%)

Author Year	Study details	Age group	Inclusion / exclusion criteria	N analysed / N recruited	Age (years) Gender	FEV_1% predicted	FeNO	Atopic
	Non-industry **Design:** Prospective, consecutive cohort study		shortness of breath. All of included patients did not receive inhaled short-acting ß2-agonists in the 8 h prior to the measurements and were also not receiving a regular treatment with controller medications for 3 months or more before evaluation of FeNO and lung function.		Atopy astmatic: 11.7±2.4 (n=129) non-asthmatic 12.6±2.6 (n=60) **Code for population:** Children and adolescents **Male n (%):** Non-atopy asthmatic 20/38 (52.6%); non-astmatic 9/18 (50%); Atopy asthmatic 92/129 (71.3%); non-asthmatic 42/60 (70%)		Cl, 10.9-14.5 ppb)	
		FeNO, fractional exhaled nitric oxide; FEV_1, forced expiratory volume in first second; +ve, positive; -ve, negative; MCT, methacholine challenge test; ICS, inhaled corticosteroids; h, hours; LTFU, lost to follow-up; n, number; N total number; NR, not reported; CI, confidence interval; EIB, exercise induced bronchoconstriction; GM, geometric mean; SD, standard deviation						

Table 24: Diagnostic review: Description of interventions in studies recruiting children and adolescents

Author, year	**Population**	**Age group**	**Device**	**Cut-off values**	**Reference standard**	**Details of reference standard**	**Position of FeNO in the pathway**	**Relevance to decision problem (Relevant/not relevant)**
Linkosalo 2012[97]	Position A with confirmed atopy	Children and adolescents	Sievers NOA 280 (chemiluminescence)	10, 20, 30, 40, 50	Airway hyper-responsiveness to exercise	Exercise induced bronchoconstriction - free running test with goal of 80% maximum heart rate according to age. Spirometry 4, 10 and 15 mins after exercise and after salbutamol inhalation given 20 mins after exercise. EIB positive if maximal decrease in FEV_1 was 12% or higher.	FeNO at Position 1	Relevant
Ramser 2008[98]	Position A	Children and adolescents	CLD 77 AM, Eco Physics, Durnten, Switzerland (chemiluminescence)	10, 20, 30, 40, 50	Airway hyper-responsiveness (MCT or exercise)	Spirometry, bodyplethysmography and methacholine challenge according to ATS/ERS guidelines. EIB was defined by decrease in FEV_1 by ≥15% of baseline. MCH challenge was done using a panel of incremental dosages of MCH, and a dose of 1.8 mg was defined as threshold of PD20 to differentiate normal airway hyper responsiveness from BHR	FeNO replaces whole pathway	No relevant – uses reference standard not used in the UK
Sivan 2009[99]	Position A	Children and adolescents	Eco Physics CLD88, EcoMedics (chemiluminescence)	15 ppb, 19 pbb, 25 ppb, >20 or <15 ppb	Exacerbation history, airway reversibility, airway hyper-	Patient's history of 2 or more clinical exacerbations of wheezing documented by a physician, dyspnea, or cough	FeNO replaces whole pathway	Relevant – uses long term follow-up.

Author, year	**Population**	**Age group**	**Device**	**Cut-off values**	**Reference standard**	**Details of reference standard**	**Position of FeNO in the pathway**	**Relevance to decision problem (Relevant/not relevant)**
					responsiveness	relieved by bronchodilators, documented variability in $FEV_1 \geq 15\%$ in response to bronchodilators at any time during the follow-up period (reversibility), r documented variability in $FEV_1 \geq 15\%$ over time with or without controller medications: inhaled corticosteroids (ICS) or montelukast. Results of provocation tests were included when available. Children in whom asthma did not manifest within 18 months of follow-up were considered as not having asthma.		
Woo 2012[100]	Position A	Children and adolescents	NIOX MINO (Aerocrine AB, Solna, Sweden)	5, 10, 15, 20, 25, 30, 34, 40, 45, 50 ppb (optimum at 22 ppb)	Airway reversibility, airway hyper-responsiveness (MCH)	Relevant symptom history and reversible airflow obstruction (≥12% improvement in FEV_1 in response to inhaled β2-agonist) and/or airway hyper-responsiveness.	FeNO replaces whole pathway	Relevant

FeNO, fractional exhaled nitric oxide; FEV_1, forced expiratory volume in first second; FVC, forced vital capacity; ATS, American Thoracic Society; ERS, European Respiratory Society; MCT, methacholine challenge test; ICS, inhaled corticosteroids; BHR, bronchial hyperresponsiveness; NR, not reported; EIB, exercise induced bronchoconstriction;

Summary: The study of greatest relevance for this assessment is Woo 2012[100] which recruited patients in Position A on the pathway and used the NIOX MINO device versus a reference standard that roughly equates to UK practice.

- In this study, FeNO replaces the whole pathway prior to ICS use

Two of the remaining studies were of some relevance to the UK context:

- Sivan 2009, which used an EcoMedics device in patients at Position A in the pathway versus a reference standard similar to UK practice
    - In this study, FeNO replaces the whole pathway prior to ICS use
- Linkosalo 2012[97], which used a Sievers NOA 280 chemiluminescence device for patients in Position A on the pathway with a reference standard of exercise challenge test, which not all presenting patients would receive in UK practice
    - FeNO would be Positioned before exercise challenge test and could triage patients away from this.

Ramser 2008[98] is not of relevance to the UK context as body plethysmography is not always available in the UK. This study used an EcoPhysics device for patients in pA on the pathway with a reference standard of spirometry, body plethysmography and airway hyper-responsiveness to exercise or methacholine.

No studies in children were identified which incorporated ICS responsiveness in the reference standard.

Estimates of diagnostic accuracy: The sensitivity and specificity values for each of the studies are presented in Appendix 13.

Table 25 also displays three sets of sensitivity and specificity for each of the studies. These are:

- The highest sum of sensitivity and specificity as reported by the authors of the study
- The highest sensitivity – in this scenario a negative test result rules out a diagnosis. This was selected as the cut-off that provided the highest sensitivity. Where 100% sensitivity was reported for more than one cut-off, the cut-off that maintained the highest specificity was selected. Where the cut-off with the highest sensitivity was not also the cut-off with the highest positive predictive value (PPV), this latter cut-off was also presented.
- The highest specificity – in this scenario, a positive test result rules in a diagnosis of asthma. Selected as for the highest sensitivity, but for specificity. Where the cut-off with the highest specificity was not also the cut-off with the highest negative predictive value (NPV), this latter cut-off was also presented.

It should be noted that superior sets of sensitivity and specificity may have in fact been achieved, but selection was limited to the range of cut-off points reported within studies.

Table 25: Diagnostic review: Diagnostic accuracy of FeNO tests in children and adolescents

					Highest sum of sens and spec					Rule-out					Rule-in				
Study author, year	**Population**	**Device**	**Reference standard**	**N analysed**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**Cut-off**	**Sens**	**Spec**	**PPV**	**NPV**
Position A versus whole pathway																			
Linkosalo 2012[97]	Position A with confirmed atopy	Sievers NOA 280 (chemiluminescence)	Airway hyper-responsiveness to exercise	30	20ppb	72	83	86.7	66.7	10 ppb	89	33	66.7	66.7	30	50	92	90	55
Ramser 2008[98]	Position A	CLD 77 AM, Eco Physics, Durnten, Switzerland (chemiluminescence)	Airway hyper-responsiveness (MCT or exercise)	169	20 ppb	49	76	74	51	10 ppb 20 ppb	76 49	36 76	63 74	51 51	50 ppb	20	93	80	45
Sivan 2009[99]	Position A	Eco Physics CLD88, EcoMedics (chemiluminescence)	Exacerbation history, airway reversibility, airway hyper-responsiveness	150	19 ppb >20 or <15 ppb	86 89	89 88	92.2 93.5	79.6 82.1	15 ppb	90	70	82.7	81.6	As highest sum				
Woo 2012[100]	Position A	NIOX MINO (Aerocrine AB, Solna, Sweden)	Airway reversibility, airway hyper-responsiveness (MCT)	245	21 ppb	56.9	87.2	90.5	50.0	5 ppb	94	14.1	70.0	50	41 ppb	23.4	100.0	100	37.9

Sens, sensitivity; Spec, specificity; PPV, positive predictive values; NPV, negative predictive value; ppb, parts per billion; MCT, methacholine challenge test.

There was a high degree of agreement as to the cut-off which produces the highest sum of sensitivity and specificity, despite the heterogeneity in devices and reference standards, with values between 19 and 21pbb (Table 25). However, estimates of sensitivity at these cut-off points were not similar across studies, ranging from 49 to 86%; specificity was more similar between studies ranging from 76 to 89%. Rule-out cut-off points were not similar and varied from 5 to 20ppb, and rule-in cut-offs similarly ranged from 30 to 50ppb. For ruling out, the highest sensitivity was reported by Woo 2012[100], with a paired specificity of 14.1%, PPV 70% and NPV 50%. Sensitivities ranged from 76 to 94%. For ruling in, the highest specificity was also reported by Woo 2012[100] at 100%, and with a paired sensitivity of 23.4%, PPV 100%, NPV 37.9%. Specificities varied less than sensitivities; from 89 to 100%. It should be noted that superior rule-in and rule-out sets of sensitivity and specificity may have in fact been achieved, but selection was limited to the range of cut-off points reported within studies.

In Woo 2012[100] an optimal cut-off point of 22 ppb was selected by the authors, demonstrating 56.9% sensitivity, and 87.2% specificity. However, by our calculations, a cut-off of 21ppb provides a higher sum of sensitivity and specificity with values 56.9 and 87.9% respectively. A cut-off of 41 ppb provided 100% specificity and PPV, but low sensitivity and NPV (23.4% and 37.9%, respectively). A cut-off of 5ppb provides 94% sensitivity with 70% PPV, but only 14.1% specificity with 50% NPV.

Linkosalo 2012[97] reported an optimal FeNO cut-off of 30 ppb, with sensitivity and specificity values of 50% and 92% respectively, and PPV / NPV of 90% and 55% respectively. However, the highest sum of sensitivity and specificity is achieved at a cut-off of 20ppb with sensitivity 72% and specificity 83%. The device used in this study was the Sievers NOA 280 analyser (Sievers Instruments, Boulder, CO, USA), and, as discussed in Section 5.2.1.3, results from chemiluminescent devices cannot be assumed to generalise to NIOX MINO.

In the Sivan 2009[99] study, the optimal FeNO cut-off point was 19 ppb (sensitivity, 86%; specificity, 89%; PPV, 93.5%; NPV 82.1%). Superior diagnostic properties (sensitivity, 89%; specificity, 88%; PPV, 92%; NPV 80% are achieved by use of two cut-offs, >20 or <15ppb to rule-in and rule-out patients, with the patients falling in-between presumably being given further tests for asthma.

In Ramser 2008,[98] the optimal FeNO cut-off in terms of sensitivity and specificity was 20 ppb (49% and 76%, respectively). The PPV was 74%, and the NPV was 58%.

No meta-analysis was performed on these data due to heterogeneity in FeNO measurement devices and reference standards.

5.2.2.2.2 Studies using FeNO in conjunction with another test as the index test

One study recruiting children reported estimates of diagnostic accuracy for FeNO in conjunction with another test. The study (Sivan 2009[99]) is described in more detail in Section 5.2.2.2.1, but in summary recruited children at Position A in the pathway, and used FeNO in conjunction with sputum eosinophilia against a reference standard of evidence of airway reversibility in response to ICS or bronchodilators, or airway hyper-responsiveness at any time during 18 months follow-up. As with Fortuna 2007[76] (a study in adults) sputum eosinophilia is not a test in widespread use in the UK and is therefore of limited generalisability to UK practice. Results showed that improvements in diagnostic accuracy were very small; sensitivity increased from 86% to 87% and specificity remained the same at 89%.

5.2.2.3 Studies providing data on subgroups of interest to the review

5.2.2.3.1 Adult smokers

Malinovschi 2012,[101] a diagnostic cohort study, was identified which specifically investigated the effects of smoking on the usefulness of FeNO in diagnosing asthma. This study was not included in the main diagnostic review as the method of recruitment was unusual: patients were recruited from a population sample who responded to a letter with a validated, self-administered asthma screening questionnaire, and reported two or more symptoms of asthma and subsequently accepted an invitation to join the study. As such, patients were unlikely to represent the usual spectrum of patients presenting to a GP practice. In addition, the reference standard is problematic in that it used "asthma medication use" as a positive test; this would allow patients who have been previously misdiagnosed as having asthma to be considered asthmatic. Whilst both of these differences make the study heterogeneous in comparison to the other diagnostic studies included in the review, it is the best level of evidence identified for assessing the use of FeNO in smokers and non-smokers.

Study design and setting: The study was a prospective consecutive cohort of patients from a random sample of the population, conducted in Denmark.

Population: The study did not recruit people on the basis of presentation to a GP, but rather on the basis of the presence of two or more symptoms of asthma as reported on a mailed questionnaire, regardless of smoking status. Patients with spirometric results suggestive of COPD were excluded. Patients were adults and adolescents aged 14 to 44 years, and 31.9% of patients had Immunoglobulin E (IgE) sensitisation. Out of a total of 282 participants, 108 (38%) had never smoked, 62 (22%) were ex smokers and 112 (40%) were current smokers. Mean age was 32.7 years and 40% were male. FEV_1% was 94% for asthmatics, and 97% for non-asthmatics. There were 282 participants in total; 108 had never smoked, 62 were ex-smokers and 112 were current smokers.

Intervention: the study used NIOX MINO to measure FeNO values.

Reference standard: to be diagnosed with asthma, patients had to exhibit symptoms and test positive by one of: MCT; airway reversibility to bronchodilator; daily use of steroids or SABA; or asthma symptoms during but not outside the pollen season, supported by allergic rhinitis. As the cohort was recruited from a random sample of the population with asthma symptoms, and did not exclude existing asthmatics, the reference standard in part depends on a previous diagnosis of asthma in that a patient already prescribed steroids or SABA is automatically classed as asthmatic. In practice, this reference standard may therefore include patients who have been wrongly diagnosed in primary care.

Estimates of diagnostic accuracy: From Table 26 it can be seen that the highest sum of sensitivity and specificity results in a cut-off of 15ppb when only never smokers are analysed. The sensitivity in this group is higher than the specificity, at 77.8% compared to a specificity of 63.5%. For current smokers, the cut-off moves to 17ppb, and sensitivity and specificity also change. For this group, a higher sensitivity is achieved than specificity; a change in sensitivity from 77.8% in never smokers to 56.3% in current smokers, and a change in specificity from 63.5% in never smokers to 82.5% in current smokers. Ex smokers produce a cut-off of 22ppb, and similar to smokers, sensitivity is lower than specificity, at 62.2% and 86.1% respectively.

When considering the best cut-offs for ruling out and ruling in, these appear fairly similar across the subgroups; the best rule-in cut-off is 50 ppb in all groups, and the best rule-out cut-off for smokers, when considering the cut-off with the highest sensitivity rather than the highest negative predictive value, is 3ppb lower than in other groups at 7ppb compared to 10ppb. As a rule-out test, FeNO achieves the highest sensitivity in ex-smokers at 95%, and is very similar in both smokers and never smokers (90.6 and 91% respectively). rule-outWhen considering the results for the cohort as a whole, sensitivity is low, and specificity high compared to the subgroups, but negative predictive value is comparatively good at 83.8%. For rule-in scenarios, the picture is similar, and it is unclear as to how the fairly minor differences in cut-off points and diagnostic properties of FeNO across groups would affect cost-effectiveness and clinical utility in practice.

Table 26: Diagnostic review: Diagnostic accuracy in adult and adolescent smokers, non-smokers, ex-smokers and studies recruiting all ages

					Highest sum of sens and spec					Rule-out					Rule-in				
Study author, year	**Device**	**Reference standard**	**Population**	**N analysed**	**cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**cut-off**	**Sens**	**Spec**	**PPV**	**NPV**
Malinovschi 2012[101]	NIOX MINO	Symptoms, plus one of airway reversibility, airway hyper-responsiveness (MCT), prescribed steroids of SABA, symptoms in pollen season plus allergic rhinitis.	All	282	20 ppb	52.08	82.8	61	77	10 ppb	87.5	33.3	40.4	83.8	50ppb	15.6	96.8	71.4	69.0
			Smokers	112	17 ppb	56.3	82.5	56.3	82.5	7 ppb 10ppb	90.6 78	15.0 48	29.9 37.3	80.0 84.4	50ppb 35ppb	9 16	98 96	60 62.5	72.9 74.0
			Never smokers	108	15ppb	77.8	63.5	60.3	80.0	10ppb	91	27	47.1	81.0	50ppb	20	95	75	62.5
			Ex Smokers	62	22ppb	62.2	86.1	66.7	84.1	10ppb	95	16	33.3	87.5	35ppb	37	98	87.5	77.8

Sens, sensitivity; Spec, specificity; PPV, positive predictive values; NPV, negative predictive value; MCT, methacholine challenge test; ppb, parts per billion; SABA, short-acting β2-agonist.

5.2.2.3.2 – Children exposed to tobacco smoke

We were unable to identify any studies which evaluated the diagnostic accuracy of FeNO in children exposed to tobacco smoke. However, evidence from the review on the use of FeNO measurements in the management of asthmatic children exposed to tobacco smoke may provide some insight (albeit limited) on how environmental tobacco smoke may impact on mean FeNO values and therefore FeNO cut-off points. Mahut[102] and Hanson[103] both reported that FeNO levels were not statistically significantly different between those exposed or not exposed to tobacco smoke. Whilst De La Riva-Velasco [104] reported that FeNO values were lower in ICS treated children who were exposed to tobacco smoke (see Section 5.2.3.3.4). Similarly, evidence from a diagnostic cohort study (Malinovschi,[101] Section 5.2.2.3.1) which investigated the effects of smoking on the usefulness of FeNO in diagnosing asthma in adults and adolescents (rather than children) reported that FeNO could differentiate asthmatic subjects from non-asthmatic subjects with asthma-like symptoms equally well in both never- and current smokers. However, the FeNO cut-off levels were lower in current and ex smokers.

The findings from the above studies suggests that it may be necessary to consider a child's exposure status when interpreting results of FeNO for the diagnosis of asthma, as FeNO may be lower in children exposed to tobacco smoke.

5.2.2.3.2 – Pregnant women

Although no studies were identified which evaluated the diagnostic accuracy of FeNO in pregnant women, one cross-sectional study by Tamasi[105], conducted in Hungary, compared FeNO levels in pregnant asthmatic and healthy women. A total of 102 females were recruited from an outpatient clinic of which 35 were healthy non pregnant women, 27 healthy pregnant women, 20 asthmatic non pregnant women and 20 asthmatic pregnant women. Exclusion criteria included the following: current smokers or had more than 5 pack years of smoking history, other chronic diseases (e.g. chronic rhinitis, hypertension), acute infection within 3 weeks of measurement or body mass index >30 kg/m^2. Asthma was diagnosed using the GINA guidelines and all asthmatic patients had persistent disease. All asthmatic patients were receiving ICS. In addition, 14 patients were on long acting β-agonist and 7 patients received added leukotriene- receptor agonist therapy. The mean age ranged from 27 years in the non-pregnant healthy women to 31 years in non-pregnant asthmatic women. FeNO was measured using the NIOX MINO device.

The authors found no significant difference in median FeNO levels between healthy pregnant (16 [IQR 9 to 35] ppb) and healthy non pregnant subjects (16 [IQR 8 to 31] ppb). Similarly, no significant difference was observed in the level of asthma control between pregnant and non-pregnant asthmatic and there was no significant difference in the total asthma control test scores (20.78±2.96

vs. 19.17±3.1, respectively, p=0.17). In contrast, FeNO levels in pregnant asthmatic women was significantly higher than healthy pregnant women (28[10 to 56] ppb versus 16 [9 to 35] ppb p<0.05). Similarly, the FeNO levels in non-pregnant asthmatic women was significantly higher compared with non-pregnant healthy women (38[9 to 54] ppb versus 16[8 to 31] ppb respectively, p<0.0001). In addition the authors reported that there was no significant difference between the two groups of asthmatics, the mean FeNO values (estimated from graph in published paper) for pregnant asthmatic women was 29 ppb and for non-pregnant asthmatics 32 ppb.

Overall, the study authors concluded that pregnancy itself does not alter FeNO levels either in healthy or in asthmatic patients and FeNO levels of pregnant asthmatic patients correlate with asthma control levels.

5.2.2.3.3 – The elderly

No diagnostic studies which used FeNO to diagnose asthma in the elderly were identified; however, one study (Simpson[106]) which examined FeNO levels and eosinophilic airway inflammation in elderly subjects with airflow obstruction was idenitifed. In asthma diagnosis, FeNO's main use is to identify patients with eosinophilic airway inflammation who are likely to respond to ICS, as a surrogate for other methods of ascertaining eosinophilic inflammation such as sputum counts. As such, this study should provide some evidence as to whether FeNO still acts as a surrogate marker for eosinophilic inflammation in the elderly. This observational case-control study was conducted in Australia and was reported in abstract form only and thus provided limited data. The study recruited 65 elderly patients with or without fixed airflow obstruction with 32 healthy controls. The setting from which the patients were recruited was unclear and the majority of patients (86%) with air flow obstruction were on ICS treatment.

The authors found that participants with eosinophilic airway inflammation (sputum eosinophil count (Eos) >3%) had similar FeNO levels to those with non-eosinophilic inflammation (16.1 (10.0-29.2) ppb vs.19.1 (13.2-24.3) ppb, respectively, p=0.762). Those with a diagnosis of asthma had similar FeNO levels to those with COPD. There was no correlation between FeNO and sputum eosinophils or any clinical markers. The authors concluded that FeNO was not a surrogate marker of eosinophilic airway inflammation in older people and showed no relationship with clinical outcomes.

5.2.3 Management review

This section is broken down into a number of subsections by population age and subgroup. Briefly these are:

- 5.2.3.1 FeNO-guided management in adults
    a. Quality assessment

b. Study details

c. Estimates of efficacy

- 5.2.3.2 FeNO-guided management in children
  - a. Quality assessment
  - b. Study details
  - c. Estimates of efficacy
- 5.2.3.3 FeNO-guided management in subgroups defined in the scope
  - 5.2.3.3.1 Pregnant women
  - 5.2.3.3.2 The elderly
  - 5.2.3.3.3 Adult smokers
  - 5.2.3.3.4 Children exposed to tobacco smoke

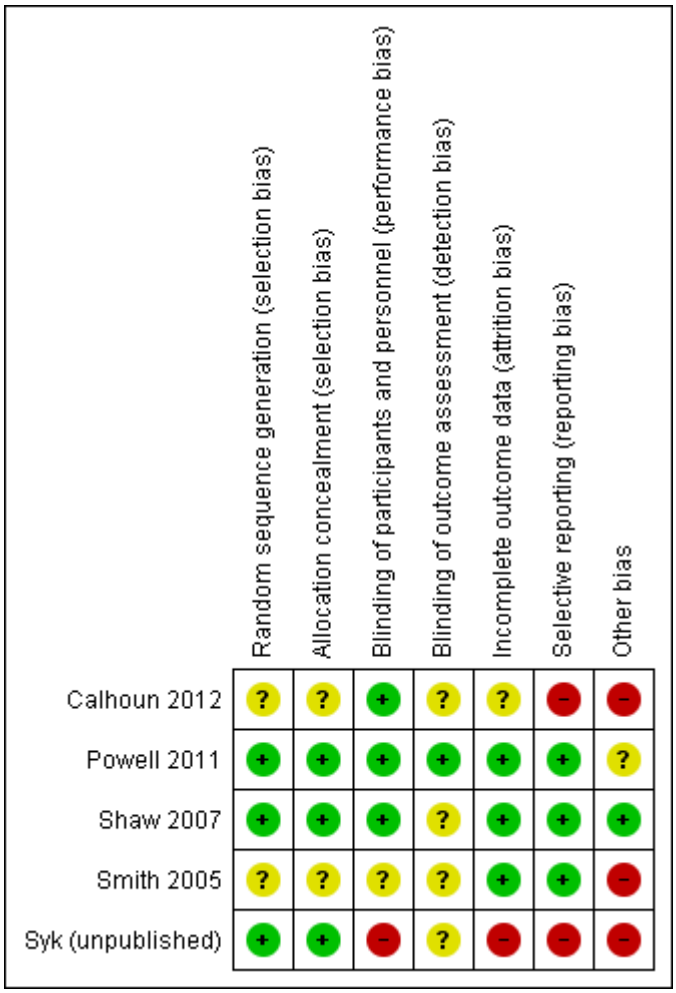
5.2.3.1 Adults

Four studies which recruited adults and compared FeNO-guided management to non-FeNO-guided management were included in the review.[107-110] Shaw 2007 was based in the UK,[108] Smith 2005[107] in New Zealand the unpublished study by Syk 2013 in Sweden[109] and Calhoun 2012[110] in the USA. Smith 2005[107], Syk[109] and Calhoun 2012[110] were at least partly supported by Aerocrine, and the Syk[109] study was submitted as part of Aerocrine's sponsor's submission. An additional study by Powell 2011[111] was conducted in adult pregnant women and is discussed separately in Section 5.2.3.3.1 as this group was defined *a priori* as a distinct group.

a. Quality assessment

The quality of the five adult's management studies (the Powell 2011[111] study in pregnant women was included in the quality assessment of adult studies) was assessed according to criteria proposed in the Cochrane Handbook and CRD Handbook. Powell 2011,[111] and Shaw 2007[108] appeared to be the highest quality articles, with each containing only one potential source of bias (industry sponsorship and uncertain outcome assessor blinding, respectively). The study at highest risk of bias was the unpublished study by Syk;[109] this was due to the lack of participant / personnel blinding, incomplete outcome data, and selective reporting.

Figure 13: Methodological quality summary: review authors' judgements about each methodological quality item presented as percentages across all included studies



Green circles with + sign, low risk of bias; red circles with – sign, high risk of bias; yellow circles with ?, unclear risk of bias.

Risk of selection bias

All the included studies were described as randomised, and three of the five studies provided satisfactory information on both random sequence generation and allocation concealment. In Shaw,[108] allocation was performed by an independent individual; Syk[109] drew lots from sealed envelopes; and Powell 2011[111] allocated in computer generated blocks of four, stratified as above or below 800µg/day budesonide. There may have been adequate randomisation procedures in the remaining two studies,[107,110] but this could not be confirmed on the basis of the reports.

Risk of performance bias

The study at highest risk from lack of blinding was the unpublished Syk study,[109] which was described as open-label. Smith[107] was rated as 'unclear' on this item, since the study was single-blind (participants only). As many of the outcomes were patient-reported, patient blinding may have been

the most important source of bias to avoid, though the blinding of other study personnel who were deciding whether to step patients up or down may also have been important. The remaining three studies,[108,110,111] were all double-blind, and therefore at low risk of performance bias.

Risk of detection bias

Powell 2011[111] was the only study to clearly state that outcome assessment blinding had been performed. The poor reporting of outcome assessment blinding in the other studies means that unblinded outcome assessment may be a potentially important source of bias throughout this body of literature. However, as outcome assessment blinding often goes unreported in journal articles, it was unclear whether any potential bias was due to reporting practices, or methodological shortcomings in the conduct of the studies themselves.

Risk of attrition bias

Powell,[111] Shaw,[108] and Smith[107] all appeared to be at low risk of attrition bias. Dropout rates from these studies were low, adequately reported, and corrected for in the statistical analyses. In addition, Smith 2005[107] performed analysis by intention to treat (ITT) and extrapolated missing data. There was a potentially high risk of bias in the Syk 2013[109] study, in that patients were missing and not corrected for in multiple analyses. There were two possible sources of bias in the Calhoun 2012[110] study: it was unclear how missing data was corrected, and there were more drop-outs in the intervention arm. If these patients were dropping out because of unsatisfactory outcomes (which was not clear from the report), this could skew the results in favour of FeNO.

Risk of reporting bias

Three of the five studies appeared to have provided data on all the prespecified outcomes.[111,107,108] However, there was some evidence of selective reporting in Calhoun[110] and Syk.[109] Calhoun 2012[110] failed to report oral prednisone levels, although this had been specified as an outcome in the study protocol, and Syk 2013[109] did not report number of severe exacerbations. Syk 2013[109] also used medians rather than means in several of the outcomes, precluding these from meta-analysis. However, these data were supplied by the manufacturer (Aerocrine) upon request.

Risk of other bias

There were a number of further potential sources of bias in each of the studies. Smith[107] reported receipt of commercial sponsorship, while Syk 2013[109], Calhoun 2012[110] and Powell 2011[111] reported at least partial commercial funding. Three studies (Smith 2005,[107] Syk 2013[109] and Calhoun 2012)[110] also conducted a run-in period prior to randomisation. It is unclear whether this may have introduced bias to the results in Smith 2005[107] and Syk 2013[109] whilst in Calhoun 2012[110] patients were excluded if their asthma did not remain controlled when administered two puffs b.i.d. of beclomethasone HFA

(40ug/puff). This is likely to have influenced the spectrum of patients recruited to this trial towards those with less severe asthma. However, this is likely to affect external validity rather than internal validity as both arms are subject to the same run-in period. No further sources of bias were identified in the remaining studies.

Summary

The quality of the sampled literature was variable, with Powell 2011[111] and Shaw 2007[108] being at the lowest risk of bias. Indeed, where there were potential sources of bias in these studies, they were unlikely to have affected the results in a substantial way; only two authors in Powell 2011[111] reported receipt of lecture fees and meeting / travel expenses from commercial entities, and no other authors reported any conflicts of interest. The only potential source of bias we identified in the Shaw 2007[108] article pertained to their failure to explicitly state blinding of outcome assessors. However, it was unclear whether this was an actual methodological flaw, or merely inadequately reported, and at least some of the outcomes were patient reported (patients were blinded). Among the remaining literature, the most important potential source of bias was selective reporting in the Calhoun 2012[110] and Syk 2013[109] studies, both of which failed to report some prespecified outcomes. The study at highest overall risk of bias was the open-label Syk 2013[109] investigation. In addition to lack of blinding, and the aforementioned selective reporting, this study may have been subject to attrition bias. In the absence of information on why data were missing from this study, it is difficult to ascertain how likely this may have biased the results, and in what direction.

b. Study details

Study design and timeline of studies: Table 27 provides details of study design and the timelines of the studies. All four were RCT studies. Smith 2005[107] and Shaw 2007[108] were both single blind, whilst Syk[109] was open label. Calhoun 2012[110] was described as "multiply blinded," though it is not entirely clear who was blinded. Table 27 details the timelines of the studies; no two studies followed the same timeline exactly. Smith 2005,[107] Syk[109] and Calhoun 2012[110] had a run-in period pre-randomisation where LABA was reduced or withdrawn and/or doses of ICS were standardised. Post-randomisation, all studies had an initial period of time where visits were more frequent. In Calhoun 2012[110], visits were made every two weeks for the first six weeks post-randomisation, then six-weekly after that. In Smith 2005[107] initial visits were monthly for four months (Shaw 2007)[108], then every two months up to 12 months (Shaw 2007).[108] In the study by Smith 2005,[107] treatment comprised two phases: a optimisation phase of three to twelve months, and a titration phase of a further twelve months. Patients were randomised before both phases, and both phases managed patients according to protocols which either did or did not incorporate FeNO measurements. However, data on eacerbations was only reported for the titration phase, and it was this data that was incorporated into the analysis. Syk[109] had an initial visit two to four weeks after the initial titration visit, then every two months up to

four months, then every four months up to 12 months. All studies titrated doses for at least 12 months except Calhoun 2012[110] where doses were titrated for nine months only. Calhoun 2012[110] included a third intervention arm that was not relevant to this review where ICS dose was controlled by matching ICS use on a puff by puff basis to the rescue use of albuterol in response to occurrence of symptoms.

Population: Table 28 provides details of patient characteristics across studies. All studies were of a moderate size, with numbers analysed ranging from 94 (Smith 2005)[107] to 229 (Calhoun 2012).[110] All patients were recruited from primary care, except for Calhoun 2012[110] where it was not clear whether patients were recruited from primary or secondary care settings. All had either a doctor's diagnosis of asthma, or asthma diagnosed according to guidelines. In Calhoun 2012[110], the doctor's diagnosis was confirmed with either a positive MCT or demonstration of airway reversibility. Inclusion and exclusion criteria varied across studies, but where compatible data is reported, study populations seem broadly similar in terms age (mean ranged from around 34.5[110] to 45[107] years), FEV_1% (mean ranged from 81.4% to 87.7%) and FeNO values (range of geometric means 18.88 ppb to 29.0 ppb). It is difficult to determine the comparability of study populations in terms of severity at baseline as different scales for severity and different metrics for medication use have been used. Inclusion and exclusion criteria suggest that at least three studies[107,110] recruited populations with mild to moderate asthma; Smith 2005[107] excluded those with ≥ four severe exacerbations in previous 12 months and those ever admitted to intensive care for asthma, whilst Syk[109] and Calhoun 2012[110] stated that all patients were mild to moderate asthmatics. Smith 2005[107] and Syk[109] also required patients to have been receiving ICS treatment for more than six months, and Calhoun 2012[110] only recruited patients who were well controlled when prescribed two puffs b.i.d. of beclomethasone HFA (40ug/puff) and were ≥75% compliant with medication during the two week run in period. Shaw 2007[108] on the other hand, may have recruited a wider spectrum of severity. Patients were only required to have had one prescription for asthma medication in the previous 12 months, making it possible that patients with comparatively less severe or less well documented asthma were included, but only excluded those with severe exacerbations in the previous 4 weeks, making it possible that severe asthmatics were included.

Smith 2005[107] included smokers (current or ex) with a history of <10 pack years whilst Shaw 2007,[108] Syk[109] and Calhoun 2012[110] all excluded current smokers, but included ex-smokers with a past smoking history of <10 pack-years. Smith 2005[107], Shaw 2007[108] and Calhoun 2012[110] all included a mix of atopic and non-atopic patients whilst Syk[109] included only atopic patients. It is unclear whether studies in atopic patients will over or underestimate efficacy, or have no impact at all, though clinical input to the assessment suggested that it would be expected to increase estimates of efficacy as atopy is correlated with ICS responsiveness.

Overall, patient populations recruited by Smith 2005[107] and Shaw 2007[108] are likely to be more representative of the general asthma population in the UK as they have included atopic and non-atopic patients, and Smith 2005[107] has also included a potentially broader spectrum of severity and some smokers. Calhoun 2012[110] has also recruited a mix of atopic and non-atopic asthmatics, but the run-in requirements for treatment tolerance and compliance may mean that generalisation to a wider population is difficult. However, were the application of FeNO management to be limited in the UK to certain populations (e.g. only atopic patients, only stable patients, only mild to moderate patients), data from Calhoun 2012[110] or Syk[109] may be more appropriate.

Interventions: Table 29 provides details of the interventions used in each study. It is not possible to determine whether any studies used the same devices as this information is not clearly reported in three studies.[107,108,110] Syk[109] used NIOX MINO. Smith 2005[107] used an usual flow rate, but justified their conversion to 35ppb equivalent at 50ml/sec. None of the studies used the same protocol or cut-off points for management of asthma with FeNO. Syk[109] and Calhoun 2012[110] used FeNO only to guide management, Smith 2005[107] used FeNO only, with a safety measure based on symptoms, bronchodilator use and spirometry and Shaw 2007[108] used FeNO in addition to the Juniper score, which gauges control through symptoms. Doses and medications used also varied from study to study, with Smith 2005[107] and Calhoun 2012[110] only titrating ICS, Shaw 2007[108] titrating ICS, LTRA and bronchodilators and Syk[109] titrating ICS and LTRA. The number of cut-offs also varied. Smith 2005[107] used only one cut-off of 35ppb (equivalent). Shaw 2007[108] and Calhoun 2012[110] each used two cut-offs, but at different cut points; one for titrating down (<16ppb and <22ppb respectively) and one for titrating up (>26ppb and >35ppb respectively) with an intermediate area in-between where symptoms also guided treatment (Shaw 2007[108]) or dose remained the same (Calhoun 2012[110]). Syk[109] used three cut-offs, with different values for men and women (<19ppb, ≥24ppb and ≥30ppb for men, <21ppb, ≥25ppb and ≥32ppb for women). Given the uncertain comparability in FeNO measurements between devices, it is difficult to assess how similar these cut-off points may in fact be.

Control: Table 30 provides details of the control interventions used in each study. As with the interventions, none of the studies used the same criteria, protocols or treatment doses for the management of asthma in the control arm of the study. Generally speaking, the control arms considered symptoms, self-reported medication use and sometimes lung function to guide titration. In terms of similarity to UK practice, Shaw 2007[108] states that BTS/SIGN guidelines were followed, using the Juniper scale to score symptoms. It is not clear how similar to UK practice other studies may be.

Table 27: Adult management review: Study design and timelines

Author Year	Study design	Timeline of study
Smith 2005[107]	RCT: single blind, single centre, placebo-controlled	Visit 1: enrolment, start of 2-week run-in period where LABA withdrawn, reinstated at fixed dose if not tolerated Visit 2 (week 2): FeNO and spirometry. Patients begin 4 weeks of 750µg/day fluticasone or 500µg/day if previous dose < 200µg/day Visit 3 (week 6): **Randomisation** and start of phase 1. **Titration** phase 1: (3 to 12 months after randomisation) visits every 4 weeks, FeNO and spirometry, dose adjustment to optimal dose by downward titration until FeNO ≥15ppb (equiv to 35ppb at 50ml flow rate) or uncontrolled, then uptitrated until controlled/ ≤15ppb. This dose deemed "optimal dose". **Titration** phase 2: (12 months after completion of phase 1) visits every 2 months: upwards adjustments when control lost/FeNO >15ppb, downwards adjustment if controlled/FeNO≤15ppb for two consecutive visits, but not below optimal dose. Treatment orders assigned by blinded investigator. Compliance assessed by inhaler weight.
Shaw 2007[108]	RCT: single blind, parallel group	**Titration** at each visit Visit 0: **Randomisation**; FeNO, FEV_1, FVC, PC20, induced sputum analysis, skin prick test, Juniper score Visit 1: Two weeks after visit 0. FEV_1, FeNO, Juniper score. Visit 2 to 5 - monthly visits to 4 months. FEV_1, FeNO, Juniper score. Visit 6 - at 6 months - FEV_1, FeNO, Juniper score, PC20, Sputum. Visit 7, 8: at 8 and 10 months - FEV_1, FeNO, Juniper score. Visit 9: 12 months - FEV_1, FeNO, Juniper score, PC20, Sputum.
Syk unpublished[109]	RCT: open label, parallel group, multicentre	Visit 1: eligibility and consent. Capillary blood for IgE confirmation. LABA withdrawn, ICS continued (Salbutamol inhaler with dose counter) Visit 2 **(Titration):** 2 to 4 weeks later. FeNO, Spirometry, reversibility, Juniper mini-AQLQ, generic QoL, Juniper 6 item ACQ, and a questionnaire on allergen exposure. Venous blood for IgE analysis. ICS and LTRA altered according to a) FeNO levels and six fixed treatment steps in FeNO group and b) FeNO recorded (blind) and treatment adjusted according to usual care (patient report, SABA use, physical exam, pulmonary function tests) Visit 3 **(Titration)**: (2 months) ACQ, FeNO and treatment altered visit 4 **(Titration)**: (4 months) mAQLQ, ACQ FeNO and treatment altered visit 5 **(Titration)**: (8 months) as visit 3. Visit 6**(Titration)**: (12 months) identical to visit 2.

Author Year	Study design	Timeline of study
		Outcomes recorded at visits 2-6
Calhoun 2012[110]	RCT: mulitply-blinded, multi-centre study	Visit 1: (Week 0) consent and start of run-in period of 2 weeks: 2 puffs b.i.d. of beclomethasone HFA (40ug/puff). If asthma acceptably controlled at this level, enrolled in trials. Visit 2&3: (weeks 2 to 8) Pre-randomisation period. Patients given 2 pairs of inhalers to facilitate blinding - one with beclomethasone (2 x 40ug b.i.d.) and a placebo counterpart, one with albuterol and a placebo counterpart (taken together on demand). Visit 4: (week 8) **Randomisation** to group 1 or group 2. Vistis 5 to 12 **(Titration)**: 2, 4, 6, 12, 18, 24, 30 and 36 weeks post-randomisation. Dose adjustments made at time of clinic visits, monitoring of secondary outcomes

Table 28: Adult management review: Study and population characteristics

Author, year	Study details	Inclusion/exclusion criteria	n analysed/N recruited	Age (years (SD)) Gender	Spirometry Mean (SD)	Severity	FeNO	Smokers; Atopic	Medication use
Smith 2005[107]	**Setting** New Zealand, primary care **Funding** Mix[a]; equipment from aerocrine	Chronic asthma (Am Rev Respir Dis 1987; 136:225) managed in primary care; regular inhaled corticosteroids for ≥ six months, no dose change in previous six weeks. If could not tolerate removal of LABA during run-in allowed to participate if could tolerate a fixed dose. **Exclusions:** ≥ four courses oral prednisone in previous 12 months; admission to hospital for asthma in previous six months; ever admitted to IC for asthma; smokers (current or ex) with history of > 10 pack years.	94/110 WBR: 13 **I:** 46/48 **C:** 48/49	**Adolescents and adults** (12 to 75) **Mean age** 44.8 (range 12 to 73) **Male** 41/110 (37.3%)	**FEV_1% mean (95%CI)** **I:** 86.4 (80.6 to 92.2) **C:** 83.1 (76.5 to 89.7)	**Symptom score**[b] **mean (95% CI)** **I:** 0.6 (0.4 to 0.8) **C:** 0.8(0.6 to 1.1)	**FeNO 250ml** [c] **GM(95% CI)** **I:** 7.8 (6.6 to 9.3) **C:** 6.4 (5.5 to 7.5)	**Smokers** NR **Atopic** NR	**Bronchodilator use, mean per day previous 7 days (95% CI)** **I:** 0.5 (0.2 to 0.8) **C:** 0.6 (0.3 to 0.8) **ICS** NR
Shaw 2007[108]	**Setting** UK, recruited from primary care **Funding** Asthma UK grant.	Patients with GP diagnosis of asthma who received ≥1 prescription for antiasthma medication in last 12 months. Current non-smokers with a past smoking history of less than 10 pack-years. **Exclusions:** Those	118(ITT LOCF) /119 WBR: 1 **I:** 58 **C:** 60	**Adults** >18 years **Mean age** NR **Male** 54/118 (46%)	**FEV_1%** **I:** 81.4 (20.9) **C:** 84.9 (20.1) **FEV_1/FVC** **I:** 71 (10.7) **C:** 72 (9.9)	**Juniper score mean (SD)** **I:** 1.32 (0.65) **C:** 1.26 (0.75)	**log FeNO GM (68% CI)** **I:** 29.2 (14.0 to 61.0) **C:** 31.2 (13.3 to	**Ex-smokers** I: 22% C: 25% **Atopic** 78/118 (66.1%)	**Mean daily dose ICS (SD)** **I:** 697µg (708) **C:** 652µg (533)

Author, year	Study details	Inclusion/exclusion criteria	n analysed/N recruited	Age (years (SD)) Gender	Spirometry Mean (SD)	Severity	FeNO	Smokers; Atopic	Medication use
	Speakers fees but not from Aerocrine	poorly compliant; those with severe asthma exacerbation (needing prednisolone) in previous 4 weeks					73.1)		
Syk unpublished[109]	**Setting** Sweden, primary care **Funding** Mix[a]; some from aerocrine	Doctor's diagnosis of asthma and ICS treatment for ≥ six months, IgE sensitisation to at least one major airborne perennial allergen (dog, cat or mite). Non-smokers for ≥1 year and with smoking history of <10 pack-years. Patients all had mild to moderate asthma.	165/187 WBR: 6 **I:** 87/93 **C:** 78/88	**Adults** (18 to 64 years) **Mean age** 41 (12.4) **Male** 94/181 (51.9.0%)	**FEV_1%** **I:** 84.3 (14.1) **C:** 83.7 (12.5) **FEV_1/FVC** **I:** 0.78 (0.08) **C:** 0.79(0.08)	NR	**FeNO ppb GM (95% CI)** **I:** 22.0 (19.3 to 25.2) **C:** 21.6 (18.7 to 25.0)	**Smokers** 0/165 (0%) **Atopic** 165/165 (100%)	**Median Budesonide equivalent ICS dose (µg/day)** 400 (IQR 400 to 800) **LABA before study entry** 54/180 (30.0%)
Calhoun 2012[110]	**Setting** USA, care level , NR **Funding** Mix[a]; equipment from Aerocrine	Patients with mild to moderate, well controlled persistent asthma with compliance rates ≥75%, who could tolerate treatment of 2 puffs b.i.d. of beclomethasone HFA (40ug/puff) during two week run in period	363 recruited to trial WBR: 21 **I:** 115/115[d] **C:** 114/114[e]	**Adults** (assumed from mean age) **Mean age:** I: 34.8 (11.3); C: 34.2 (11.9)	**FEV_1%** **I:** 86.3% (10.4) **C:** 87.7% (12.1)	**ACQ** **I:** 0.79 (0.54) **C:** 0.72 (0.50) **AQLQ** **I:** 6.16 (0.77) **C:** 6.27 (0.76)	**FeNO ppb GM** **I:** 18.88 (0.66) **C:** 21.38 (0.62)	**Smokers** NR **Atopic** 196/229 (85.6%)	**Albuterol rescue use median (IQR)** **I:** 0.07 (0 to 0.43) **C:** 0.04 (0 to 0.29)

Author, year	Study details	Inclusion/exclusion criteria	n analysed/N recruited	Age (years (SD)) Gender	Spirometry Mean (SD)	Severity	FeNO	Smokers; Atopic	Medication use
			Other study arm (not included in review): 113/113	**Male** 75/229 (32.8%)		**ASUI** **I:** 0.88 (0.12) **C:** 0.90 (0.10)			

ASUI, asthma symptom utiity index; IC, intensive care; n, number; N, total number; SD, standard deviation; FeNO, fractional exhaled nitric oxide; LABA, long acting β2 agonist; WBR, withdrew before randomisation; I, intervention group; C, control group; CI, confidence interval; GM, geometric mean; NR, not reported; ICS, inhaled corticosteroids; ITT, intention to treat; LOCF, last observation carried forward; ppb, parts per billion; b.i.d. twice per day; ACQ, asthma control questionnaire; AQLQ, asthma quality of life questionnaire;

[a]Mix of industry and non-industry funding eg. Research council grants
[b] Daily score, previous 7 days: Asthma symptoms were scored for each 24-hour period as follows: 0 indicated no symptoms, 1 symptoms for one short period, 2 symptoms for two or more short periods, 3 symptoms most of the time that did not affect normal daily activities,4 symptoms most of the time that did affect normal daily activities, and 5 symptoms so severe as to disrupt daily activities.
[c] FeNO measured at 250ml/s gives lower values than FeNO at 50ml/s
[d]37 withdrew, imputation method NR
[e] 13 withdrew, imputation method NR

Table 29: Adult management review: Description of the intervention management strategies

Author, year	Decisions based on Flow rate, device Cut-off points	Step-up/step-down protocol	Doses
Smith 2005[107]	**Based on** FeNO, with a safety measure based on symptoms, bronchodilator use and spirometry **Flow rate, device** 250ml/sec, ATS 1999, Assume Niox device[a] **Cut-offs** Equivalent to 35ppb 50ml/second (≥15ppb at 250ml/second)[b]	FeNO <35ppb (equivalent at 50ml/s) = asthma controlled FeNO ≥35ppb = asthma uncontrolled **Safety measure:** If one or more of the following clinical criteria met, increase one step: 1. symptom score for previous 7 days ≥one point more than mean run-in AND minimum score of 2/5 2. nocturnal wakening on ≥3 nights/week more than mean run-in 3. mean daily bronchodilator use ≥3 times that of mean run-in AND minimum use 15 occasions during prior 7 days 4. diurnal peak flow variation ≥30% AND/OR FEV_1 of <85% of baseline	Dose steps: placebo, inhaled fluticasone 100µg, 250µg, 500µg, 750µg, 1000µg Phase 1: until optimal dose reached. Optimal dose = one step higher than that at which control lost. Phase 2: up titrate one step at a time; down titrate if controlled for two visits, but not lower than optimal dose. Patients had personalised self-management plan which instructed them to take oral prednisone 40mg per day, when morning peak flows fell below 70% of mean run-in values, until it reached >85%, at which time they took 20mg per day for the same number of days.
Shaw 2007[108]	**Based on** FeNO plus symptoms (Juniper score) **Flow rate, device** 50ml/sec, device NR **Cut-offs** Intermediate: 16 ppb to 26ppb High: >26ppb.	Exhaled NO < 16 ppb on first occasion or exhaled NO 16-26 ppb on second occasion: AND Juniper score ≤1.57 = step down anti-inflammatory treatment, step down bronchodilator treatment once off steroids. AND Juniper score >1.57 = Step down anti-inflammatory treatment, Step up bronchodilator treatment Exhaled NO >26ppb: AND Juniper score ≤1.57 = step up anti-inflammatory treatment, no change in bronchodilator treatment AND Juniper score >1.57 = Step up anti-	Hierarchy of Anti-Inflammatory Treatment: 1) Low dose inhaled steroid (100-200µg BDP b.i.d) 2) Moderate dose inhaled steroid (200-800µg BDP b.i.d) 3) High dose inhaled steroid (800-2000µg BDP b.i.d) 4) High dose inhaled steroid (800-2000µg BDP b.i.d) plus leukotriene antagonist 5) Higher dose inhaled steroid (2000µg BDP b.i.d) plus leukotriene antagonist 6) Higher dose inhaled steroid (2000µg BDP b.i.d) plus leukotriene antagonist plus oral Prednisolone 30mg 2/52, then titrating dose reducing by 5mg/week Hierarchy of Bronchodilator Treatment

Author, year	Decisions based on Flow rate, device Cut-off points	Step-up/step-down protocol	Doses
		inflammatory treatment, Step up bronchodilator treatment once on maximum anti-inflammatory treatment **Safety measure:** patients on 2000µg becolmethasone per day with >26ppb FeNO and had not fallen to 60% of baseline had sputum checked. If no eosinophilic inflammation, treatment reduced stepwise, unless FeNO increased by >60% of baseline.	1) PRN short-acting β2-agonists 2) Long acting β2-agonists 3) Long acting β2-agonists plus theophylline 4) Long acting β2-agonists plus theophylline plus nebulised bronchodilator
Syk unpublished[109]	**Based on** FeNO only **Flow rate, device** ATS 2005, NIOX MINO **Cut-offs** <19ppb (men), <21ppb (women) 19-23 (men), 21-25 (women) ≥24ppb (men), ≥26ppb (women) ≥30ppb (men), ≥32ppb (women)	FeNO <19ppb (men), <21ppb (women) - decrease one step FeNO 19-23 (men), 21-25 (women) - no change FeNO ≥24ppb (men), ≥26ppb (women) - increase one step (no change in treatment step if on step 4 or 5 and using ≤2 inhalations of short-acting β2-agonist per week.) FeNO ≥30ppb (men), ≥32ppb (women) - increase two steps (only if on treatment step 1). Grey zone of 5ppb applied to avoid frequent dose changes.	Steps 1 - 6: Budesonide (µg/day): 0, 200, 400, 800, 800+LTRA, 1600+LTRA Fluticasone ((µg/day): 0, 100, 250, 500, 500+LTRA, 1000+LTRA Mometasone (µg/day): 0, 100, 200, 400, 400+LTRA, 800+LTRA

Author, year	Decisions based on Flow rate, device Cut-off points	Step-up/step-down protocol	Doses
Calhoun 2012[110]	**Based on** FeNO only **Flow rate, device** Flow rate NR, device NR (protocol states Niox) **Cut-offs** Well controlled: <22ppb Controlled: 22 ppb to 35ppb Under-controlled: >35ppb.	<22ppb = well controlled = down one level 22 to 35 = controlled = maintain current level >35ppb = undercontrolled = up 1 level	Dosing Beclamethasone HFA: Level 1 = 0ug/day Level 2 = 80ug qd Level 3 = 160ug b.i.d. Level 4 = 320ug b.i.d. Level 5 = 640ug b.i.d.

FeNO, fractional exhaled nitric oxide; ppb, parts per billion; ATS, American Thoracic Society; FEV, forced expiratory volume in first second; NR, not reported; LTRA. Leukotriene receptor antagonist; qd, once per day; b.i.d, twice per day; BDP, beclomethasone dipropionate; PRN, not defined in source article.

[a] Donated by Aerocrine
[b] Discussed and supported in journal article[107]

Table 30: Adult management review: Description of the control group management strategies

Author, year	Decisions based on	Step-up/step-down protocol	Doses
Smith 2005[107]	GINA 2002; symptoms, bronchodilator use, spriometry	GINA 2002 Uncontrolled asthma criteria: 1. Symptoms present >2 days/wk with 24 hour asthma score ≥2/5 2. >1 nighttime waking/wk 3. Bronchodilator use >4 occasions/week or on >2 days per week 4. Variation in PEFR >20(aplitude % of mean over previous 7 days) 5. FEV_1 <90% of baseline	As for intervention, but without the personalised management plan
Shaw 2007[108]	BTS/SIGN guidelines using Juniper scale to score symptoms	Scored by Juniper scale. Treatment doubled if score >1.57, Treatment halved if score <1.57 for 2 consecutive months	Step 1: SABA as required Step 2: Add inhaled steroid 200 to 800mcg/day BDP equivalent Step 3: Add inhaled LABA Step 4: increase ICS up to 2000mcg/day and addition of 4th drug, eg LTRA, theophylline, LABA Step 5: oral prednisolone, high dose ICS, refer to specialist care.
Syk unpublished[109]	Symptoms, lung function, b-agonist use.	Usual care (patient symptom report, SABA use, physical exam, pulmonary function tests).	Assume same doses as intervention
Calhoun 2012[110]	NHLBI guidelines (USA version of SIGN guidelines)	Use severity classification chart, assessing both domains of impairment and risk, to determine initial treatment. Use asthma control chart, assessing both domains of impairment and risk, to determine if therapy should be maintained or adjusted (step up if necessary, step down if possible). Use	As intervention.

Author, year	Decisions based on	Step-up/step-down protocol	Doses
		multiple measures of impairment and risk: different measures assess different manifestations of asthma; they may not correlate with each other; and they may respond differently to therapy. Obtain lung function measures by spirometry at least every 1–2 years, more frequently for not-well-controlled asthma. Asthma is highly variable over time, and periodic monitoring is essential. In general, consider scheduling patients at 2- to 6-week intervals while gaining control; at 1–6 month intervals, depending on step of care required or duration of control, to monitor if sufficient control is maintained; at 3-month intervals if a step down in therapy is anticipated. Assess asthma control, medication technique, written asthma action plan, patient adherence and concerns at every visit	

GINA, Global Initiative for Asthma; wk, week; PEFR, peak expiratory flow rate; FEV_1, forced expiratory volume in the first second; BDP, beclomethasone dipropionate; BTS, British Thoracic Society; SIGN, Scottish Intercollegiate Guidelines Network; SABA, short acting ß2 agonists; LABA, long acting ß2 agonists; LTRA, leukotriene receptor antagonist; NHLBI, National Heart, Lung and Blood Institute; ICS, inhaled corticosteroid.

c. Estimates of efficacy

Exacerbations

Exacerbations were reported in all studies, but definitions varied (Table 31) and results were not entirely consistent across studies.

Major or severe exacerbations: This outcome was defined differently across studies. Smith 2005[107] reported two such outcomes: "major exacerbations" defined according to global daily asthma scores; and exacerbations leading to a course of oral prednisone (OCS). A similar outcome "worsening requiring a course of OCS" was also reported in Syk 2013.[109] Shaw 2007[108] did not report rates of OCS use alone, but did report a composite outcome of "exacerbations resulting in the use of OCS or antibiotics". Calhoun 2012[110] reported an outcome called "exacerbations" which included exacerbations leading to OCS use, increased ICS use or additional medication for asthma. This latter definition may incorporate exacerbations that other studies would have classified as moderate or minor, though the study does define an additional outcome called "treatment failure" which is likely to incorporate minor, moderate and major exacerbations. As such, the outcome "exacerbations" in Calhoun 2012[110] will be considered in this analysis.

Smith 2005[107] and Shaw 2007[108] both reported lower rates of major/severe exacerbations in the intervention arms, but in neither case did the level reach statistical significance in comparison to the control arms. Syk 2013[109] reported higher levels of OCS use in the intervention arm, but the level of significance was not reported. Calhoun 2012[110] showed very similar rates of exacerbations in both arms of the trial with no statistically significant difference between them. The best improvement in major/severe exacerbations was seen in Shaw 2007,[108] at -21% (95% CI -57 to 43%, p=0.43) (reviewer-calculated rate ratio 0.79 (95% CI 0.44 to 1.41) and the worst in Syk 2013[109] which reported higher rates, though not statistically significantly so, in the intervention arm at 0.113 versus 0.0875 (p value not reported) (reviewer-calculated rate ratio 1.29 (95% CI 0.51 to 3.30)).

Despite the high level of heterogeneity between study characteristics, an exploratory meta-analysis of the rates of major/severe exacerbations using fixed effects methods was conducted and showed no heterogeneity, with an I^2 statistic of 0% (Figure 14). The pooled estimate was 0.87 (95% CI 0.64 to 1.19), with a p value 0.38. This indicated a trend towards fewer major exacerbations in the intervention arm, but this did not reach statistical significance.

In sensitivity analysis, Calhoun 2012[110] and Shaw 2007[108] were both removed, leaving only studies which reported the number of exacerbations resulting in the use of OCS (Figure 15). In this analysis, heterogeneity statistics remained low (I^2=0%), and the pooled estimate continued to show a statistically non-significant trend toward a positive effect of FeNO for asthma management, with a

risk ratio of 0.97 (95% CI 0.61 to 1.54) and a p value of 0.90. The two studies included in this analysis showed opposite directions of effect, which may be due to the different step-up/step-down protocols employed in the studies, or due to the populations being slightly different.

Figure 14: Fixed effects meta-analysis of the effects of FeNO guided asthma management on major/severe exacerbation rates

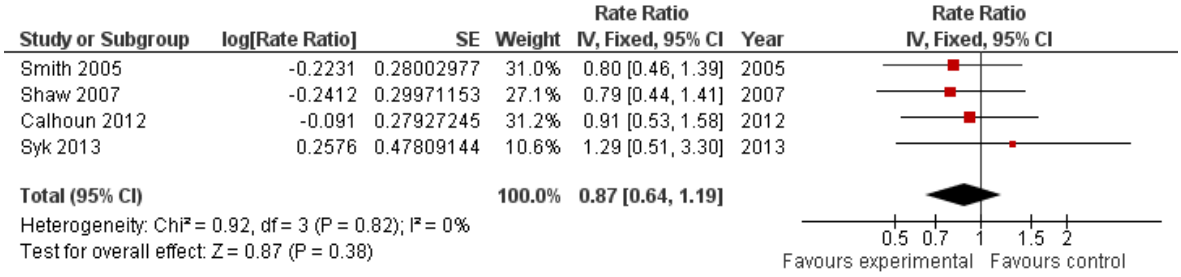
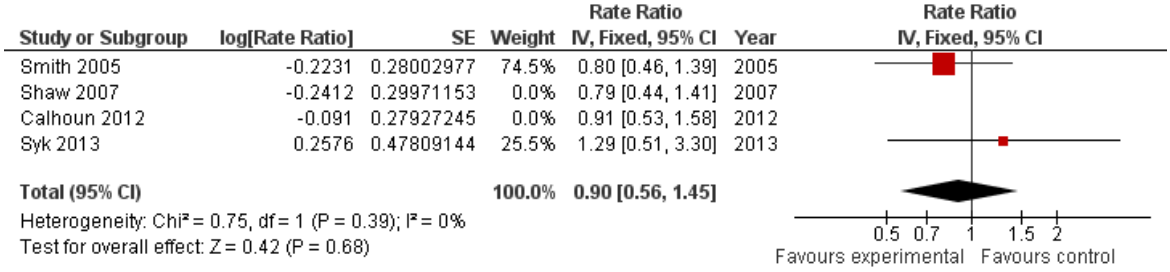


Figure 15: Sensitivity analysis removing studies with wider definitions of major/severe exacerbations



All exacerbations or treatment failures: When considering other, wider definitions of exacerbation, as described in Table 31, three studies report composite outcomes which can be considered broadly similar, and represent what may be termed "treatment failure". In Smith 2005[107] and Syk 2013[109] this is "any major or minor exacerbation", whilst is Calhoun 2012[110] it is exacerbation or any loss of control by a variety of measures (see footnotes to Table 31 for details). In Smith 2005[107] and Calhoun 2012,[110] FeNO-guided groups showed numerically but not statistically significantly lower rates of treatment failure. In Syk 2013[109] the improvement was statistically significant, with a rate of 0.22 in the intervention arm versus 0.41 in the control arm (p=0.024) (reviewer calculated rate ratio 0.52 (95% CI 0.30 to 0.91).

Despite the high level of heterogeneity between study characteristics, an exploratory meta-analysis of these rates (Figure 16) using fixed effects (the I^2 statistic was 0%) was conducted. The pooled risk ratio was 0.58 (95% CI 0.43 to 0.77), which represents a statistically significant effect in favour of using FeNO guided management in asthmatics for this outcome.

Moderate and minor exacerbations: Smith 2005[107] and Syk unpublished[109] both reported the rates of less severe exacerbations separately from the rates of all exacerbations and from the rates of major/severe exacerbations (Table 31). In both cases, the point estimate reduction in minor/moderate exacerbations was far greater than the reduction in severe/major exacerbations. Smith 2005[107] reported 0.36 events per person year in the intervention arm, and 0.75 events per person year in the control arm, with a p value of 0.24. Syk 2013[109] reported 0.1 moderate exacerbations per person year in the intervention arm and 0.325 in the control arm. The p value was not reported. When considering the results reported by Calhoun 2012[110] for exacerbations alone and the composite outcome treatment failure, it can be seen that the larger difference in rates of treatment failures in favour of the intervention arm is not driven by the exacerbation rates which are very similar 0.21 (97.5% CI 0.1 to 0.32) and 0.23 (97.5%CI 0.1 to 0.37), and must therefore be due to a decrease in the intervention arm of less severe exacerbations/loss of control. The impact on QoL and costs of such exacerbations is much lower than for major/severe exacerbations.

Figure 16: Meta-analysis of the effects of FeNO guided asthma management on the composite outcome of major/severe, moderate and minor exacerbation rates/treatment failures

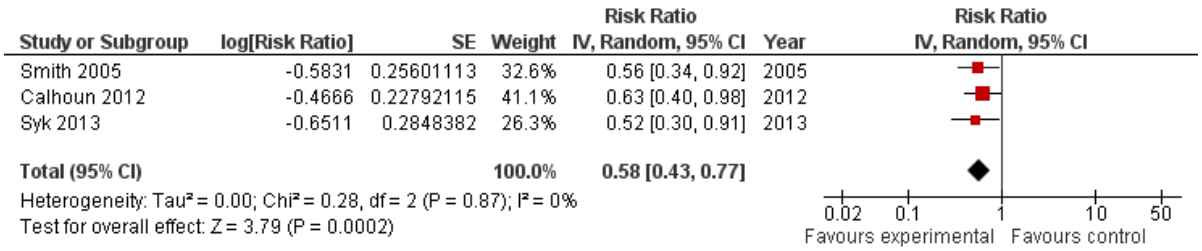


Table 31: Adult management review: Exacerbation and oral corticosterone use rates in adult patients with or without FeNO guided management

Author, year **Time of outcome**	**Definition of outcomes**	**N**	**Intervention** **per person year**	**Control** **per person year**	**Between group comparison**
Smith 2005[107] 3 to 12 months optimisation (exacerbation rates not reported for this period) plus 12 months titration	Minor: global daily asthma score[a] of 2 on ≥2 consecutive days.	94	Minor[b]: 0.36	Minor [b]: 0.75	Minor: p=0.24
	Major: global daily asthma score[a] of 3 on ≥2 consecutive days (or in one day, in the context of a minor exacerbation). Major exacerbation or medical emergency: global daily asthma score[a] of 4 in one day.		Major [b]: 0.13	Major [b]: 0.14	Major: p = 0.91
	Any minor or major exacerbation		0.49 (95% CI 0.20 to 0.78)	0.90 (95% CI 0.31 to 1.49)	-45.6% (95% CI -78.6 to 54.5, p = 0.27)
	Course of oral prednisone		0.48	0.60	p=0.60
Shaw 2007[108] **12 months**	Course of oral steroids or antibiotics	118	0.33 (SD 0.69)	0.42 (SD 0.79)	-21% (95% CI -57 to 43%, p=0.43)
Syk unpublished[109] Endpoints analysed from visit 2 to visit 6 (2 - 4 weeks, 12 months)	Moderate exacerbation[c] - need to step up controller treatment for at least two days with or without clinic visit. Prophylactic use before pollen season excluded.	165	0.1	0.325	NR
	Severe exacerbation[c] - worsening requiring a course of oral corticosteroids.		0.113	0.0875	Not significant

Author, year Time of outcome	Definition of outcomes	N	Intervention per person year	Control per person year	Between group comparison
	Moderate or severe exacerbation		0.22	0.41	Total: p = 0.024
Calhoun 2012[110]	Exacerbation: Unscheduled medical contact for increased asthma symptoms that results in the use of oral corticosteroids, increased inhaled corticosteroids or additional medication for asthma.	229	0.21 (97.5% CI 0.1 to 0.32)	0.23 (97.5%CI 0.1 to 0.37)	"did not differ"
	Treatment failure defined as exacerbation or loss of control[d]		0.27 (97.5%CI 0.14 to 0.39)	0.43 (97.5%CI 0.23 to 0.64)	"were not different"

CI, confidence interval; SD, standard deviation; NR, not reported; PEFR, peak expiratory flow rate; PEF, peak expirator flow; FEV_1, forced expiratory flow rate in first second.

[a] Asthma scores:
0 (stable): Morning PEFR >75% of best PEFR in 14-day run-in period without deterioration in any symptom scores †
1 (mildly unstable): One or more of the following: a. Bronchodilator use on 2 or more occasions in 24 hr more than rounded mean number of occasions during the run-in period; b. Increase in symptom score of 1 point or more as compared with rounded mean during run-in period; c. Onset of or increase in nocturnal waking: 1 or more times in previous 7 nights more than rounded mean no. of times during the run-in period OR Morning PEFR of 61–75% without deterioration in any of the above categories
2 (minor deterioration): Morning PEFR of 61–75% of best PEFR during run-in period and one or more criteria for an asthma score of 1 OR Morning PEFR of 41–60% without deterioration in any criteria for an asthma score of 1
3 (major deterioration) Morning PEFR of 41–60% of best PEFR during run-in period and one or more criteria for an asthma score of 1
4 (major exacerbation or medical emergency) Morning PEFR of 40% or less than best PEFR during run-in period regardless of symptoms OR Attendance at clinician's office or emergency department because of severe asthma
[b] Estimated off graph
[c] ATS/ERS Task Force Criteria 2009
[d] **At-home measurements:** any of the following 3 criteria, when not associated with the increased asthma symptoms, satisfies treatment failure criteria: a) Prebronchodilator AM peak expiratory flow (PEF) of less than 65% of baseline on 2 consecutive mornings, scheduled measurements; b) Postbronchodilator PEF of less than 80% of baseline despite 60 minutes of rescue ß-agonist treatment; c) Postbronchodilator PEF may be taken at any time of day. An increase in albuterol use of more than 8 puffs per 24 hours over baseline use for a period of 48 hours, or more than 16 puffs per 24 hours for more than 48 hours. **In-clinic measurements:** a) Prebronchodilator forced expiratory volume in the first second of expiration (FEV_1) values on 2 consecutive sets of spirometric determinations measured 24 to 72 hours apart that are less than 80% of the baseline prebronchodilator value (baseline value for adherence period: FEV_1 value at visit 3; baseline for randomization period: FEV_1 value at visit 4). All participants found to have an FEV_1 of less than 80% of baseline at any center visit but who are not considered to meet treatment failure or exacerbation criteria must be seen again within 72 hours to have FEV_1 measured; OR b)Physician judgment for patient safety; OR c)Patient dissatisfaction with asthma control achieved by study regimen; OR d) Requirement for open-label inhaled corticosteroids or another (nonsystemic corticosteroid) new asthma medication (eg, montelukast) without the addition of systemic corticosteroids.

ICS use

All studies provided some data on ICS use, and this is presented in Table 32. Smith 2005[107] and Shaw 2007[108] reported ICS use as a mean per day at the end of the study with mean differences of -270µg per day (95% CI -112 to -430, p=0.003) and -338µg per day (95% CI -640 to -37µg, p= 0.028) respectively, in favour of FeNO-guided management. Syk 2013[109] reported a median value, with the means being supplied upon request, and showed a small increase in ICS use in the intervention arm (586µg (SE 454) versus 540 µg (SE 317) in the control arm. Calhoun 2012[110] reported mean per month, though it is unclear if this was an average over the whole course of the study, or the means for the final month of the study. The means were very similar at 1617µg/month and 1610µg/month. It should also be noted that this study managed and followed patients for only nine months where the other studies did so for 12 months.

When looking at mean use over time (graphical data not reproduced here) in Smith 2005[107] and Syk,[109] ICS use fell initially in the FeNO arm (both when compared to baseline and in comparison to the control arm) and then rose at the final measurement, to a level above the control arm in Syk,[109] but staying below the control arm in Smith 2005.[107] Conversely, in the Shaw 2007[108] study, initial ICS use rose, then fell at the final two measurement points to below baseline and below the control arm. Only Shaw 2007[108] reported the area under the curve for ICS use, and this showed an 11% greater use of ICS in the FeNO group. Based on the "mean use over time" Figures this is unlikely to be true for Syk 2013[109] where a visual interpretation of the area under the curve would suggest very similar levels of total ICS use in both arms, with little change over time. Appropriate data were not available for Calhoun 2012[110] or Smith 2005.[107] These differences may be due to the different titration protocols and cut-off values used in the studies, and it is difficult to draw a generalised conclusion as to the direction of effect and the trends over time for ICS use. However, it would seem most likely that ICS use will either remain the same or fall in FeNO-managed groups when taken as an average over the course of the first year. The first year of titration is likely to be when the greatest gains are made as patients reach a stable dose. It is unclear how ICS use will change in the following years as no study reported results beyond 1 year follow-up, as the severity of the disease may progress, stay stable or remiss over time.

Despite the high level of heterogeneity between study characteristics, an exploratory meta-analysis of ICS use incorporating data from all four studies was conducted (Figure 17). As studies reported values for different ICSs (fluticasone, beclomethasone, budesonide) a standardised mean difference analysis was performed. A random effects model was used to account for the very high degree of heterogeneity between studies, but the I^2 statistic remained high at 75%. Standard deviations for Calhoun 2012[110] were imputed based on consideration of the other three studies. Sensitivity analyses where the imputed SD was altered by an order of magnitude in either direction, and where values of

10000 for the intervention arm and 5000 for the control arm were used (to mirror the SDs of Syk 2013[109]) did not alter the analysis results, with the pooled analysis confidence intervals crossing the line of no effect in every case, and the pooled mean value ranging from -0.25 to -0.23 SMD. The results of the meta-analysis agree with the conclusions drawn from the narrative consideration of the data; it would seem most likely that ICS use will either remain the same or fall in FeNO-managed groups, probably depending on factors such as step-up/step-down protocols, cut-off values selected, treatments incorporated in the treatment protocol and comparator interventions.

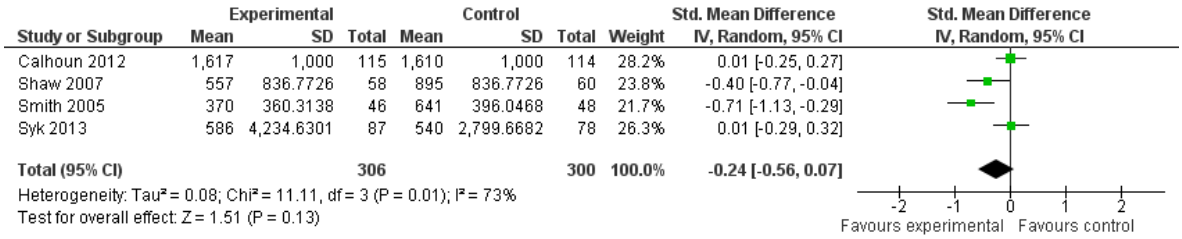
Table 32: Adult management review: ICS use

Author, year		Intervention	Control	Between group difference Expressed as Intervention minus Control (negative values indicate FeNO lower)
Smith 2005[107]	Final value ICS use[a]	Baseline: Mean 411µg per day (95%CI 344 to 478) End of phase 2: Mean 370µg per day (95% CI 263 to 477)	Baseline: Mean 491µg per day (95% CI 403 to 579) End of phase 2: Mean 641µg per day (95% CI 526 to 756)	Mean difference -270µg per day (95% CI -112 to -430, p=0.003)
Shaw 2007[108]	Final value ICS use[b]	557µg	895µg	Mean difference -338µg per day (95% CI -640 to -37µg, p= 0.028)
	Total used in study (area under the curve):			11% greater in FeNO group (-15 to 37%)
Syk unpublished[109]	ICS use[c]	Median 0 (IQR -400 to 400) Baseline: Mean 604 (SE 370) Final value: 586 (SE 454)	0 (IQR -200 to 200) Baseline: Mean 626 (SE 391) Final value: 540 (SE 317)	0.945
Calhoun 2012[110]	ICS use (unclear if mean over whole study or final value)[b]	Mean 1617µg/month	Mean 1610µg/month	NR

NS, non-significant difference; FeNO, fractional exhaled nitric oxide; ICS, inhaled corticosteroid; CI, confidence interval; IQR, interquartile range; SE, standard error; NR, not reported.

[a]fluticasone or the equivalent
[b]Beclomethasone diproprionate or equivalent
[c] Budesonide equivalent

Figure 17: Meta-analysis of the effects of FeNO-guided asthma management on mean ICS use (standardised mean difference analysis)



HRQoL

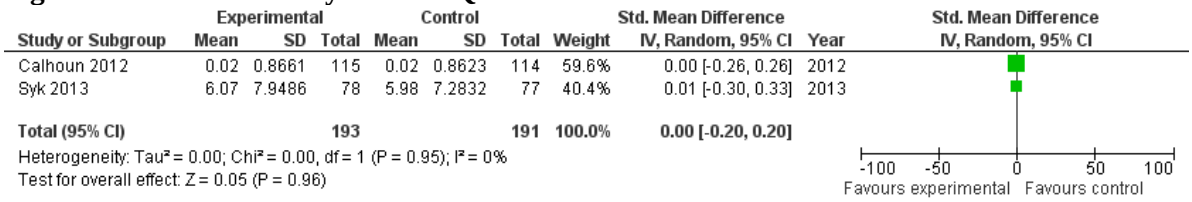
Only Syk 2013[109] and Calhoun 2012[110] reported quality of life data and this was as measured by the mini Asthma Quality of Life Questionnaire (mAQLQ) and the Asthma Quality of Life Questionnaire (AQLQ) respectively. In both studies the overall score, and in Syk 2013[109] three of four domains, did not show a statistically significant change over time. The symptoms domain did, however, show a relatively small but statistically significant between group difference in change from baseline of 0.10 (Table 33) in Syk 2013.[109] An exploratory meta-analysis of the overall scores (shown in Figure 18) shows no effect with a standardised mean difference of 0.00 (-0.20 to 0.20).

Table 33: Adult management review: HRQoL

Author, year	Intervention	Control	Between group difference
Syk unpublished[109]	Appears to be some data missing (n=78-86). Total change over time (n = 80/87): Median 0.23 (IQR 0.07 to 0.73) Final Mean value 6.07 (SE 0.90) Visit 2 and visit 6 data, median (IQR) mAQLQ symptoms: Visit 2: 5.60 (4.80 to 6.20) Visit 6: 6.00 (5.60 to 6.60) Activity limitation Visit 2: 6.50 (5.75 to 6.75) Visit 6: 6.75 (6.00 to 7.00) Emotional Function Visit 2: 6.00 (4.67 to 6.67) Visit 6: 6.33 (5.67 to 7.00) Environmental stimuli Visit 2: 6.00 (5.00 to 6.67) Visit 6: 6.33 (5.67 to 6.67) GQLI change (n = 85/88): 0.06 (-0.22 to 0.28)	Appears to be some data missing (n=77-85). Total change over time (77/78): Median 0.07 (IQR -0.20 to 0.80) Final Mean Value 5.98 (SE 0.83) Visit 2 and visit 6 data, median (IQR) mAQLQ symptoms: Visit 2: 5.70 (4.80 to 6.40) Visit 6: 6.00 (5.20 to 6.40) Activity limitation Visit 2: 6.25 (5.50 to 7.00) Visit 6: 6.50 (5.75 to 7.00) Emotional Function Visit 2: 6.00 (4.67 to 6.67) Visit 6: 6.00 (5.33 to 6.67) Environmental stimuli Visit 2: 5.67 (5.00 to 6.67) Visit 6: 6.33 (5.33 to 6.67) GQLI change (n= 78/78): 0 (-0.39 to 0.39)	Analyses of Median (IQR) change between visit 2 and visit 6 mAQLA overall: p= 0.197 (NSD) mAQLQ symptoms: p = 0.041 Activity limitation: p= 0.544 Emotional Function: 0.596 Environmental stimuli: 0.193 GQLI: p = 0.666
Calhoun 2012[110]	AQLQ change from baseline 0.02 (-0.14 to 0.18), p=0.75	AQLQ change from baseline 0.02 (97.5% CI -0.14 to 0.17), p=0.80	AQLQ between group 0.00 (97.5% CI -0.22 to 0.23) p=0.96

NR, not reported; IQR, interquartile range; SE, standard error; mAQLQ, mini-Asthma Quality of Life Questionnaire; AQLQ, Asthma Quality of Life Questionnaire; GQLI, Gothenburg quality of life instrument.

Figure 18: Meta-analysis of HRQoL outcomes



Asthma control and other medication use

All four studies reported data for asthma control. Smith 2005[107], Calhoun 2012[110] and Shaw 2007[108] reported no change in asthma control (Table 34), whilst Syk[109] reports a statistically significant difference in change in ACQ from visit 2 to visit 6 between the two trial arms. This matches the change seen in the AQLQ symptoms domain previously mentioned. Smith 2005[107], Calhoun 2012[110] and Syk[108,109] reported use of other medications; Smith 2005[107] and Calhoun 2012[110] reported no significant difference between groups for bronchodilator use, though in Calhoun there was a trend towards less use in the intervention arm, and Syk[109] reported non-significant trends towards greater numbers and mean use of LTRA and SABA (significance not reported) in the FeNO controlled arm.

Adverse events, mortality, compliance and test failure rates

No data were reported in three studies for adverse events or mortality, whilst Calhoun 2012[110] reported one unrelated adverse event (hip surgery) in the control arm. Compliance was reported by Smith 2005[107] and was 85% and 89% in the intervention and control arms respectively, and in Calhoun 2012[110] the median was ≥95% in both groups. No test failure rates for NIOX MINO or NObreath were reported.

Table 34: Adult management review: Other outcomes

Author, year	Outcome	Intervention	Control	Between group difference
Asthma control				
Smith 2005[107]	Symptom score (daily score previous 7 days): final scores, mean (95% CI)	0.4 (0.1 to 0.7	0.6 (0.4 to 0.9)	p=0.23
	Nocturnal waking (nights/week, previous 7 days): final scores, mean (95% CI)	0.2 (0.0 to 0.6)	0.2 (0.0 to 0.4)	p=0.89
	Asthma score (% of days)	Score 0: 85.2 (78.4 to 92.0) Score 1: 14.0 (7.4 to 20.6) Score ≥2: 0.8 (0.3 to 1.3)	0: 78.5 (70.4 to 86.6) 1: 19.9 (12.3 to 27.5) ≥2: 1.7 (0.3 to 3.1)	between-group final scores, (not change from baseline)p=0.19
Shaw 2007[108]	Asthma control	Data NR No difference between groups in Juniper score throughout study. However, in both groups the score decreases from baseline. Significance NR		
Syk unpublished[109]	ACQ Median (IQR), change between visit 2 and 6	-0.17 (-0.67 to 0.17) (n = 81/88)	0 (-0.33 to 0.50) (n = 74/78)	p= 0.045
Calhoun 2012[110]	Nighttime symptoms	0.01 (-0.00 to 0.02), p=0.07	0.01 (-0.00 to 0.02), p=0.11	0.00 (-0.02 to 0.02), p=0.86
	Daytime symptoms	-0.00 (-0.02 to 0.02), p=0.86	0.01 (-0.00 to 0.03), p=0.06	-0.01 (-0.04 to 0.01), p=0.17
	ACQ	-0.01 (-0.15 to 0.12), p=0.81	0.03 (-0.10 to 0.16), p=0.64	-0.04 (-0.23 to 0.15), p=0.62
	ASUI	0.01 (-0.02 to 0.04) p=0.40	0.01 (-0.02 to 0.03) p=0.64	0.00 (-0.04 to 0.04) p=0.79

Author, year	Outcome	Intervention	Control	Between group difference
Other Medication use				
Smith 2005[107]	Bronchodilator use (occasions/day, previous 7 days):	0.4 (0.1 to 0.7)	0.4 (0.1 to 0.6)	p=0.98
	Safety buffer criteria used	16/436 assessments.	NA	NA
Shaw 2007[108]	Medication use	NR	NR	NR
Syk unpublished[109]	LTRA - number of patients:	33/92 (35.9%)	19/85 (22.4%)	0.069
	Mean months on LTRA	2.87 (4.42)	1.81 (3.89)	0.094
	SABA use between visit 5 to 6 (8 to 12 months	1.56 (0.06 to 5.18)	0.94 (0.03 to 2.81	NR
Calhoun 2012[110]	Albuteral rescue use (puffs/day)	-0.04(-0.10 to 0.02), p=0.15	0.02 (-0.03 to 0.08), p=0.30	-0.06(-0.14 to 0.02), p = 0.08

CI, confidence interval; NR, not reported; ACQ, asthma control questionnaire; ASUI, asthma symptom utility index ; NA, not applicable; NR, not reported; LTRA, leukotriene receptor antagonist; SABA, short acting β2 agonist.

5.2.3.2 Children

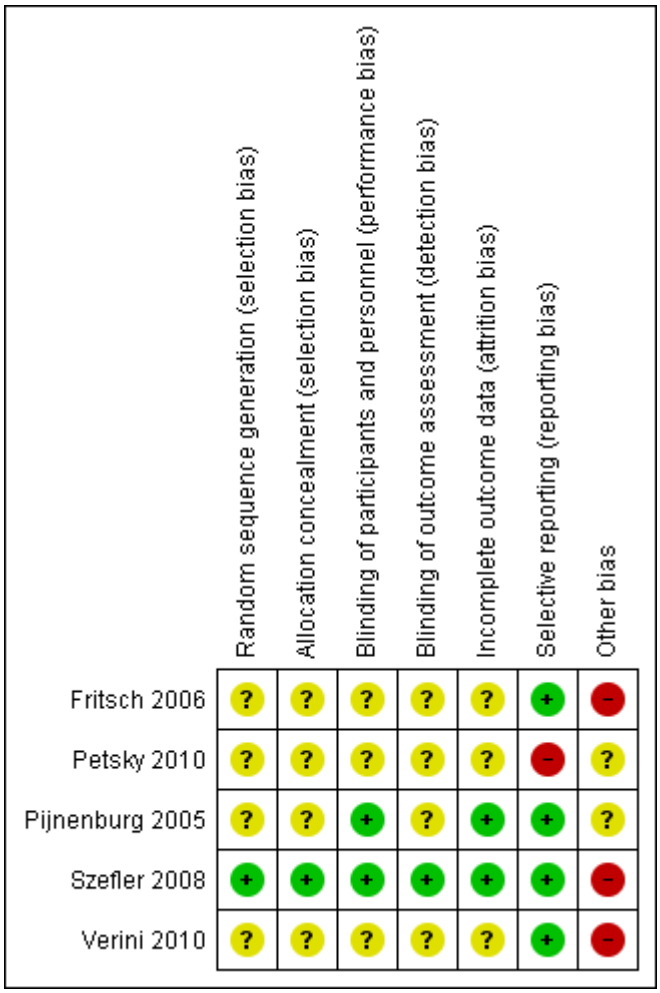
Children

Five studies which recruited children (plus adolescents and/or young adults) and compared FeNO-guided management to non-FeNO-guided management were identified. Fritsch 2006,[112] was based in Vienna, Austria; Szefler 2008,[113] was based in the US; Verini, 2010,[114] was based in Italy; Pijnenburg 2005,[115] was based in Rotterdam, Netherlands; and Petsky 2010,[116] was based in Australia. Fritsch 2006[112] received technical and analytical support from Aerocrine; one of the authors in Szefler 2008[113] received speaker's fees from Aerocrine; Pijninsburg was supported by the Kroger Foundation/Sophia Children's hospital Foundation, though Aerocrine had provided a grant to the department. Petsky 2010[116] was funded by the Royal Children's Hospital Foundation, Asthma Foundation of Queensland, while Verini 2010[114] did not report their source of funding.

a. Quality assessment

The quality of the five children's management studies, Fritsch 2006,[112] Szefler 2008,[113] Verini 2010,[114] Pijnenburg 2005[115] and Petsky 2010[116] was assessed according to criteria proposed in the Cochrane Handbook and CRD Handbook.[32,33] The study quality varied, with no one study scoring well in every item, and no item scoring well in every study. Of the studies included in the review, the study of highest overall quality appeared to be Szefler 2008[113] in which the only potential source of bias identified was study funding by a company with a commercial interest in FeNO measurement. The studies of lowest quality appeared to be Verini 2010[114] and Petsky 2010[116] neither of which were scored as 'low risk' on any of the quality assessment items. As it was a conference abstract, Petsky 2010[116] was at especially high risk of selective reporting, while Verini 2010[114] was at risk of bias as the statistical comparison data were presented poorly (as discussed below). Potential sources of bias for the evidence base as a whole are discussed below.

Figure 19: Methodological quality summary: review authors' judgements about each methodological quality item presented as percentages across all included studies



Green circles with + sign, low risk of bias; red circles with – sign, high risk of bias; yellow circles with ?, unclear risk of bias.

Risk of selection bias

All of the included studies were described as randomised. However, only Szefler 2008[113] provided sufficient information on sequence generation and allocation concealment. Randomisation was generated via a random number table and stratified by site, and group allocation was concealed from study investigators. In all other studies, the method of sequence generation and allocation concealment was not reported.

Risk of performance bias

In terms of blinding, only Szefler 2008[113] and Pijnenburg 2005[115] appeared to have performed adequate blinding for both participants and study personnel. Fritsch 2006[112] had blinded participants, but did not report whether this was the case for study personnel, and so was rated as 'unclear' on this item. Neither Petsky 2010[116] nor Verini 2010[114] provided sufficient information to make a judgment

on participant and personnel blinding, and so were rated as 'unclear'. Since much of the step up / step down protocols and criteria for exacerbations were based on participant symptom reporting and physician judgement, any potential lack of blinding in the studies could significantly affect the direction and size of outcomes.

Risk of detection bias

Szefler 2008[113] was the only study to clearly state that outcome assessment blinding had been performed. The poor reporting of outcome assessment blinding in the other studies means that unblinded outcome assessment may be a potentially important source of bias throughout this body of literature.

Risk of attrition bias

In terms of outcome data completeness, Pijnenburg 2005[115] and Szefler 2008[113] appeared to be at low risk of bias. There may have been some bias in terms of outcome data completeness in the remaining three studies. Fritsch 2006[112] did not report reasons for participant withdrawal or correction for missing FeNO values; there may have been missing outcome data in Petsky 2010,[116] but this is unknown as only a conference abstract of this study was identified. In Verini 2010,[114] it was reported that 64 patients were recruited. However, it was unclear whether this was the total n after dropout, or whether no participants dropped out. Consequently, the study was rated as 'unclear' on outcome data completeness.

Risk of reporting bias

Selective reporting risk may also have been present in some of the data. Petsky 2010[116] was a conference abstract, so this was rated as 'high risk'. There was some potential reporting bias in Fritsch 2006[112], in that medication usage was reported as median (IQR) rather than mean values, and so these values could not be be used in the planned meta-analysis. However, though planned, a meta-analysis was not possible due to study heterogeneity, so this bias did not affect the synthesis of data. Verini 2010[114] was rated as having a low risk on selective reporting; respiratory function and immunoallerlogical parameters were inadequately reported (i.e. no numerical data were provided, it was only stated that there were no significant between-group differences on these outcomes), but these outcomes were not of relevance to this review. Both Szefler 2008[113] and Pijnenburg 2005[115] appear to have reported all the outcomes they set out to measure, and so were rated as 'low risk'.

Risk of other bias

There were a number of further potential sources of bias in each of the studies. Fritsch 2006[112] and Szefler 2008[113] were both in receipt of sponsorship from the pharmaceutical industry, and there was evidence of some such sponsorship in Pijnenburg 2005.[115] The statistical comparison data reported in

Verini 2010[114] was of a poor quality, in that most comparisons were presented as within-group longitudinal trends. Pijnenburg 2005,[115] Szefler 2008[113] and Fritsch 2006[112] all conducted run-in periods before randomisation, which have an unknown risk of bias attached to them. Finally, it was unclear whether there may have been additional sources of bias in Petsky 2010[116], as this research was presented as a conference abstract only.

Summary

The quality of the sampled literature was variable, with selective reporting being the most common potential source of bias. Fritsch 2006[112], Petsky 2010[116], and Verini 2010[114] were rated as 'unclear' on the majority of quality items. Such ratings were given for those aspects of study design which were not clearly presented within the articles themselves, so it was unclear whether there is likely to be bias in the conduct of the studies themselves, or whether the lack of clarity was a result of inadequate reporting. Other common potential sources of bias were random sequence generation, allocation concealment, and outcome assessor blinding (4 out of 5 studies each). Pharmaceutical industry sponsorship was declared as the source of funding for Fritsch 2006[112] and Szefler 2008,[113] and at least partially funded the activities of one author in Pijnenburg 2005.[115] The studies at the lowest risk of overall bias appeared to be Szefler 2008[113] (low risk on six of seven items), and Pijnenburg 2005[115] (low risk on three of seven items). Indeed, only these two studies could be confirmed to have blinded both participants and study personnel, and to be free of reporting bias. Fritsch 2006[112] was the only other study to present any usable information on blinding procedures: they blinded participants, but it was unclear whether this was also the case for study personnel. Possible lack of participant blinding may be a particularly important source of bias, given that the study outcomes were largely based on subjective measurements (ie. self-reporting of symptoms); blinding of study personnel may also be an important source of bias as they decide whether a patient's medication step should be changed and there is some degree of interpretation in this decision. Finally, there was the possibility of selective reporting in Fritsch 2006[112] and Petsky 2010[116], which may predispose the results to favour intervention over control.

b. Study and patient characteristics

Unlike other reviews of FeNO for asthma management, in this review, our primary analysis of studies which assess the efficacy of guiding treatment by FeNO measurement in adults considers the study in pregnant women separately (Powell 2011[111]), as this subgroup of patients was defined *a priori* as a separate group. This study is described and discussed in Section 5.2.3.3.1. This current section considers the other four studies in adults.

Study design and timeline of studies: All four studies were RCTs. Szefler 2008[113] and Pijnenburg 2005[115] both reported double-blinding, while Fritsch 2006[112] reported blinding of participants only.

In the remaining study, [114] blinding status of participants and outcome assessors was unclear. Study timelines are presented in Table 35. Study duration ranged from 6 months[112] to 12 months. [114,116] Pijnenburg 2005,[115] Szefler 2008[113] and Fritsch 2006[112] reported run-in periods of two, three, and four weeks, respectively. In Pijnenburg 2005[115] and Fritsch 2006[112] details of the run-in period were not stated. In Szefler 2008[113] patients were provided with a treatment programme based on previous treatment, adherence and control. Verini, 2010[114] and Petsky 2010[116] reported no run-in. In addition, the frequency of visits varied from study to study. Fritsch 2006[112] and Szefler 2008[113] reported visits every 6 to 8 weeks; Pijnenburg 2005[115] every 3 months; and Verini, 2010[114] every six months. Petsky 2010[116] did not report frequency of visits, but provided outcomes for 12 months only.

Table 35: Children management review: Study timelines

Author Year	Timeline of study	Final assessment
Fritsch 2006[112]	Visit 1. 4 week run-in Visit 2. Randomisation Visit 3. Vists at 6, 12, 18 and 24 weeks. Symptoms, short-acting ß-agonist use, anti-inflammatory treatment, FeNO and spirometry recorded. Bronchial challenge test (4.5% hypertonic saline) between 1st and 2nd visit.	24 weeks
Szefler 2008[113]	Visit 1: assessed asthma symptoms, pulmonary function, skin-test, sensitivity, adherence and level of asthma control (NHLBI guidelines). Run-in period of 3 weeks on a regimen based on standard treatment - physicians selected a treatment programme (one of six) based on previous treatment, adherence and asthma control. 10 minute session about adherence. Adherence measured during run-in period with diskus inhaler and questionnaire. Centralised block randomisation and visit every 6 to 8 weeks for 46 weeks. Each visit: FeNO, days of symptoms, use of rescue drugs, pulmonary function, use of health care, adherence to treatment, missed days of school from asthma. Data entered into computer and treatment option computed based on random allocation. 546 were then randomly assigned to 46 weeks of either standard treatment or standard treatment modified on the basis of measurements of FeNO	46 weeks
Verini 2010[114]	Visit 1. Baseline Visit 2. 6 months Visit 3. 12 months ASS, Asthma exacerbation frequency, Asthma Therapy Score and immunoallergological and functional data recorded at each visit.	12 months
Pijnenburg 2005[115]	Visit 0: 2 week run-in. Visit 1: randomisation, FeNO, FEV_1, PD20, diary card. visit 2 to 4: visit every 3 months, FeNO and symptoms (diary card previous 2 weeks) recorded at each visit. Visit 5: FeNO, FEV_1, PD20, diary card	9 months
Petsky 2010[116]	Spirometry, FeNO QoL and asthma/cough diary every visit.	12 months

FeNO, fractional exhaled nitric oxidel; NHLBI, National Heart, Lung and Blood Institute; ASS, asthma severity score; FEV_1, forced expiratory volume in first second; PD20, The dose of methacholine causing a 20% fall in forced expiratory volume in 1 sec (FEV1) from baseline; QoL, quality of life.

Population: Eligibility criteria varied from study to study. With the exception of Verini, 2010[114] all studies included children with confirmed or persistent asthma. It is difficult to determine whether severity was comparable in the studies as scores have not been reported in a way that allowed comparison. Szefler 2008[113] reported asthma control test scores; Verini 2010[114] classified participants on the basis of GINA scores; and Pijnenburg 2005[115] reported mean daily symptom scores. Some insight into severity can be gained from considering the inclusion criteria and setting of each study. Three studies appeared to recruit patients who were uncontrolled. These were Verini 2010[114], Szefler 2008[113] and Petsky 2010.[116] Verini 2010[114] and Petsky 2010[116] recruited patients attending a specialist clinic, perhaps suggesting difficult to control patients, though in Verini 2010[114] patients had not yet started ICS therapy; an alternative explanation would be that all patients are sent to a specialist clinic before starting ICS therapy in this region, and the patients were therefore not necessarily uncontrolled. Szefler 2008[113] recruited only patients with evidence of persistent or uncontrolled disease. Studies recruiting patients who are difficult to control are only likely to capture the efficacy of FeNO for increasing control, but not for decreasing ICS use in patients who are well controlled. Fritsch 2006[112] recruited mild to moderate persistent asthma patients at an outpatients clinic. Pijnenburg 2005[115] recruited atopic asthma patients who were attending a children's hospital, though it is not clear if this was a scheduled appointment, an emergency admission or just the location of the study follow-up, and therefore unclear what level of severity these patients may have. As they had all had a stable dose of ICS for the previous 3 months, it may be reasonable to assume these patients were reasonably well controlled.

Interestingly, atopy is a known confounder to FeNO measurements as atopic subjects have raised FeNO levels regardless of asthma status. Asthmatic atopic patients are thought to have the highest levels overall, and so FeNO is theoretically able to distinguish between controlled and uncontrolled atopic asthmatics as well as controlled and uncontrolled non-atopic asthmatics.[117] However, atopic asthma patients tend to have ICS responsive asthma more often than non-atopic patients, and so studies recruiting atopic patients will have limited generalisability. It is unclear whether studies in atopic patients will over or underestimate efficacy, or have no impact at all, though clinical input to the assessment suggested that it would be expected to increase estimates of efficacy as atopy is correlated with ICS responsiveness. Atopic patients were recruited by Fritsch 2006,[112] Verini 2010[114] and Pijnenburg 2005.[115]

Petsky 2010[116] was the only study to include young children only, with all other studies including adolescents and/or young adults as well as young children, or adolescents only in the case of Szefler 2008[113] (Table 36). The studies also varied in terms of size (range: n=52[112] to n=546[113]), and baseline FeNO was inconsistently reported: Fritsch 2006[112] and Szefler 2008[113] utilised median ppb, Verini, 2010[114] reported the mean, and Pijnenburg 2005[115] reported the geometric mean and range.

Table 36: Children management review: Study and population characteristics

Author Year	Study details	Inclusion / exclusion criteria	n analysed / N recruited	Age (years) Gender	Spirometry mean (SD)	Severity	FeNO	Atopic / smokers	Medication use
Fritsch 2006[112]	**Setting** Vienna, Austria; outpatient clinic **Funding** Technical and analytic support from Aerocrine **Study design** RCT, single blind, parallel groups	Patients with mild to moderate persistent asthma with a positive skin prick test (SPT) or radioallergosorbent test (RAST>1) to at least one of seven common aeroallergens (cat, dog, house dust mite, alternaria, birch-, hazelnut-, and mixed grass-pollen) in past medical history or at the time of recruitment.	47 / 52 **I:** 22 analysed, n recruited unclear **C:** 25/ analysed, n recruited unclear	**Children and adolescents** (6 to 18) **Mean age** **I:** 11.3±3.4 C: 12.1±2.8 **Male** 28/47 (59.6%)	**FEV_1% (median, IQR)** **I:** 101 (91.1 to 107.5) **C:** 93.7 (83.8 to 99.6)	NR	**ppb Median (IQR)** **I:** 34.6 (17.5 to 58.6) **C:** 31 (20.8 to 54.8) Difference NS	**Atopic:** Assume 100% **Smokers:** NR	**ICS dose µg/day Median (25% - 75%)** **I:** 230 (100 to 400) **C:** 140 (0 to 400) **ß-agonist puffs x day-1 Median (IQR)** **I:** 1 (0 to 7) **C:** 0 (0 to 2)
Szefler 2008[113]	**Setting** USA, 10 sites **Funding** Non-industry, though one author received speaker fees from aerocrine. **Study design** randomised, multicentre, double-blind, parallel-group trial	Eligibility restricted to residents of urban census tracts in which at least 20% of households had incomes below the federal poverty threshold. Eligible participants had been diagnosed to have asthma by physicians. Those on long term control included only if they had persistent asthma or evidence of uncontrolled disease, and all others must have both symptoms of	534/546 **I:** 267/276 **C:** 267/270	**Children and young adults** (12 to 20) **Mean age** 14.4 (2.1) **Male** 288/546 (53%)	FEV_1 (proportion of best FEV_1) **I:** 95.9% (15·5) **C:**95.7% (15·9) FEV_1/FVC **I:** 79·8 (9·0) **C:** 80·4 (8·3)	**Asthma control test score in last month (range 5 to 25)** **I:** 21.1 (3.6) **C:** 21.3 (3.2)	**ppb Median (25% - 75%)** **At enrolment** 31.7 (14.1-65.4) ppb **At randomisation** **I:** 20.5 (11.5-	**Atopic:** NR **Smokers:** 0%	NR

Author Year	Study details	Inclusion / exclusion criteria	n analysed / N recruited	Age (years) Gender	Spirometry mean (SD)	Severity	FeNO	Atopic / smokers	Medication use
		persistent asthma and evidence of uncontrolled disease. Excluded if adherence less than 25% or smoker.					45.3) **C:** 19.7 (10.9-38.0)		
Verini 2010[114]	**Setting** Italy, secondary care hospital **Funding** NR **Study design** Single centre, parallel group RCT	Patients referred for allergic asthma - not all already on ICS.	64/64 **I:** 32/32 **C:** 32/32	**Children and adolescents** (6 to 17) **Mean age** **I:** 10.7 ±2.4 **C:**11.3 ±2.1 **Male** 36/64 (56.3)	NR	**I:** GINA score 0: **I:** 7 **C:** 7 GINA score 1: **I:** 18 **C:** 19 GINA score 2,3: **I:** 7 **C:** 6	**Mean ppb ± SD** **Baseline** **I:** 13.78 (±12.31) **C:** NR	**Atopic:** 64/64 (100%) **Smokers:** NR	NR
Pijnenburg 2005[115]	**Setting** Rotterdam, Netherlands, Children's hospital **Funding** Some non-industry plus grant from Aerocrine to department.	Children with atopic asthma who had been using ICS at a constant dose for previous 3 months	85/89 **I:** 39/42 **C:** 46/47	**Children and adolescents** (6 to 18) **Mean age** 11.9 ± 2.9 (int); 12.6 ± 2.8 (control)	**FEV_1, % pred** **I** 96 (14) **C** 99 (20)	**Mean daily symptom score** **I** 1.4 (2.0) **C** 2.0 (2.4)	**ppb Geometric mean (range).** **I** 26.4 (5.6–134.9) **C** 29.8	**Atopic: I** 39/39 (100%) **C** 46/46 (100%) **Smokers:** NR	**mean daily ß-2 agonists, puffs** **I** 0.4 (±0.6) **C** 0.4 (±0.5) **Initial ICS dose, µg**

Author Year	Study details	Inclusion / exclusion criteria	n analysed / N recruited	Age (years) Gender	Spirometry mean (SD)	Severity	FeNO	Atopic / smokers	Medication use
	Study design Randomised, parallel group, double blind			**Male** n= 55			(3.1–117.5)		**I** 762 (335) **C** 746 (410)
Petsky 2010[116]	**Setting** Australia (clinical setting NR) **Funding** Non-industry **Study design** RCT	Children with asthma and on ICS attending a paediatric specialist clinic Excluded if other cardiorespiratory illness, if poorly compliant, or if unable to take ICS or LABA.	63/63 **I:** 31/31 **C:** 32/32	**Children** (min 4 years, max NR) **Mean age** NR **Male** NR	**FEV_1% predicated (95% CI)** **I** Baseline: 101.3 (90.9 to 111.7) At 12 months: 106.1 (91.9 to 120.2) **C** Baseline: 84.28 (63.9 to 104.7) At 12 months: 91.16 (84.7 to 97.7)	NR	NR	**Atopic:** NR **Smokers:** NR	NR

n, number; N, total number; SD, standard deviation; FeNO, fractional exhaled nitric oxide; IQR, interquartile range; I, intervention group; C, control group; ICS, inhaled corticosteroids; FEV^1, forced expiratory volume in first second; NR, not reported; RCT, randomised controlled trial; LABA, long acting β2 agonist; ppb, parts per billion; GINA, Global Initiative for Asthma; CI, confidence interval.

Interventions: Table 37 describes the interventions used in each study. The NIOX device was used in three of the studies (Fritsch,[112] Szefler 2008[113] Pijnenburg 2005[115]); Verini, 2010[114] used the Ecomedics CLD 88, and it was unclear what device was used by Petsky 2010.[116] None of the studies used the same protocol or cut-off points for management of asthma with FeNO. The protocol of Fritsch 2006[112] was based on FeNO readings only, with a cut-off >20ppb. Szefler 2008[113] specified four levels of cut-off: level 1 was determined by the presence of 0-3 days of symptoms, 0-1 nights of symptoms, % of personal best FEV_1 $\geq$80% and a FeNO of 0-20 ppb, while level 4 was determined by 14 days of symptoms, five to 14 nights of symptoms, <70% personal best FEV_1, and FeNO >40ppb. Verini 2010[114] based their cut-off on a FeNO of 12ppb and GINA guidelines; Pijnenburg 2005[115] used a cut-off of >30ppb, and Petsky 2010[116] did not specify a cut-off (Table 37). Treatments indicated at each step were also highly heterogeneous across studies, with Fritsch 2006[112] and Szefler 2008[113] indicating doses for ICS, LTRA and LABA, but at different doses and at different steps in the pathway and Pijinburg only indicating doses for ICS. Verini 2010[114] and Petsky 2010[116] did not report doses. Importantly, in the Szefler 2008[113] study, step-down could not occur on the basis of low FeNO if the patient was still experiencing symptoms. This may have limited any potential reduction in ICS use for patients who are non-ICs responsive.

Control: Table 38 provides details of the control interventions. As with the interventions, none of the studies used the same criteria, protocols or treatment doses for the management of asthma in the control arm of the study. Management was typically guided by symptom severity and/or FEV_1. Verini 2010[114] used GINA guidelines as a control.

Table 37: Children management review: Description of the intervention management strategies

Author Year	Decisions based on flow rate, device cut-off points	Step-up / step-down protocol	Doses
Fritsch 2006[112]	**Based on** FeNO readings only **Flow rate, device** 50 ml/ sec with the single breath online method using the NIOX **Cut-offs** FeNO >20 ppb	In patients with stable asthma increased FeNO was considered a sign of insufficient anti-inflammatory treatment (either due to insufficient dosing or low adherence to prescribed therapy). Hence, aimed to improve the adherence to therapy in patients on anti-inflammatory treatment with raised FeNO. These patients were provided with 2-week diary cards to record daily symptoms, b-agonists use and controller medication requirement, and telephone calls were regularly performed to check adherence to therapy. Asymptomatic patients on therapy with b-agonist on demand only, with normal lung function but increased FeNO were prescribed low dose steroids. Step up was performed irrespective of FeNO level if FEV1%predicted was <80% and/or there were severe symptoms over the last 4 weeks and/or b-agonist use was ≥6 puffs over the last 14 days. If FeNO was raised in these patients, they received 2-week diary cards as well. Step down was performed if FEV1%predicted was ≥80% and there were no or mild symptoms over the last 4 weeks and b-agonist use was <6 puffs over the last 14 days and FeNO was ≤20 ppb.	Low dose ICS: (2 x 100 µg fluticasone or 2 x 200 µg budesonide) Low dose ICS + leukotriene receptor agonists: (2 x 100 µg fluticasone or 2 x 200 µg budesonide + 5 mg montelukast once daily p.o.) Low dose ICS + long acting ß –agonists (2 x 100 µg fluticasone + 2 x 50 µg salmeterol or 2 x 200 µg Budesonide + 2 x 12 µg Formeterol); High dose ICS + leukotriene receptor agonists (2 x 250 µg fluticasone or budesonide 2 x 400 µg + 1 daily 5 mg montelukast p.o.) High dose ICS + long acting ß -agonists (2 x 250 µg fluticasone + 2 x 50 µg salmeterol or 2 x 400 µg budesonide + 2 x 12 µg formeterol).
Szefler 2008[113]	**Based on** Days/ nights of symptoms (patient recall over past 2 weeks); FEV_1 as % of personal best; FeNO **Flow rate, device** flow rate	Step up/down based on predefined levels of control: Level 1 - days of symptoms in past 2 weeks 0 to 3; nights of symptoms in past 2 weeks 0 to 1; % of personal best FEV_1 ≥80%; FeNO 0 to 20ppb. Medication would not change at this level, or if 2	Step 0 - no controller medication; rescue treatment with salbutamol as needed Step 1 - Fluticasone by dry powder inhaler 100µg per day Step 2 - Fluticasone by dry powder inhaler 100µg b.i.d. Step 3 - Fluticasone by dry powder inhaler 100µg per day and salmetarol 50µg b.i.d. Step 4 - Fluticasone by dry powder inhaler 250µg per day

Author Year	Decisions based on flow rate, device cut-off points	Step-up / step-down protocol	Doses
	50 mL/s; NIOX **Cut-offs**	consecutive visits at level 1, may be stepped down. Level 2 - days of symptoms in past 2 weeks 4 to 9; nights of symptoms in past 2 weeks 2; % of personal best FEV_1 ≥80%; FeNO 20.1 to 30ppb. Increase medication 1 step. Level 3 - days of symptoms in past 2 weeks 10 to 13; nights of symptoms in past 2 weeks 3 to 4; % of personal best FEV_1 70 to 79%; FeNO 30.1 to 40ppb Increase medication two steps. Level 4 - days of symptoms in past 2 weeks 14; nights of symptoms in past 2 weeks 5 to 14; % of personal best FEV_1 <870%; FeNO >40ppb Increase by three steps or two steps + prednisone.	and salmetarol 50µg b.i.d. Step 5 - Fluticasone by dry powder inhaler 500µg per day and salmetarol 50µg b.i.d. Step 6 - Fluticasone by dry powder inhaler 500µg per day and salmetarol 50µg b.i.d. plus either low-dose theophylline or montelukast every day.
Verini 2010[114]	**Based on** GINA guidelines plus FeNO values. **Flow rate, device** Ecomedics CLD 88 (flow rate NR) **Cut-offs** 12 ppb	Values above 12ppb lead to increased medication Values below 12ppb lead to a reduction or maintenance of amount of drugs. Changes in drugs not reported. Unclear whether FeNO used at visit 2 only to guide therapy.	NR
Pijnenburg 2005[115]	**Based on** FeNO and symptoms **Flow rate, device** NIOX analyser (presume 50 mL/s) **Cut-offs** FeNO >30ppb = ICS increased; FeNO ≤30ppb AND symptoms >14 = ICS stays same; FeNO ≤30ppb AND	FeNO >30ppb = ICS increased; FeNO ≤30ppb AND symptoms >14 = ICS stays same; FeNO ≤30ppb AND symptoms ≤ 14 = ICS decreased	ICS doses: 100µg: increase to 200µg, decrease to 0µg 200µg: increase to 400µg, decrease to 100µg 400µg: increase to 800µg, decrease to 200µg 500µg: increase to 100µg, decrease to 250µg 800µg: increase to 1200µg, decrease to 400µg 1000µg: increase to 1500µg, decrease to 500µg 1200µg: increase to 1600µg, decrease to 800µg 1600µg: increase to 2000µg, decrease to 1200µg 2000µg: no further increase, decrease to 1000µg

Author Year	Decisions based on flow rate, device cut-off points	Step-up / step-down protocol	Doses
	symptoms ≤ 14 = ICS decreased		
Petsky 2010[116]	**Based on** FeNO and atopy **Flow rate, device** NR **Cut-offs** NR	NR (Treatment adjusted according to exhaled nitric oxide result, monthly for 4 months, then second monthly for 8 months)	NR

FeNO, fractional exhaled nitric oxide; ICS, inhaled corticosteroids; ppb, parts per billion; b.i.d, twice per day

Table 38: Children management review: Description of the control group management strategies

Author Year	Decisions based on	Step-up/step-down protocol	Doses
Fritsch 2006[112]	Asthma control (symptoms, short-acting b-agonist use, lung function), as recommended in current (German) asthma guidelines.[118]	A step down in therapy was performed if FEV_1 % predicted was ≥80% and there were no or mild symptoms over the last 4 weeks and b-agonist use was <6 puffs over the last 14 days. A step up was performed in every other case	Low dose ICS: (2 x 100 µg fluticasone or 2 x 200 µg Budesonide) Low dose ICS + leukotriene receptor agonists: (2 x 100 µg fluticasone or 2 x 200 µg budesonide + 5 mg montelukast once daily p.o.) Low dose ICS + long acting ß-agonists(2 x 100 µg fluticasone + 2 x 50 µg salmeterol or 2 x 200 µg budesonide + 2 x 12 µg Formeterol); High dose ICS + leukotriene receptor agonists (2 x 250 µg fluticasone or budesonide 2 x 400 µg + 1 daily 5 mg montelukast p.o.) High dose ICS + long acting ß -agonists (2 x 250 µg fluticasone + 2 x 50 µg salmeterol or 2 x 400 µg budesonide + 2 x 12 µg formeterol).
Szefler 2008[113]	National Asthma Education and Prevention Program (symptoms, treatment use, lung function)[119]	Control group received standard treatment, based on the guidelines of the National Asthma Education and Prevention Program (ie as intervention but without FeNO measurements)	Step 0 - no controller medication; rescue treatment with salbutamol as needed Step 1 - Fluticasone by dry powder inhaler 100µg per day Step 2 - Fluticasone by dry powder inhaler 100µg b.i.d. Step 3 - Fluticasone by dry powder inhaler 100µg per day and salmetarol 50µg b.i.d. Step 4 - Fluticasone by dry powder inhaler 250µg per day and salmetarol 50µg b.i.d. Step 5 - Fluticasone by dry powder inhaler 500µg per day and salmetarol 50µg b.i.d. Step 6 - Fluticasone by dry powder inhaler 500µg per day and salmetarol 50µg b.i.d. plus either low-dose theophylline or montelukast every day.

Author Year	Decisions based on	Step-up/step-down protocol	Doses
Verini 2010[114]	GINA (symptoms, SABA use, lung function)	GINA guidelines (specific step up/down protocol not described)	NR
Pijnenburg 2005[115]	Symptoms	Symptom score >14 = ICS increase symptoms score ≤14 for first time - ICS stays same symptoms score ≤14 for second time - ICS decrease Symptom score calculated as mean daily scores for dyspnea, wheezing, cough, daytime and nighttime, each scored 0-3. Max score 18. Also use of ß2 agonsits and % symptom-free days. Calculated over previuos 2 weeks for monitoring, over previous 4 weeks for end-point evaluation.	ICS doses: 100µg increase to 200µg, decrease to 0µg 200µg increase to 400µg, decrease to 100µg 400µg increase to 800µg, decrease to 200µg 500µg increase to 100µg, decrease to 250µg 800µg increase to 1200µg, decrease to 400µg 1000µg increase to 1500µg, decrease to 500µg 1200µg increase to 1600µg, decrease to 800µg 1600 increase to 2000µg, decrease to 1200µg 2000µg no further increase, decrease to 1000µg
Petsky 2010[116]	Symptoms / FEV_1	Unclear	NR

FeNO, fractional exhaled nitric oxide; ICS, inhaled corticosteroids; ppb, parts per billion; b.i.d, twice per day; GINA, the global initiative for asthma; SABA, short acting β2 agonists; NR, not reported.

c. Estimates of efficacy

Exacerbations

All five studies provided some data on asthma exacerbations, although it was unclear in some cases what the precise definition of an exacerbation was (Table 39).

Exacerbations resulting in OCS use

Data on exacerbations resulting in OCS use were reported in Pijnenburg 2005,[115] Szefler 2008[113] and Fritsch 2006.[112] Only Szefler 2008[113] and Pijnenburg 2005[115] reported the number as (or data in a way that allowed calculation of) the rate per person year. In both cases, rates were lower in the intervention arm; In Szefler 2008[113] the reviewer-calculated rate per person year in the intervention arm was 0.746 and in the control arm 0.950. In Pijnenburg 2005[115] the reviewer-calculated rates were 0.21 and 0.39 respectively. There is a large difference in the rates between studies, which is probably due to the inclusion criteria applied; Szefler 2008[113]only included patients with evidence of uncontrolled disease and Pijnenburg 2005[115] included only patients who had been on a stable dose of ICS for three months. As this data was reviewer-calculated, significance of the differences is unclear, though the mean difference comparison provided by Szefler 2008[113] indicated a non-significant difference and the difference in the number of people experiencing an exacerbation in Pijnenburg 2005[115] was also non-significant. Rate ratios calculated by the reviewer were 0.79 in Szefler 2008[113] and 0.54 in Pijnenburg 2005.[115]

All exacerbations or treatment failures

Fritsch 2006,[112] Szefler 2008[113] Verini 2010[114] and Petsky 2010[116] all reported outcomes that were not defined as either major or minor, and used different definitions to one another. All four studies reported fewer exacerbations or treatment failures in the intervention arm, but these differences were not always statistically significant. In Fritsch 2006[112] the outcome was a composite that appeared to include moderate to severe exacerbations. There were 17 exacerbations out of 88 observations in the intervention group (18.2%), vs 22 exacerbations out of 99 observations in the control group (21.2%). These data was not convertible to rates as some data points were missing in the study. The difference was not significant at the $p<0.05$ level. Szefler 2008[113] also used a composite outcome which appeared also to relate to moderate to severe exacerbations. This study reported the percentage of patients in each group with more than one exacerbation; these were 37% in the intervention group and 43.6% in the control group (RR: −6·5(95% CI: −14·4 to 1·4, p=0·11)). Verini, 2010[114] reported events leading to use of SABA, which appeared to be likely to incorporate minor to major exacerbations. Rates per person year were 0.83 ± 0.98 and 1.85 ± 1.34, in the intervention arm and control arm respectively. No between-group comparisons were presented for this outcome. The only study to report a significant between-group difference in the number of participants with exacerbations was the conference abstract Petsky 2010.[116] Exacerbations were not clearly defined, but

occurred in 6/31 participants in the control group (19.4%), and 15/32 in the control group (46.9%, p=0.021). Pijnenburgh 2005[115] did not report this outcome.

Table 39: Children management review: Exacerbation and oral corticosteroid use rates in children and adolescents with or without FeNO guided management

Author Year Time of outcome	Definition of outcomes	Intervention	Control	Between-group comparison
Fritsch 2006[112] **4-week run-in and 24-week trial**	**Unclear severity:** Exacerbation defined as: Oral steroid courses because of asthma symptoms over the last 4 weeks and/or unscheduled visit because of asthma symptoms over the last 4 weeks And/or Increase of asthma symptoms from a symptom score of 0 or 1 to a score of 2 And/or Decline in FEV >10% compared with last visit (n patients / group)	17/88 observations (18.2%) Unable to calculate rate per person year due to missing data points	22/99 observations (21.2%) Unable to calculate rate per person year due to missing data points	p=NS
	Major exacerbations (Oral corticosteroid use, (n patients / group)	2 people	2 people	p=NS
Szefler 2008[113] **Run-in of 3 weeks plus 46-week trial**	**Major exacerbations:** Prednisone courses	mean 0.66 (SE: 0.085) events over 46 weeks = 0.746 per person year	0.84 (SE: 0.085) events over 46 weeks = 0.950 per person year	Mean diff: 0.17 (95% CI: -0.08, 0.41), p=0.14
	Unclear severity: Exacerbation (defined as a composite outcome consisting of admissions to hospital, unscheduled visits, and prednisone use).	37.0% (SD 4.83) people	43.6% (SD 4.96) people	Mean diff: -6.5% (95% CI: -14.4, 1.4), p=0.11
	≥1 hospital admission	3·3% (SD 1·78) people	4·1% (SD 1·98) people	Mean diff: -0.8 (95% CI: -4.0, 2.3), p=0.61
	≥1 unscheduled use of health care	21·3% (SD 4·09) people	22·7% (SD 4·19) people	Mean diff: -1.4 (95% CI: -9.3, 6.7), p=0.74
	≥1 prednisone course	32.1% (SD 4.67) people	42.0% (SD 4.94) people	Mean diff: -10.3% (95% CI: -18.5, -2.2), p=0.01

Author Year Time of outcome	Definition of outcomes	Intervention	Control	Between-group comparison
Verini 2010[114] **12 months**	**Exacerbations (appears to be minor to major):** ATS 2005 definition - number of episodes of coughing, dyspnea, wheezing requiring SABA	0.83 (SD±0.98) per person year	1.85 (SD±1.34) per person year	Between-group comparison NR
Pijnenburg 2005[115] **2-week run-in; 12-month trial**	**Major exacerbations:** Prednisone courses	8 events in one year =8/39 = 0.21 per patient year	18 events in one year = 18/46 = 0.39 per patient year	p=0.60[a]
Petsky 2010[116] **12 months**	**Unclear severity:** Asthma exacerbations (severity not described):	6/31 (19.4%) people	15/32 (46.9%) people	p=0.021[b]

[a] p-value is for comparison of number of patients with exacerbations (7 in FeNO group, 10 in control group), rather than total exacerbations
[b] Values based on number of patients with exacerbations, not total events

NS, not statistically significant; FEV, forced expiratory volume; SE, standard error; SD, standard deviation; CI, confidence interval; SABA, short acting β2 agonist; ATS, American Thoracic Society; NR, not reported

ICS use

Table 40 provides details of ICS use in each study. With the exception of Petsky 2010[116] all the studies provided some indication of between-group differences in ICS use. Fritsch 2006[112] and Szefler 2008[113] reported statistically significantly higher ICS use in the FeNO group, Pijnenburg 2005[115] reported very similar levels, while the values in the remaining study (in terms of absolute N using ICS) were difficult to interpret. Fritsch 2006[112] reported median (IQR) values for ICS use in the intervention and control groups; these were 316 µg (200 to 500), and 241 µg (26 to 607), respectively ($ß_1$ 0.20, $p<0.01$). Pijnenburg 2005[115] reported mean cumulative ICS dose from visits one to five (±SEM). The values were 4,407 (SEM 367)g for the intervention and 4,332 (SEM 383)g for control ($p=0.73$). Szefler 2008[113] reported the between-group difference in use of fluticasone, which was 119µg/day greater in the FeNO group by the final visit (95% CI 49 to 189, $p=0.001$). Finally, Verini 2010[114] reported the absolute number of participants using ICS from each group at each time point. In the intervention group, these numbers were: T1: 20; T2: 19; T3: 19; in the control group, these were: T1: 15; T2: 22; T3: 19. However, the baseline values for the groups in this study were not comparable, and the absolute number of participants using ICS, without concomitant data on dosages used, provides little understanding of between-group ICS use.

When studies recruiting patients who are or are likely to be difficult to control were considered (Fritsch 2006[112] and Szefler 2008[113]) versus the study that recruited patients who had been on a stable dose of ICS for three months, (Pijnenburg 2005[120]) it was seen that the two studies recruiting the difficult to control groups saw an increase in ICS useage, whilst the study that recruited stable patients saw similar levels of ICS use across both arms. This would be expected as the difficult to control group of patients are unlikely to need a dose reduction, whereas patients who are stable may be eligible for such a reduction. In addition to this, the Szefler 2008[113] protocol did not allow step-down of ICS on the basis of low FeNO levels if symptoms were still present, making step-down less likely to occur.

Table 40: Children management review: ICS use

Author Year	ICS type and measurement definition	Intervention	Control	Between-group difference
Fritsch 2006[112]	Fluticasone and budesonide permitted. Data reported as mean ICS dose (µg/day); unclear which ICS type the doses refer to	Baseline median (IQR) dose: 230 (100 to 400) Endpoint median (IQR) dose: 316 µg (200 to 500)	Baseline median (IQR) dose: 140 (0 to 400) Endpoint median (IQR) dose: 241 µg (26 to 607)	Repeated measures analysis ß10.20, p<0.01
Szefler 2008[113]	Fluticasone	NR	NR	difference 119µg/day (95% CI 49 to 189, p=0.001) (higher in intervention group)
Verini 2010[114]	Unclear what ICS used Measured in terms of absolute n of patients using per group at each time-point	Number of patients using: T1: 20 T2: 19 T3: 19	Number of patients using: T1: 15 T2: 22 T3: 19	NR
Pijnenburg 2005[115]	Budesonide (2mg max daily dose permitted) Cumulative steroid dose (sum of mean daily steroid doses visits 1 to 5)	Cumulative endpoint: 4,407 (367)µg	Cumulative endpoint: 4,332 (383)µg	p=0.73
Petsky 2010[116]	NR	NR	NR	NR

ICS, inhaled corticosteroid; IQR, interquartile range; NR, not reported; CI, confidence interval; T, Time; max, maximum.

HRQoL

Table 41 provides HRQoL data. Only Petsky 2010[116] provided data on HRQoL, and it was unclear which QoL tool was used. In the intervention group, the baseline mean was 84.38 (95% CI 77.27 to 91.48), which rose to 86 (95% CI: 74.84 to 97.1) at 12 months. Conversely, in the control group, the baseline mean of 86 (95% CI 81.49 to 90.51) dropped to 83.75 at 12 months (95% CI: 78.6 to 88.9). If QoL was measured with the EQ-5D, which seems likely, then higher values would indicate better QoL, and thus FeNO would be favoured. The endpoint difference was statistically significant (p=0.042), although it was unclear whether this comparison was for the change or for absolute end values.

Asthma control and other medication use

Table 41 provides details of outcomes relating to asthma control and medication use. Four studies provided some data on asthma control, none of which demonstrated any significant effects favouring either intervention or control, though in Verini 2010[114] significance was not reported. Furthermore, there was lack of uniformity in how asthma control was measured. Fritsch 2006[112] recorded the absolute number of participants per group whose symptom severity score increased to 2 (i.e. severe symptoms). Ten participants in the intervention (11.4%), and 11 in the control group (11.1%) fulfilled this criterion (difference NS). These researchers also reported the absolute number of participants per group who experienced a decline in FEV greater than 10%: there were seven in the intervention group (8%) and 13 in the control group (13.1%, p=NS). Szefler 2008[113] presented between-group differences for a number of symptomatic indicators of control, with higher numbers favouring intervention. The measure comprised days of wheeze (Risk ratio [RR]: 0·04 95% CI: −0·22 to 0·29, p=0·78), days of interference with activities (RR: 0·03, 95% CI: −0·21 to 0·26 p=0·83), nights of sleep disruption (RR: −0·08 95% CI: −0·26 to 0·10 p=0·38), days of school missed (RR: 0·03 −0·11 to 0·16 p=0·71), and asthma control test score in the last month (RR: −0·04 −0·12 to 0·05 p=0·38). Pijnenburg 2005[115] calculated symptom scores based on diary card data for dyspnea, wheezing and cough. Day and night were scored separately, and each symptom was scored between 0-3, giving a maximum possible total score of 18. The change in mean daily scores between baseline and visit 5 was 0.1 in the intervention group, and 0.6 in the control group (between group difference, p=0.4). Finally, Verini 2010[114] measured symptom control using the GINA scale, which classified participants as having remission asthma (GINA score: 0), intermittent asthma (GINA score: 1), or persistent asthma (GINA score: 2, 3). The means for these categorical data were presented for both groups at all three time points: at time one the values were 1.09 $\pm$ 0.81 (intervention), vs 1.09 $\pm$0.77 (control); at the second visit they were 0.56 $\pm$0.75 vs 0.93 $\pm$0.61 respectively; and at time three they were 0.75 $\pm$0.96 vs 0.92 $\pm$ 0.82. No inter-group comparisons were conducted, though means in the intervention arm are numerically lower than in the control arm.

With respect to additional medication use, three studies provided data using different metrics and mostly without formal comparison statistics. Szefler 2008[113] reported mean difference from baseline in salmeterol usage; These were -6.5μg/day in the intervention group, and -12 μg/day in the control group (p=NR). Verini, 2010[114] created a categorical measure of medication use, where antihistamines, ketotifen, cromones = 1; Specific Immuno therapy, LABA, LTRA = 2; ICS = 3. The means of these data were presented for both groups: for the intervention group, the means were 1.5 ± 0.7 (visit 1), 1.43 ± 0.7 (visit 2), and 1.53 ± 0.6 (visit 3). For comparison, the values for the control group at the same time-points were 1.03 ±0.9; 1.62 ± 0.6, and 1.4 ± 0.7, respectively. Verini, 2010[114] also provided absolute numbers using LTRA. For the intervention group these were 8, 8, and 11 at visits one, two, and three respectively. The control numbers for comparison were three, eight and seven. Fritsch 2006[112] reported the median (IQR) number of β-agonist puffs per day. For the intervention group, the number was one (IQR: 0 to 7), while in the control group the number was 0 (IQR: 0 to 2). Overall, two studies report numerically higher additional medication use in the intervention arm (Szefler 2008[113] and Verini 2010).[114]

Adverse events, mortality, compliance and test failure rates

Data on adverse events, mortality, and compliance were only reported in Szefler 2008.[113] No statistically significant differences between the groups were reported for any adverse events or for mortality. For the intervention and control groups, respectively, adverse events were reported for eyes, ears, nose and throat (8.3% vs 8.1%, p=0.87); gastrointestinal disorders (13.4% vs 14.1%, p=0.78); haematology disorders (27.2% vs 28.9%, p=0.44); infections (55.8% vs 52.2%, p=0.46); musculoskeletal symptoms (15.9% vs 18.5%, p=0.44); nervous system disorders (34.4% vs 33.7%, p=0.20); respiratory signs and symptoms (33.7% vs 34.1, p=0.92); and skin symptoms (15.6% vs 17.8%, p=0.18). Medication compliance was shown to be 86% for the intervention group and 92% for the control group; no mortality was observed in either group.

Table 41: Children management review: Other outcomes

	Outcome	**FeNO** Mean (SD)	**Control** Mean (SD)	**Comparison**
Asthma control				
Fritsch 2006[112]	Increase of symptoms to a score of 2 (severe)	10/88 (11.4%)	11/99 (11.1%)	NS at p<0.05
	Unscheduled visits	5/88 (5.7%)	5/99 (5.1%)	NS at p<0.05
	FEV_1 decline >10%	7/88 (8.0%)	13/99 (13.1%)	NS at p<0.05
	median β-agonist puffs/day	1 (IQR: 0-7)	0 (IQR:0-2)	NR
Szefler 2008[113]	Maximum days with symptoms	1·93 (2·60).	1·89 (2·69).	Mean diff: 0·04 (95% CI −0·22 to 0·29) p=0·78.
	Days of wheeze	1·71 (2·52)	1·69 (2·64)	Mean diff: 0·03 (95% CI −0·21 to 0·26) p=0·83
	Days of interference with activities	0·87 (1·79)	0·95 (1·98)	Mean diff: −0·08 (95% CI −0·26 to 0·10) p=0·38
	Nights of sleep disruption	0·52 (1·30)	0·50 (1·25)	0·03 (95% CI −0·11 to 0·16) p=0·71
	Days of school missed	0·19 (0·79)	0·23 (0·84)	−0·04 (95% CI −0·12 to 0·05) p=0·38
	Asthma control test score in the last month	21·89 (2·83)	21·83 (2·88)	0·06 (95% CI −0·28 to 0·40) p=0·72
Verini 2010[114]	Symptom score (mean for ordinal data: intermittent asthma =1; mild/moderate persistent asthma=2; severe persistent asthma=3)	T1: 1.09 (±0.81) T2: 0.56 (±0.75) T3: 0.75 (±0.95)	T1: 1.09 (±0.77) T2: 0.93 (±0.61) T3: 0.92 (±0.82)	NR
Pijnenburg 2005[115]	Change in mean symptom severity scores between visits 1 and 5	0.1	0.6	mean daily scores change p=0.40
HRQoL				
Petsky 2010[116]	HRQoL (metric not specified)	Baseline: Mean 84.38 (95% CI 77.27 to 91.48) 12 months: Mean 86 (74.84 to 97.1)	Baseline: mean 86 (95% CI 81.49 to 90.51) 12 months: 83.75 (78.6 to 88.9)	NR

	Outcome	**FeNO** Mean (SD)	**Control** Mean (SD)	**Comparison**
Other medication use				
Fritsch 2006[112]	-	-	-	-
Szefler 2008[113]	Salmeterol (mean difference from baseline)	-6.5µg/day difference	-12µg/day difference	NR
Verini 2010[114]	Mean of ordinal data, where Antihistamines, ketotifen, cromones = 1 Specific Immuno therapy, LABA, LTRA = 2 ICS = 3	T1: 1.5 (±0.7) T2: 1.43 (±0.7) T3: 1.53 (±0.6)	T1: 1.03 (±0.9) T2: 1.62 (±0.6) T3: 1.4 (±0.7)	NR
	Number of patients using LTRA	T1: 8 T2: 8 T3: 11	T1: 3 T2: 8 T3: 7	NR
	Number of patients using no anti-inflammatory drugs	T1: 4 T2: 5 T3: 2	T1: 14 T2: 2 T3: 6	NR
Pijnenburg 2005[115]	B-agonist use	NR	NR	p=0.28
Petsky 2010[116]	-	-	-	-

FeNO, fractional exhaled nitric oxide; NS, not statistically significant; FEV_1, forced expiratory volume in first second; SD, standard deviation; mean diff, mean difference; T, time; HRQoL, health related quality of life; CI, confidence interval; LABA, long acting β2 agonist; LTRA, leukotriene receptor antagonist; ICS, inhaled corticosteroids; 95% CI, 95% confidence interval

5.2.3.3 Subgroups relevant to the review as defined in the scoping workshop

5.2.3.3.1 Pregnant women

Only one randomised controlled trial of FeNO-guided management of asthmatic pregnant women was found. An assessment of the quality of this study is included in Section 5.2.3.1a.

Study and patient characteristics: The study, Powell 2011,[111] was conducted in Australia, was funded from a number of industry and non-industry sources (including lecture fees from Aerocrine), was double-blind and multicentre with 203 participants analysed out of 242 recruited. In the run-in period, patients not already using ICS were started on budesonide at a dose of 200 µg b.i.d. After randomisation, patients underwent monthly review and titration of ICS dose. Patients were phoned two weeks after each visit to assess symptoms and encourage adherence (Table 42). The study included no current smokers, but did include 156/206 (75.7%) atopic patients (Table 43).

Intervention: An Ecometrics device was used to measure FeNO and cut-off points were <16ppb and >29ppb. Step-up and step-down criteria are described in Table 44; decisions were based on a combination of FeNO and asthma control, with asthma control levels directing the dose of LABA (Formoterol) and FeNO controlling the dose of ICS (budesonide).

Comparator: (Table 45) patients were managed according to asthma control as assessed by ACQ. Medication doses were somewhat different to the intervention arm, with different doses of budesonide and formoterol indicated.

Table 42: Pregnant women: study design and timelines

Author Year	Study details	Timeline of study
Powell 2011[111]	**Setting** Australia, Antenatal outpatients **Funding** Mix; lecture fees from Aerocrine. **Study design** RCT: double-blind, parallel group, multicentre	Visit 1: FeNO, spirometry, ACQ, atopy (specific IgE to aeroallergen). Optimised self-management skills including inhaler technique, knowledge, action plan, and adherence. 2-week run-in period after visit 1 - continued use of ICS budesonide turbuhaler OR uncontolled patients not using ICS started on budesonide 200µg b.i.d. Visit 2: Randomisation. Asthma symptoms, FeNO, spirometry, ACQ, QoL questionnaires. Then monthly review with research assistant: recorded clinical symptoms, ACQ, present treatment, FeNO, FEV_1. FeNO, ACQ and treatment sent to algorithm keeper who applied algorithm and sent treatment recommendation to research assistant who informed patient. If patient uncontrolled and at maximum dose, seen by investigator Telephone assesments 2 weeks after visit to assess symptoms and encourage adherence.

RCT, randomised controlled trial; FeNO, fractional exhaled nitric oxide; ACQ, asthma control questionnaire; IgE, immunoglobulin E; ICS, inhaled corticosteroids; b.i.d. twice per day; FEV_1, forced expiratory volume in first second.

Table 43: Pregnant women: study and patient characteristics

Author, year	Age (years)	N analysed/N recruited	Inclusion/exclusion criteria	Spirometry Mean (SD)	Severity	FeNO	Smokers; Atopic	Medication use
Powell 2011[111]	**Pregnant Adults** >18 years **Mean age** 28 years (5.4) **Male** 0/220 (0%)	203/242 WBR: 22 **I:** 100/111 **C:** 103/109	Doctor's diagnosis confirmed by respiratory physician's diagnosis of asthma. Non-smoking pregnant women between 12 and 20 weeks gestation with doctor's diagnosis of asthma and who were using inhaled therapy in last year.	**FEV_1% mean (95%CI)** **I:** 95.1 (92.8 to 97.4) **C:** 96.1 (93.5 to 98.7) **FEV_1/FVC mean (95%CI)** **I:** 79.7 (75.4 to 78.0) **C:** 80.63 (79.3 to 82.0)	**AQLQ-M Median (IQR)** **I:** 0.8 (0.4 to 1.5) **C:** 1.0 (0.5 to 1.6) **ACQ Mean (read off graph)** **I:** 0.98 **C:** 1.01	**FeNO ppb median (IQR)** **I:** 13.9 (6.6 to 32.0) **C:** 13.1 (7.5	**Smokers** Current: 0/203 (0%) Ex-smokers: 80/203 (39.4%) **Atopic:** 156/206 (75.7%)	**Median days ß2 agonist in past week (IQR)** **I:** 1.0 (0 to 5) **C:** 2.0 (0 to 6) **ICS users** **I:** 46/111 (41.4%) **C:** 47/109 (43.1%) **Median BDP equivalent ICS dose (µg per day)** **I:** 800 (400 to 800) **C:** 800 (400 to 1600)

n, number; N, total number; SD, standard deviation; FeNO, fractional exhaled nitric oxide; FEV_1, forced expiratory volume in first second; CI, confidence interval; I, intervention group; C, control group; AQLQ-M, asthma quality of life questionnaire-marks; IQR, interquartile range; ppb, parts per billion;

Table 44: Pregnant women: detail of intervention group management strategies

Author, year	Decisions based on; Flow rate, device; Cut-off points	Step-up/step-down protocol	Doses
Powell 2011[111]	**Based on** FeNO and ACQ **Flow rate, device** ATS 2005, Ecomedics **Cut-offs** <16ppb 16 to 29 ppb >29 ppb	FeNO concentration use to adjust dose of ICS ACQ used to adjust dose of LABA FeNO >29 - ICS increase one step, LABA no change FeNO 16 to 29 and ACQ ≤1.5 - ICS no change, LABA no change FeNO 16 to 29 and ACQ >1.5 - ICS no change, LABA increase one step FeNO <16 and ACQ ≤1.5 - ICS decrease one step, LABA no change FeNO <16 and ACQ >1.5 - ICS decrease one step, LABA increase one step If a patient had undergone two ICS dose increments and FeNO remained >29ppb, ICS was not increased further. If still symptomatic (ACQ >1.5) formoterol 6µg b.i.d. was added. For patients taking formoterol, the ICS dose could never be 0, but would be reduced to 100µg b.i.d. Patients who remained uncontrolled at maximum doses referred to respiratory physician.	Steps 1 to 5 ICS: budesonide µg b.i.d.: 0, 100, 200, 400, 800 LABA: step 1 - slabutamol as required; step 2 to 5 - formoterol µg b.i.d.: 6, 12, 24, 24.

ATS, American Thoracic Society; FeNO, Fractional exhaled nitric oxide; ACQ, Asthma Control Questionnaire; ICS, inhaled corticosteroids; LABA, long acting ß2 agonist; ppb, parts per billion; b.i.d. twice per day;

Table 45: Pregnant women: detail of the control group management strategies

Author, year	Decisions based on	Step-up/step-down protocol	Doses
Powell 2011[111]	ACQ-guided	Well controlled asthma ACQ<0.75 - reduce treatment one step Partially controlled asthma ACQ 0.75 to 1.50 - no treatment change Uncontrolled asthma ACQ>1.5 - increase one step Those at maximum dose referred to respiratory physician	Step 1 - salbutamol as required Step 2 - budesonide 200µg b.i.d. plus salbutamol as required Step 3 - budesonide 400µg b.i.d. plus salbutamol as required Step 4 - budesonide 400µg and formoterol 12 µg b.i.d. Step 5 - budesonide 800µg b.i.d. and formoterol 24µg b.i.d.

ACQ, Asthma Control Questionnaire; b.i.d. twice per day.

Outcomes: Table 46 provides details of all outcomes of relevance to the review.

Exacerbations

As with the other studies in adults described in Section 5.2.3.1, the composite outcome of all exacerbations (in this case defined as an unscheduled visit to a doctor, presentation to the emergency

room or admission to hospital, or when OCS used) was reduced in the intervention arm, with a rate ratio of 0.496 (95% CI 0.325 to 0.755), p=0.001. Data for each element of the composite outcome were reported individually, and from this it can be seen that this difference is mostly driven by the rate of OCS use and the rate of doctor's visits. Mean OCS was 0.08 (95% CI 0.03 to 0.133) in the intervention arm, and 0.19 (95% CI 0.08 to 0.31) in the control arm: unlike other studies in adults, this did reach statistical significance with a p value of 0.042. Similarly, the rate of doctors' visits was 0.26 (95% CI 0.16 to 0.36) in the intervention arm and 0.56 (95% CI 0.40 to 0.72) in the control arm, with a p value of 0.002. The other components of the exacerbation outcome (hospitalisations and Emergency Room/labour ward visits) did not differ between groups.

ICS use

The change in mean values from baseline to final visit for ICS use was a decrease of 210µg/day in the intervention arm and an increase of 50µg/day in the control arm. This difference was statistically significant at p = 0.043. Interestingly, more women in the intervention arm were taking ICS (68.5% versus 42.2%), and median ICS dose as BDP equivalent µg/day was higher in the intervention arm (200 µg/day (IQR 0 to 400) compared to 0 µg/day (IQR 0 to 800)), but not statistically significantly so (p = 0.079).

HRQoL

Median scores and p values indicate a small but statistically significant difference in the Short form-12 (SF-12) mental component summary in favour of the intervention arm with a median score of 56.9 (IQR 50.2 to 59.3) versus 54.2 (IQR 46.1 to 57.6) (p = 0.037), but no statistically significant differences in the SF-12 physical summary or the AQLQ-M total score.

Asthma control and other medication use

The ACQ score at the end of the study indicated good control in both groups, with the mean indicating statistically significantly better control in the intervention group (0.56 (SD0.67)) than in the control group (0.72 (SD 0.80), p=0.046). ß2 agonist use in past week was higher in the intervention arm (p =0.024) and there were statistically significantly more LABA users in the intervention arm (p<0.0001).

Adverse events, mortality, compliance and test failure rates

None of these outcomes were reported in the study.

Conclusion

In pregnant women, the composite outcome of exacerbation rates, as well as OCS use on its own and doctors' visits were all statistically significantly better in the intervention arm. ICS use and ß2 agonist

use were also lower, though LABA was used by more patients in the intervention arm than the control arm. Asthma control was marginally better in the intervention arm, and the mental component of the SF-12 was better also, though the physical summary and the total score for the AQLQ-M were not statistically significantly different between groups. In summary, the use of FeNO to guide asthma management in pregnant women appears to be as effective if not more so than in other adults, and appears to reduce exacerbations and ICS use. This may be due to increased efficacy in pregnant women, or due to differences in step-up/step-down protocols. Notably, this protocol allowed for the step-down of ICS use on the basis of FeNO levels alone, regardless of whether symptoms were still present or not.

5.2.3.4 FeNO to assess compliance to treatment

One further study was identified for the management review (Beck-Ripp 2002).[121] This open-label RCT recruited patients aged 6–16 yrs with mild-to moderate persistent asthma (n=54), and excluded patients who had received oral or inhaled steroid treatment during the last 2 months, had acute upper airway infection, had a history of bad compliance (<65% of prescribed medication) or had any other serious illness. The trial consisted of a four-week run-in period, a four-week washout phase, and a final randomised treatment phase in which only one group was treated with inhaled budesonide, and FeNO was used to attempt to distinguish these groups – i.e., the study explored FeNO as a tool to assess patient compliance. The study showed that FeNO was able to distinguish those who had been treated with ICS more successfully than conventional lung-function parameters. However, as the study did not examine the efficacy of FeNO for guiding management per se, the data could not be compared with other studies in the management review. It should also be noted that this potential benefit of using FeNO in management of asthma will have been captured in the other RCT trials, if the management protocol included investigations of compliance before stepping treatment up or down.

Table 46: Pregnant women: all outcomes

Author, year Time of outcome	Definition of Outcomes	Intervention Mean (95% CI)	Control	Between group comparison
Exacerbations				
Monthly until birth (max approx 30 weeks)	Exacerbations: An unscheduled visit to a doctor, presentation to the emergency room or admission to hospital, or when OCS used. Events separated by 7 days or more were counted as a second event.	0.288 per pregnancy (mean study time 17.8 weeks(SD 5.5))	0.615 per pregnancy (mean study time 18.8 weeks (SD 3.8))	Incidence rate ratio 0.496 (95% CI 0.325 to 0.755), p=0.001
	OCS mean use (95% CI)	0.08 (0.03 to 0.133)	0.19 (0.08 to 0.31)	p value OCS use: 0.042
	Hospitalisations mean (95% CI)	0 (0 to 0)	0.03 (-0.004 to 0.06)	p = 1.0
	ER/labour ward visits mean (95% CI)	0.04 (0.001 to 0.07)	0.02 (-0.01 to 0.04)	p = 0.399
	Unplanned or unscheduled doctors' visits mean (95% CI)	0.26 (0.16 to 0.36)	0.56 (0.40 to 0.72)	p = 0.002
ICS use				
	Difference in means (from baseline to last visit) (read off graph):	- 210µg/day	50µg/day	p = 0.043
	BDP equivalent ICS dose (µg/day) Median (IQR)	200 (0 to 400)	0 (0 to 800)	p=0.079
	Users	76/111 (68.5%)	46/109 (42.2%)	p<0.0001
Other outcomes				
	HRQoL Median (IQR) SF-12 physical summary (low 0, high 100): SF-12 mental summary (low 0, high 100): AQLQ-M: total score (good 0, poor 10):	 47.7 (40.8 to 52.0) 56.9 (50.2 to 59.3) 0.75 (0.38 to 1.25)	 46.9 (38.2 to 51.8) 54.2 (46.1 to 57.6) 0.81 (0.38 to 1.63)	p values 0.89 0.037 0.54
	Asthma control : ACQ mean (SD)	0.56 (0.67)	0.72 (0.80)	p=0.046
	ß2 agonist use in past week: Median (IQR)	0 (0 to 3)	1 (0 to 5)	p = 0.024
	LABA users	45/111 (40.5%)	19/109 (17.4%)	p<0.0001
	Adverse events, mortality, compliance and test failure rates	NR	NR	NR

CI, confidence interval; OCS, oral corticosteroids; SD, standard deviation; ER, emergency room; BDP, beclometasone diproprionate; IQR, interquartile range; ACQ, Asthma Control Questionnaire; LABA, long acting ß2 agonistHRQoL, health realted quality of life; SF-12, short form 12; AQLQ-M, Asthma Quality of Life Questionnaire-Marks;

5.2.3.3.2 The elderly

In the absence of randomised control trials, other study designs were included to gather evidence for the use of FeNO in the management of asthma in the elderly. Five observational studies (three cohort,[122-124] one nested case control[125] and one cross-sectional study[126]), published between 2010 and

2012, were identified which evaluated the measurement of FeNO in elderly asthmatics.[123,123-126] A summary of the design, patient characteristics and outcomes of the included studies are summarised in Table 47. Columbo[123], Ross[126] and Smith[125] were based in the USA, Inoue[122] in Japan and Roh[124] in Korea. Only two studies reported where the patients were recruited from: Inoue[122] recruited patients from an outpatient clinic, whilst Smith[125] recruited from a primary care community based family practice. All patients in the intervention group had a diagnosis of asthma; however, how asthma was defined was not reported in any of the studies apart from Smith,[125] which used the GINA guidelines. The mean disease duration of asthma was reported in three studies,[122,123,126] and ranged from 14.4 years[122] to 35 years.[123] The mean age in the intervention group ranged from 68[125] to 73[126] years and all studies included both sexes with a slightly higher proportion of females apart from Roh,[124] which had a higher proportion of males. There were only two studies with a comparative control group. Inoue[122] compared FeNO values between elderly asthmatics (≥65 years) and non-elderly asthmatics (<60 years,) and Smith[125] compared asthmatics to non-asthmatics. The device for measuring FeNO was only reported in two studies,[125,126] and both used NIOX MINO.

The results of the included studies are summarised in Table 47. Four studies generally indicated a trend showing that FeNO measurements may not be clinically valuable in elderly asthmatics. Smith 2012[125] found no statistical significant difference (p=0.5) in FeNO levels in elderly asthmatic subjects (20.8±17.3 ppb) compared with elderly non-asthmatics (18.3±9.8 ppb). Furthermore no statistically significant difference was observed in FeNO levels between ICS treated (21.4±20.4 ppb) and untreated asthmatics (19.8±14.3 ppb) p=0.8. Moreover, the definition of asthma used in this study appeared to be vaguely defined; hence non-asthmatics might have been recruited, which may result in a lowermean FeNO value in the asthmatic group. Columbo[123] followed up stable elderly asthmatic patients for one year and evaluated FeNO measurements at baseline and every three months for one year. The results showed that FeNO levels were not elevated at baseline (18.2±14.3 ppb) and did not significantly change during the follow-up. In addition, no significant correlation was found between FeNO levels and spirometric values, ICS usage or asthma control. Two further studies[124,126] showed no correlation between FeNO levels and asthma control. Furthermore, one study, Inoue,[122] which evaluated the pathophysiological characteristics of asthma in elderly patients, found that there was no difference in FeNO levels, percentage of induced sputum eosinophils and neutrophils, or methacholine airway sensitivity or reactivity between elderly asthmatics and non-elderly asthmatics.

Overall, these results should be interpreted with caution as data were derived from studies with lower quality designs which have greater potential to produce biased results. In addition, four of these studies were reported in abstract form only, hence providing limited data. However, all studies appear to indicate that FeNO is not useful in the elderly.

Table 47: Study design, patient characteristics and outcomes of elderly asthmatics

Author, year, source of funding	Study details	Population (No. of patients analysed/recruited, Age, Gender, Disease duration, Medication use)	FeNO measurements (ppb)	Summary of outcomes
Inoue, 2010[122] (abstract) Funding: NR	Observational retrospective study Country: Japan Setting: Outpatient clinic Pathophysiological characteristic of asthma in elderly	Clinically stable asthmatics, never or ex-smokers (<5 pack-years) and receiving ICS therapy for at least 3 months N=136 EA ≥65 yrs: N=49 NEA ≤60 yrs: N=51 Mean age NR Gender (Male): EA: 12/49 (24.5%) NEA: 21/51 (41.2%)	NR	There was no difference in FeNO levels, percentage of induced sputum eosinophils and neutrophils, or methacholine airway sensitivity or reactivity between EA and NEA.
Columbo, 2012[123] (abstract) Funding: NR	Observational study Country: USA Setting : NR Role of FeNO measurements in elderly asthmatics Follow-up for one year with evaluation at baseline and every 3 months	Stable elderly asthmatics N=30 (all asthmatics) Mean age: 71.6±4.9 Gender: Male 12/30 (40%) Atopic: 21/30 (70%) Rhinitis: 27/30 (90%) GORD: 12/30 (40%) Medication usage: ICS: 26/30 (86.7%) 384±378mcg/day LABD: 20/30 (67%)	Baseline 18.2±14.3 ppb Device: NR Flow rate: NR measured in triplicate	FeNO was not elevated at baseline and did not significantly change throughout follow up period. No significant correlation was found between FeNO and FEV_1/FVC (p>0.55, 0.25, 0.10, 0.26 respectively at each time point), other spirometric values, ICS or ACT at any time point. Authors conclusion: In stable elderly asthmatic patients, FeNO was not elevated and did not correlate with subjects' demographics, comorbidities, treatment symptoms or spirometric values. Routine measurements of FeNO may not be

Author, year, source of funding	Study details	Population (No. of patients analysed/recruited, Age, Gender, Disease duration, Medication use)	FeNO measurements (ppb)	Summary of outcomes
		Leukotriene antagonists: 19/30 (63%) Nasal steroids: 15/30 (50%)		clinically valuable in elderly asthmatics
Ross, 2011[126] (abstract) Funding: NR	Cross-sectional study Country: USA Setting: NR Baseline data were collected of objective measures of asthma, including spirometry and FeNO and correlated with asthma QoL and asthma control	Subjects ≥65 yrs with a history of physician-diagnosed asthma N=77 (all asthmatics) Mean age: 73.2 yrs Gender: Male 16/70 (22.9%) Medication usage: NR	NR Device: NIOX MINO Flow rate: NR	No correlation between spirometric/FeNO objective data and QoL or asthma control was noted.
Smith, 2012[125] Funding: Department of Veterans Affairs	Nested case-control study Country: USA Setting: Primary care (community-based family practice) Complete medical history and physical examination was taken. SPT to panel of common aeroallergens, spirometry and measurement of FeNO levels. Geriatric QoL and health statues was assessed through	Asthmatics: Older adults ≥60 yrs with symptoms consistent with asthma. Control: ≥60 yrs age matched, without asthma Asthmatics: 32/36 Control: 39/41 6/77 were unable to perform FeNO measurements Mean age (whole cohort): 68.7 ±7.2 years Gender: Male 18/77 (23.4%)	Asthmatics: 20.8±17.3 ppb Control: 20.8±18.3±9.8 ppb p=0.5 Asthmatics treated with ICS: 21.4±20.4 ppb Asthmatics not treated with ICS: 19.8±14.3 ppb p=0.8 Device: NIOX	No statistical significant difference was seen in FeNO levels between asthmatics and control (p=0.5) No statistical significant difference was seen between ICS treated and not ICS treated asthmatics (p=0.8)

Author, year, source of funding	Study details	Population (No. of patients analysed/recruited, Age, Gender, Disease duration, Medication use)	FeNO measurements (ppb)	Summary of outcomes
	standardised questionnaire	Passive smokers: 24/77 (31.2%) Smoking history: 19±19 pack years Medication usage: Asthma medication: ICS 6/36 (16.7%) Combination inhaler: 12/36 (33.3%) LTMA 8/36 (22.2%) Theophylline 4/36 (11.1%) Albuterol 18/36 (50%)	MINO Flow rate: 50mL/s	

FeNO, fractional exhaled nitric oxide; ICS, inhaled corticosteroids; EA, elderly asthmatics; NEA, not elderly asthmatics; FEV_1, forced expiratory volume in first second; NR, not reported; GORD, gastro-oesophageal reflux disease; ppb, parts per billion; ACT, Asthma Control Test; FVC, forced vital capacity; QoL, quality of life; COPD, chronic obstructive pulmonary disease; GINA, Global Initiative for Asthma; MCT, methacholine challenge test; ppb, parts per billion;

5.2.3.3.3 Adult smokers

Four studies[5,127,127,128] were conducted in adult smokers. A summary of the design and patient characteristics of the four included studies are summarised in Table 48. Two studies (Michils 2009[127] and Schleich 2009[5]) were conducted in Belgium, one (Hromis 2012[129]) in Serbia and one (Kostika 2011[128]) in New Zealand. With the exception of Hromis[129] (which did not provide details of funding source), all studies received funding from one or more commercials sponsors; Kostika 2011[128] received an unrestricted research grant from Aerocrine; Michils 2009[127] received technical funding from Astra-Zeneca; and Schleich 2009[5] received an unrestricted grant from Glaxo SmithKline, Astra-Zeneca and Novartis.

Patients were recruited from a variety of settings and included secondary care,[5] tertiary care[127] and outpatient clinics.[128] Eligibility criteria varied from study to study but all included smokers with confirmed or persistent asthma. The sample sizes of the included studies ranged from 83[129] to 470[127] patients, with the mean age of participants ranging from 38[127] to 50 years.[128] Baseline FeNO was inconsistently reported (Schleich 2009[5] and Kostika 2011[128] utilised median ppb, Hromis 2012[129] reported the mean, and Michils 2009[127] reported the geometric mean and range) and all included patients who were being treated with ICS except in one study[5] in which treatment was unclear. In all studies asthma control was evaluated according to either the GINA guidelines and/or asthma control test and Juniper's asthma control questionnaire with the exception of Hromis 2012,[129] in which it was unclear. The NIOX MINO device was used by Kostika 2011;[128] Schleich 2009[5] used NIOX; Michils 2009[127] used LR 2000, and it was unclear which device was used by Hromis 2012.[129] None of the studies used the same protocol or cut-off points for management of asthma with FeNO.

The results are summarised in Table 48. In Schleich 2009[5] median FeNO levels (17 ppb) in smokers was significantly (p=0.003) lower than in non-smokers (35 ppb). In addition, the FeNO threshold for identifying sputum eosinophil count of ≥3% was significantly (p=0.066) lower in smokers (28 ppb, sensitivity 76% and specificity 62%) compared with non-smokers (46 ppb, sensitivity 58% and specificity 82%).

Michils 2009[127] reported baseline FeNO levels were reduced in smoking asthmatics at 18.1 ppb compared to non-smoking asthmatics at 33.7 ppb. Furthermore, when patients were treated with high to medium doses of ICS, FeNO no longer had the ability to reflect an improvement in asthma control for smoking patients, whereas for non-smoking patients its ability was only slightly reduced. However, the authors suggested that FeNO can reflect asthma control in smoking patients provided changes in FeNO values detected by repeated measurements are considered. A decrease in FeNO of <20% precludes asthma control improvement in non-smoking (NPV=78%) and smoking patients

(NPV = 72%). An increase in FeNO <30% is unlikely to be associated with deterioration in asthma control in both groups (NPV=86% and 84% in non-smoking and smoking patients, respectively).

In Kostika 2011[128] non-smokers who were either treated or not treated with ICS reported higher FeNO values in uncontrolled asthma compared with partly or well controlled asthma. In contrast, smokers not treated with ICS showed significant differences in FeNO values between uncontrolled and well controlled asthma but no difference from partly controlled asthma. Smokers treated with ICS showed no statistically significant difference ($p>0.05$) in FeNO values between the controlled, partly controlled or uncontrolled asthma groups. Diagnostic performance of FeNO for identification of not well controlled (partly or uncontrolled) asthma was better in the non-smoker groups (FeNO cutoff >22 ppb, sensitivity 87%, specificity 81%; FeNO cutoff >27ppb, sensitivity 64%, specificity 94%) compared with smoker groups (FeNO cutoff >19ppb, sensitivity 56%, specificity 75%; FeNO cutoff >23 ppb, sensitivity 45%, specificity 87%)

Hromis 2012[129] showed that FeNO levels are low in asthmatic smokers (17 ppb) compared with asthmatic non-smokers (57 ppb); however, when treated with long acting β agonist a reduction of FeNO was observed in both non-smokers and smokers (32% vs. 22%), respectively. Similar patterns were also observed when both groups were treated with antileukotriens (reduction of 22% in non-smokers and reduction of 12% in smokers). The authors concluded that the sequential changes in FeNO could be a useful marker of asthma control, regardless of smoking status. FeNO level also depends on the applied treatment.

In addition a study by Lehtimaki 2000[130] reported that smoking can cause a small and transient but statistically significant increase in FeNO at 1 and 5 minutes after smoking, highlighting the importance that smokers abstain from smoking before FeNO measurement.

Overall, the findings suggest that FeNO levels in adults tend to be lower in asthmatic smokers compared to asthmatic non-smokers and there is some evidence to suggest that when this group of patients are treated with ICS, FeNO can no longer detect asthma control. However, the use of repeated measures and within patient change from baseline cut-offs may be worthy of further investigation in higher quality studies, with two studies (Michils 2009[127] and Hromis 2012[129]) providing promising data on this approach. However, as no high quality RCT studies have been conducted in this group, the evidence is currently inconclusive as to the effectiveness of using FeNO to guide the management of asthma in smokers.

Table 48: Study design, patient characteristics and outcomes of adult smokers with asthma

Author, year, source of funding	Study details	Population (No. of patients analysed/recruited, Age, Gender, Disease duration, Medication use)	FeNO measurements (ppb)	Outcomes
Michils, 2009[127] Funding: Astra Zeneca	Retrospective, post hoc analysis study Country: Belgium, Setting: Tertiary asthma clinic Evaluation of FeNO to predict asthma control in smoking patients At each visit asthma treatment was adjusted to GINA guidelines - regardless of ACQ score or FeNO values	Adults with Persistent asthma N=470 Current Smokers N=59 Non-smokers N=411 Mean age: Smoker 38±11 yrs Non-smokers 41±16 yrs Gender: Male Smokers 34/59 (58%) Non-smokers 195/411 (47%) Medication usage: ICS dose: Smoker 500 (0-2000) Non-smokers 250 (0-2000)	(GM) Smoker 18.1 (6.9 to 47.5) Non-smokers 33.7 (14.3 to 79.2) Device: LR 2000 chemiluminescence anaslyser Flow rate: 50mL/s.	Levels of FeNO were lower in smoking asthmatics In smoking patients FeNO is unable to cross-sectionally discriminate for or against controlled vs. uncontrolled asthma p=0.39. FeNO cut-off 50 ppb in non-smokers compared to smokers 25 ppb A decrease in FeNO of <20% precludes asthma control improvement in non-smoking (NPV=78%) and smoking patients (NPV = 72%). An increase in FeNO <30% is unlikely to be associated with deterioration in asthma control in both groups. Authors conclusion: Even in smokers, sequential changes in FeNO have a relationship with asthma control. This study indicates that cigarette smoking does not obviate the clinical value of measuring FeNO in asthma among smokers
Schleich, 2009[5] Funding: Supported by Interuniversity Project, GSK, Astra-Zeneca and Novartis	Retrospective, post hoc analysis Country: Belgium Setting: Secondary care FeNO to predict sputum eosinophil count ≥3%	Asthmatic patients N=295 Smokers: N=58 Non-smokers: N=237 Mean age: 47.3 (14-83) Gender: Male 131/295 (78%)	(Median) Smokers 17 (12-37) non-smokers 35 p=0.003 Device: NIOX (Chemiluminescence) Flow rate: 50mL/s.	FeNO levels in smokers was significantly lower than in non-smokers p=0.003 (the median FeNO level was two times higher in non-smokers than current smokers) FeNO level that identified a sputum eosinophil count of ≥3% was lower in smokers than non-smokers (28 ppb vs. 46 ppb) p=0.066 and in non-smokers the sensitivity 58%, and specificity 82%. In smokers sensitivity 76%, specificity 62%.

Author, year, source of funding	Study details	Population (No. of patients analysed/recruited, Age, Gender, Disease duration, Medication use)	FeNO measurements (ppb)	Outcomes
		Medication usage: NR		When combining all variables in the logistic model , FeNO and smoking were independent predictors of sputum eosinophilia Optimum cut-off for FeNO eos >3% ranged from 15 ppb for smoking non-atopic patients receiving a high dose of ICS to 58 ppb for non-smoking atopic patients not treated with a high dose of ICS. When FeNO values were compared to values expected by Dressel equation, the observed values were much higher Authors conclusion: FeNO threshold need to be adjusted for smokers from non-smokers when identifying the presence of sputum eosinophilia in unselected asthma population
Kostika 2011[128] Funding: Aerocrine AB	Cohort study Country: New Zealand Setting: Out-patient clinic Patients sequentially undertook FeNO measurement, EBC collection, spirometry with dry spirometer according to ATS guidelines. Subjects did not smoke 2h before testing	Asthmatic patients N=274 Grp 1: ICS untreated non-smokers N=48 Grp 2: ICS untreated smokers N=32 Grp 3: ICS treated non-smokers N=144 Grp 4: ICS-treated smokers N=50	Grp 1: 30 ppb (18-111) Well controlled 16 ppb (14-21) Partly controlled 40 ppb (27-105)[a] Uncontrolled 116 ppb (63-145) [a, b] Grp 2: 19 ppb (14-22) Well controlled 16 ppb (12-19) Partly controlled 21 ppb (15-38) [a] Uncontrolled 22 (21-108)*	There was statistically significant difference in FeNO values between controlled, partly controlled and uncontrolled asthma ($p<0.05$) in non-smokers (Grp1 and Grp 3). Grp 4 (smokers treated with ICS): no statistically significant difference in FeNO values between the 3 asthma control groups $p>0.05$ Grp 2 (smokers untreated with ICS): there was significant difference in FeNO values between un- controlled and well controlled asthma but no difference from partly controlled asthma. Diagnostic performance of FeNO for identification of not well controlled (partly or uncontrolled) asthma was better in

Author, year, source of funding	Study details	Population (No. of patients analysed/recruited, Age, Gender, Disease duration, Medication use)	FeNO measurements (ppb)	Outcomes
		Mean age: 50±17 yrs Gender: Male 109/274 (40%) Medication usage: ICS 194/274 (71%)	Grp 3: 23 ppb (16-44) Well controlled 16 ppb (12-20) Partly controlled 28 ppb (20-44) [a] Uncontrolled 61 ppb (35-78) [a,b] Grp 4: 19 ppb (14-25) Well controlled 17 ppb (14-22) Partly controlled 19 ppb (13-25) Uncontrolled 23 ppb (17-74) Device: NIOX MINO Flow rate: 50 mL/s	the non-smoker groups (Grp 1 FeNO cut-off >22 ppb, sensitivity 87%, specificity 81%; Grp 3 FeNo cut-off >27ppb, sensitivity 64%, specificity 94%) compared to smoker groups (Grp 2 FeNO cut-off >19ppb, sensitivity 56%, specificity 75%; Grp 4 FeNO cut-off >23 ppb, sensitivity 45%, specificity 87%)
Hromis, 2012[129] (abstract) Funding: NR	Randomised, open label study Country: Serbia Setting: NR FeNO in assessing asthma control in smoking patients In the next 4 weeks from start of study , LABA (salmeterol, 30 non-smokers and 17 smokers) or antileucotriens (montelucast	Mild to moderate asthmatics with poor control N=83 Smokers: N=31 Non-smokers: N=52 Age: NR Gender: NR	Smokers: Baseline 57 ppb After 4 weeks treatment: 17 ppb LABA added to ICS Reduction of FeNO Smokers: 22% Non-smokers: 32% Treated with antileucotriens (reduction of FeNO)	The sequential changes in FeNO could be a useful marker of asthma control, regardless of smoking status. FeNO level also depends on the applied treatment

Author, year, source of funding	Study details	Population (No. of patients analysed/recruited, Age, Gender, Disease duration, Medication use)	FeNO measurements (ppb)	Outcomes
	10 mg, 17 non-samokers and 14 smokers)was added Changes in FeNO, FEV_1 and ACT at baseline and after 4 weeks of treatment were measured	Medication Usage: Low dose of ICS (400 µg beclomethasone dipropionate daily or equivalent)	Smokers: 12% Non-smokers: 22% Device: NR Flow rate: NR	

[a] p<0.05 compared to controlled asthma;
[b]p<0.05 compared to partly controlled asthma;
FeNO, fractional exhaled nitric oxide; ACQ, Asthma Control Questionnaire; ACT, Asthma Control Test score; GINA, global initiative for asthma guidelines; GM, geometrical mean; ICS, inhaled corticosteroids; LABA, long acting β agonist; NPV, negative predictive value; QoL, quality of life ; GM, geometric mean; NPV, negative predictive value; GSK, GlaxoSmithKlein; ppb, parts per billion; American Thoracic Society; h, hour; EBC, exhaled breath condensate; Grp, group; LABA, long acting ß2 agonist; NR, not reported.

5.2.3.3.4 Children exposed to tobacco smoke

A summary of the study design and patient characteristics of the three included studies (Mahut[102], De La Riva-Velasco 2012[104] and Hanson 2012[103]) are summarised in Table 49. All studies were conducted in the USA, except for Mahut[102] which was conducted in France. None of the studies appeared to receive funding from commercial sponsors. All the studies were observational studies. In De La Riva-Velasco 2012[104] a cohort study, the authors determined the relationship between FeNO levels and exposure to low levels of environmental tobacco smoke in children with asthma on ICS treatment. Hanson 2012[103] was a retrospective chart review study that looked at the relationship of FeNO with multiple clinical variables in children aged 4-7 years. Mahut 2011[102] was a cross-sectional study. This study evaluated whether exhaled FeNO was independently associated with underlying pathophysiological characteristics of asthma (e.g. airway tone and airway inflammation) and with clinical phenotypes of asthma. In two studies (De La Riva-Velasco 2012[104] and Mahut 2011[102]) patients were recruited from outpatient clinics and in Hanson 2012[103] from an Asthma Allergy clinic. Eligibility criteria varied from study to study but all studies included children who had been exposed to tobacco smoke and had a diagnosis of asthma. The sample sizes of the included studies ranged from 33[104] to 169[102] patients, with the mean age of participants ranging from 10[104] to 10.5 years.[102] However, Hanson 2012[103] did not report the mean age but recruited children between the age of 4 and 7 years. Baseline FeNO was inconsistently reported (Mahut 2011 [102] and De La Riva-Velasco 2012[104] utilised median and Hanson 2012[103] reported mean FeNO levels) and the studies also varied in terms of medication usage. De La Riva-Velasco 2012[104] included children who were on treatment with ICS[104], whilst Mahut 2011[102] included patients who were being treated with either ICS, LABA and beta agonist on demand. Hanson 2012[103] failed to provide details on medication usage. A range of devices were used to measure FeNO levels. Hanson 2012[103] used NIOX MINO; Mahut 2011[102] used NIOX; and De La Riva-Velasco 2012[104] used NIOX Flex. None of the studies used the same protocol or cut-off points for management of asthma with FeNO.

The results of the included studies were varied. Mahut 2011[102] reported that there was no statistically significant relationship between FeNO and smoke exposure, and concluded that FeNO is potentially helpful in asthma management. On the other hand, De La Riva-Velasco 2012[104] found that children on low to medium doses of ICS with recent low level environmental tobacco smoke exposure have lower median FeNO levels (9.6 IQR 5.1-15.8) when compared with subjects not exposed to environmental tobacco smoke (23.9 IQR 15.2-34.5) p=0.008. The authors concluded that environmental tobacco smoke exposure or third-hand smoke (that which lingers after a cigarette is extinguished) may be an important variable to consider when interpreting FeNO levels in school-aged children with asthma.

Overall, the findings suggest that the potential efficacy of using FeNO to inform the management of asthma in children exposed to environmental tobacco smoke may be similar to that in children who have not been exposed, but that it may be necessary to consider a child's exposure status when interpreting results, as FeNO may be lower in these children. Whether this should involve the setting of lower cut-off points or whether a more qualitative interpretation on a case-by-case basis or by comparing within-patient changes from baseline is unclear.

Table 49: Study design, patient characteristics and outcomes of asthmatic children exposed to smoke

Author, year, source of funding	Study details	Population (No. of patients analysed/recruited, Age, Gender, Disease duration, Medication use)	FeNO measurements (ppb)	Outcomes
Mahut, 2011[102] Funding: NR (authors declared no competing interest)	Cross-sectional study, single centre cohort Country: France Setting: Outpatient clinic To evaluate if FeNO is associated with a phenotype of childhood asthma (exposure to tobacco)	Asthmatic children N=169 Cluster 1 (asthmatic boys, well-controlled asthma, unexposed to tobacco): N=79 Cluster 2 (asthmatic girls, well-controlled asthma, unexposed to tobacco): N=44 Cluster 3 (asthmatic boys or girls, un-controlled asthma associated with increase airways tone, unexposed to tobacco): N=11 Cluster 4 (asthmatic boys or girls, un-controlled asthma associated with small airway to lung size ratio, exposed to parental smoking): N= 35 Mean Age: 10.5 ±2.6 yrs Medication usage: ICS 87/169 (51%) LABA 73/169 (43%) Beta-agonist on demand 82/169 (82%)	(Median) Cluster 1: 25[14-45] Cluster 2: 34[19-51] Cluster 3: 21[9-49] Cluster 4: 30[14-52] p=0.58 Device: NIOX Flow rate: 50mL/s	FeNO was not decreased by tobacco exposure in univariate analysis FeNO was not different in the four clusters p=0.58
De La Riva-Velasco, 2012[104] Funding: Supported by Children's Environmental Health	Cohort study Country: USA Setting: Outpatient clinic	Clinically stable, mild or moderate persistent asthma taking low or medium dose of ICS daily N=33	(Median) ETS exposed: 9.6 ppb IQR(5.1-15.8) Non-ETS exposed: 23.9 ppb IQR (15.2-34.5)	Children on low to medium doses of ICS with recent low level ETS exposure have lower FeNO levels when compared with non-ETS-exposed subjects. Low level ETS exposure or third hand smoke may

Author, year, source of funding	Study details	Population (No. of patients analysed/recruited, Age, Gender, Disease duration, Medication use)	FeNO measurements (ppb)	Outcomes
Centre of the Hudson Valley and the Maria Fareri Children's Hospital Foundation (authors declared no competing interest)	Relationship between FeNO and exposure to low level environmental tobacco smoke in children with asthma on ICS	Subjects stratified based on urine cotinine levels: ≥1 ng/ml (ETS exposed) N=10 <1 ng/ml (non- ETS exposed) N=23 Median age: ETS exposed 10.5 (9.5-10.9) Non-ETS exposed 10.0 (9-11) Gender: ETS exposed 4/10 (40%) Non-ETS exposed 12/23 (52%) Medication usage: N=26 daily medium-dose of ICS N=7 daily low dose ICS	p=0.008 Device: NIOX Flex system Flow rate: 50mL/s	be an important variable to consider when interpreting FeNO levels in school-aged children with asthma
Hanson, 2012[103] (abstract) Funding: NR	Retrospective chart review Country: USA Setting: Asthma/Allergy clinic Characterisation of FeNO in children aged 4-7 yrs and analyses its	Children age 4-7 years who underwent FeNO testing N=80 Mean age: NR Gender: NR Medical usage: NR	(Mean) 18.7 ppb (2.5-89) Device: NIOX MINO Flow rate: NR	There was no statistically significant relationship between FeNO and smoke exposure.

Author, year, source of funding	Study details	Population (No. of patients analysed/recruited, Age, Gender, Disease duration, Medication use)	FeNO measurements (ppb)	Outcomes
	relationship with multiple clinical variables.			

FeNO, fractional exhaled nitric oxide; FEV_1, forced expiratory volume in first second; ATS, American thoracic Society; ERS, European Respiratory Society; ACQ, Asthma Control Questionnaire; ETS, exposed to tobacco smoke; GINA, global initiative for asthma guidelines; ICS, inhaled corticosteroids; LABA, long acting β agonist

5.3 Discussion of clinical evidence

5.3.1 Summary of key results

5.3.1.1 Equivalence of devices

The review of the equivalence of devices revealed that the level of agreement between devices is highly variable.

NIOX MINO versus Niox: The most evidence is available for a comparison of Niox to NIOX MINO in adults. Devices give similar mean FeNO values in five of eight studies, but higher values in three studies. Bland-Altman analysis (reported in four of eight studies) suggested that the limits of agreement were around 10ppb in both directions when analysed on the absolute scale, with mean differences of between 0.5 and 1.5ppb. Analysis on the log scale produced better limits of agreement in Menzies 2007[38] but very wide limits in Korn 2010.[40] There was also evidence that not all NIOX MINO devices produce equivalent readings to one another, in a head to head comparison between two NIOX MINO devices,[39] though another study comparing three devices found them to be equivalent to one another. [44] There was also evidence that agreement between NIOX MINO and niox is worse at higher FeNO values, with all studies where cohorts had mean values below 30ppb reporting better agreement and most Bland-Altman plots showing a multiplicative relationship. Agreement looked acceptable in children, with all limits of agreement falling between -4.4 and 8.9.

NIOX MINO versus other chemiluminescent devices: correlation co-efficients were generally good showing correlation between 0.76 and 0.96. However, cohort means were far more variable with some devices reading higher and some lower than NIOX MINO. The highest difference between cohort mean FeNO values was 47ppb. Whilst individual devices may show good agreement with NIOX MINO, it is not possible to draw any solid conclusions as most devices were only tested in one or two studies, and as was seen in the comparison between NIOX MINO and Niox, results between studies appear to vary considerably. NIOX MINO

NObreath versus Niox and other chemiluminescent devices: In the one study that compared NObreath to NIOX MINO, a good level of agreement was seen in Bland-Altman analysis, but cut-off values derived by this study for diagnosis of asthma differed by 10ppb according to which device was used.[67] Other devices generally appeared to read higher than NObreath, but not by a consistent amount.

NObreath versus NIOX MINO: Both studies reporting this comparison[53,65,68] found that in most analyses NIOX MINO provided lower mean FeNO values than NObreath. This contradicts the available evidence for comparisons of NIOX MINO to Niox and NObreath to Niox, where NIOX MINO>Niox>NObreath.

NIOX VERO: Only one study provides data on this device. [redacted]

Test failure rates: The overall test failure rate for FeNO measurement in adults was generally low across all devices, and most patients appear to be able to provide FeNO readings, provided they are permitted sufficient measurement attempts. There may be a higher test failure rate in children using NIOX MINO.

Conclusion: Overall, it cannot be concluded that any two devices are equivalent in all situations. Whilst there may be situations where they are similar, it appears to depend on the characteristics of the studies and cannot be generalised to all situations. Further research is required to identify what is driving the variability between studies and devices. However, as there is mostly a high degree of correlation between measurements across all devices, estimates of sensitivity and specificity are likely to be an accurate indication of potential diagnostic accuracy of using FeNO to guide diagnosis, but the derived cut-off points are not likely to be interchangeable between devices.

5.3.1.2 Diagnostic accuracy of FeNO for the diagnosis of asthma

No end-to-end studies were identified, and no cohort study provided a comparison of using FeNO within a sequence of tests versus a suitable reference standard of the same sequence of tests without FeNO. The review included twenty-five studies estimating the diagnostic accuracy of either FeNO alone or FeNO in conjunction with another test versus a variety of reference standards and in a variety of populations.

Adults presenting with symptoms of asthma (4 studies): It is difficult to draw any conclusion as to the optimal cut-off for sensitivity and specificity due to the heterogeneity between results, study designs and devices used.

- Cut-off for the highest sum of sensitivity and specificity ranged from 20ppb to 47ppb amongst these studies, and even results produced by the same authors in studies with high levels of heterogeneity varied widely (25ppb and 46ppb in Schneider 2013[75] and Schneider 2009[77,78]). Sensitivities ranged form 32% to 88%, and specificities from 75% to 93%.

A range of cut-offs was not reported in all studies, and it was not clear if the highest sensitivity or specificity values were available. From those that were reported:

- When selecting the cut-off with the highest sensitivity, these ranged from 9 ppb to 15 ppb, sensitivity from 85% to 96% and specificity from 13% to 48%.

- When selecting the cut-off with the highest specificity, these ranged from 47 ppb to 76ppb, with sensitivities from 55.6% to 13% and specificities from 88.2% to 100%.

The consistently smaller range and higher values of specificities than sensitivities reported suggest that FeNO may be a more reliable and useful parameter to base diagnostic decisions on as a rule-in test than as a rule-out test. However, this balance will depend on the clinical and cost consequences of being TP, TN, FP and FN in each scenario.

Subset of patients at Position A versus airway reversibility or airway hyper-responsiveness (2 studies): These studies do not produce estimates of diagnostic accuracy of noticeable difference to the studies which recruited a potentially broader spectrum of patients:

- Cut-off for the highest sum of sensitivity and specificity ranged from 27 to 36ppb, sensitivities from 77.8% to 87%, specificities from 60% to 92%

A range of cut-offs was only reported in one study, Heffler 2006.[86] These reached 100% sensitivity and specificity at the highest and lowest cut-offs in this cohort versus this reference standard.

- When selecting the cut-off with the highest sensitivity, this was 25ppb, sensitivity 100%, specificity 46.7%
- When selecting the cut-off with the highest specificity, this was 100ppb , sensitivity 27.8%, specificity 100%

Difficult to diagnose patients (3 studies): these studies all used some form of airway hyper-responsiveness as the reference standard. Surprisingly, estimates of sensitivity and specificity seemed largely comparable to the studies recruiting patients presenting to primary care with symptoms of asthma versus airway reversibility, ICS responsiveness and airway hyper-responsiveness. Bobolea 2012[92] recruited a set of patients who were negative by MCT and compared FeNO to an adenosine challenge test. This study produced 100% sensitivity (29.2% specificity) at a cut-off of 30ppb, making it likely to operate well as a rule-out test.

The other studies used MCT challenge in patients who were negative for asthma in previous tests.

- Cut-offs for the highest sum of sensitivity and specificity ranged from 34ppb to 40ppb amongst the studies versus MCT, which is a little narrower than in the broader cohorts. Sensitivities ranged from 24.4% to 74.3%, and specificities from 72.5% to 98.9%, which is a similar range to the broader cohort. This perhaps reflects the fact that airway reversibility is a highly accurate test for asthma, and that the combined tests behave in a similar manner. If this

is the case, it would also suggest that FeNO has similar diagnostic properties in difficult to diagnose patients as in the broader spectrum of patients.

A range of cut-offs was not reported in these studies.

Patients with chronic cough, difficult to diagnose (3 studies): These studies recruited patients with chronic cough who had tested negative for other causes. All three studies used a reference standard of ICS responsiveness. Cut-offs for the highest sum of sensitivity and specificity were also in the same range, and sensitivity and specificities were somewhat better in two studies at 94.7% sensitivity with 76.3% specificity in Hsu 2013,[79] and 90% sensitivity with 85% specificity in Hahn 2007.[80] This is in accordance with the expectation that FeNO is a better marker of ICS responsiveness than of asthma itself. At this position in the pathway, FeNO may be a useful test to perform before ICS responsiveness to indicate which patients are likely to respond to a trial of treatment. Patients with low FeNO could go on for further asthma investigations (eg methacholine challenge), or be assumed to be non-asthmatic depending on whether a rule-in or rule-out scenario is used.

Children with symptoms of asthma (4 studies): In comparison to the adult cohorts with a similar spectrum of patients and reference standards, the cut-offs derived are generally lower, but with similar ranges of estimates of sensitivity and specificity.

- There was a high degree of agreement as to the cut-off which produces the highest sum of sensitivity and specificity, despite the heterogeneity in devices and reference standards, with values between 19 and 21pbb, which more consistently lower than in adults. Estimates of sensitivity at these cut-off points were also wide ranging and of a similar range of values as in adult studies at 49% to 86% compared to 32% to 88% in adults. Again as in adults, specificity was more similar between studies ranging from 76% to 89%, and of a similar range to adults (75% to 93%)
- When selecting the cut-off with the highest sensitivity, results were similar to adult cohorts. Cut-offs ranged from 5 to 20ppb (versus 9 ppb to 15 ppb in adults), sensitivity from 89% to 94% (versus 85% to 96% in adults), specificity from 14.1% to 70% (versus 13% to 48%).
- When selecting the cut-off with the highest specificity, results were also similar to adult cohorts. Cut-offs were perhaps a little lower again, and ranged from 30 to 50ppb (versus 47 to 76 ppb in adults), sensitivity ranged from 20% to 50% (versus 13% to 55.6% in adults, specificity from 92% to 100% (versus 88.2% to 100% in adults).

Adult smokers: One study was identified. Malinovschi 2012[101] recruited an unusual spectrum of patients via a random selection of the population of adults, who reported symptoms of asthma, rather

than a population presenting with symptoms, and the reference standard was poor. For this reason, a comparison of cut-off values and sensitivity and specificity estimates with the studies already discussed is not appropriate.

This study compared diagnostic accuracy of FeNO in smokers, non-smokers and ex-smokers. From cohort mean values, it would seem likely that FeNO is generally lower in smokers, and it may be useful to consider a patient's smoking status when interpreting results, or to select lower cut-off points for smokers. It is difficult to determine how the fairly minor differences in cut-off points and diagnostic properties of FeNO across groups would affect cost-effectiveness and clinical utility in practice. However, it would appear that FeNO is able to distinguish between asthmatics and non-asthmatics in smokers with similar accuracy to non-smokers and ex-smokers.

Children exposed to tobacco smoke: Evidence was limited and drew on studies reported in the section relating to management of children exposed to tobacco smoke and from the above study reported for adults. The overall conclusion was the same as for adults smokers: it may be necessary to consider a child's exposure status when interpreting results of FeNO for the diagnosis of asthma, as FeNO may be lower in children exposed to tobacco smoke.

Pregnant women: No diagnostic accuracy studies in pregnant women were identified. A cross-sectional study compared mean FeNO values in pregnant asthmatics, non-pregnant asthmatics, healthy pregnant women and healthy non-pregnant women. The study concluded that pregnancy does not alter FeNO levels in asthmatics or non-asthmatics, and that FeNO can distinguish between asthmatic and non-asthmatic pregnant and healthy women. However, it is unclear whether diagnostic accuracy would be equivalent to that reported in other studies in non-pregnant people or a mix of pregnant and non-pregnant people.

The elderly: No diagnostic accuracy studies in the elderly were identified. A case-control study (Simpson[106]) investigated FeNO levels in elderly patients with eosinophilic airflow obstruction (sputum cell count >3%) versus elderly healthy controls. No significant difference was found in the mean FeNO values, suggesting FeNO is not a good marker of eosinophilic airway inflammation in elderly patients. This indicates that FeNO is unlikely to act as a useful test in the diagnosis of asthma.

5.3.1.2 FeNO-guided management in asthma

5.3.1.2.1 Adults (4 studies)

There was a high degree of heterogeneity in all aspects of study design across studies, including levels of blinding, inclusion criteria, visit frequency, cut-off points selected, devices used, step-up step-down

protocols and medications controlled by the protocols. Only one study reported using UK guidelines in the comparator arm (Shaw 2007).[108]

Exacerbations:

All studies reported a fall in exacerbation rates per person year, though it appeared that this was mostly driven by mild and moderate exacerbations.

Severe exacerbations:

- Syk unpublished[109] reported higher rates of oral corticosteroid use in the intervention arm, whilst the composite outcome of moderate or severe exacerbations favoured the intervention arm.
- In other studies, the difference between the outcome OCS use and the composite outcomes which include less severe exacerbations was less pronounced. OCS use and the composite outcomes of severe and less severe exacerbations fell in intervention arms, though there was still an apparently greater effect in the composite outcomes. Reviewer-calculated rate ratios for major/sever exacerbations ranged from 0.79 (95% CI 0.44 to 1.41) to 1.29 (95% CI 0.51 to 3.30) whilst reviewer-caluclated rate ratios for composite outcomes of all severity of exacerbation ranged from 0.52 (95% CI 0.30 to 0.91) to 0.63 (95% CI 0.40 to 0.98).
- Exploratory pooled analyses showed that FeNO guided management may not be effective at reducing severe exacerbations requiring a course of OCS with the pooled risk ratio (in a sensitivity analysis removing studies with wider definitions of severe exacerbations) of 0.90 (95% CI 0.56 to 1.45). However, there were only two studies in this analysis, and one, Smith 2005,[107] showed a trend towards a reduction in OCS use whilst Syk 2013[109] conversely showed a trend towards an increase in OCS use, perhaps suggesting that differences in study characteristics, step-up/step-down protocols and patients may account for differences in direction of effect. When considering all definitions of sever exacerbation, the pooled estimate of all four studies indicated efficacy at 0.87 (95% CI 0.64 to 1.19).

All definitions of exacerbations:

- Three studies reported a composite outcome including all types of exacerbation and all showed statistically significant reviewer-calcualted rate ratios in favour of the intervention arm, but only non-significant trends in the primary analyses performed by the study authors. That is, all except Syk 2013,[109] where there were 0.22 exacerbations per person year in the intervention arm, but 0.41 in the control arms (p= 0.024).
- An exploratory pooled analysis showed a statistically significant effect with a rate ratio of 0.58 (95% CI 0.43 to 0.77).

- Heterogeneity in study results may be due to one of the confounding factors; for example step-up step-down protocols may be driving the estimates of efficacy rather than the use of FeNO.

ICS use:

All studies reported some data on ICS use. Smith 2005[107] and Shaw 2007[108] reported ICS use as a mean per day at the end of the study with mean differences of -270µg per day (95% CI -112 to -430, p=0.003) and -338µg per day (95% CI -640 to -37µg, p= 0.028) respectively, in favour of FeNO-guided management. Syk 2013[109] showed a small increase in ICS use in the intervention arm (586µg (SE 454) versus 540µg (SE 317) in the control arm. Calhoun 2012[110] reported mean per month, though it is unclear if this was an average over the whole course of the study, or the means for the final month of the study. The means were very similar at 1617µg/month and Mean 1610µg/month.

- An exploratory meta-analysis using standardised mean difference as outcome reporting was not standardised showed an overall effect of -0.24 standard deviations in favour of the intervention, though this narrowly missed significance (95% CI -0.56 to 0.07, p=0.13).
- This may indicate that some step-up step-down protocols were better at decreasing ICS use than others, or may be due to the characteristics of the study populations.

HRQoL:

Two studies used versions of the AQLQ to measure quality of life. Both showed no effect in the global score, but one investigated domains and found a statistically significant difference in the symptoms score.

Asthma control and other medication use:

- Asthma control either did not change (Smith 2005, [107]Calhoun 2012[110] and Shaw 2007[108]) or increased (Syk unpubished).[109]
- Smith 2005[107] and Calhoun 2012[110] reported no significant difference between groups for bronchodilator use, and Syk[109] reported non-significant trends towards greater numbers and mean use of LTRA and SABA (significance not reported) in the FeNO controlled arm.

Adverse events and mortality:

No asthma-related adverse events were reported. No deaths were reported.

Conclusions for FeNO-guided management in adults:

Due to the high levels of heterogeneity in multiple study characteristics and outcome definitions, it was not possible to draw any firm conclusion as to which step-up/step-down protocol or cut-off points

offer the best efficacy. However, it would seem possible to conclude that FeNO-guided management of most descriptions is likely, during the first year of management, to result in non-significant trend towards better management overall with no or a small reduction in ICS use. The least favourable results were reported in Syk unpubished,[109] but the explanation for this is not clear.

5.3.1.2.2 Children (4 studies)

There was a high degree of heterogeneity in all aspects of study design across studies, including levels of blinding, inclusion criteria, visit frequency, cut-off points selected, devices used, step-up step-down protocols and medications controlled by the protocols. No study reported using UK guidelines in the comparator arm.

- Two studies recruited patients who appeared to be poorly controlled; Szefler 2008[113] and Petsky 2010.[116]
- It was not possible to tell whether patients in Verini 2010[114] were controlled or uncontrolled
- One study recruited patients who were mild to moderate persistent asthmatics; Fritsch 2006[112]
- One study recruited patients who had a stable dose of ICs for the previous 3 months, suggesting they were reasonably well controlled; Pijnenburg 2005.[115]

Severe exacerbations:

Two studies reported this outcome in way that allowed calculation of rates per person year, both with lower rates in the intervention arm

- Szefler 2008[113] in uncontrolled asthmatics; rates 0.746 and in the control arm 0.950; and
- Pijnenburg 2005[115] in patients who had been on a stable dose of ICS for three months; rates 0.21 and 0.39 respectively.

Both were calculated by the reviewer and the statistical significance is unclear. One further study, Fritsch 2006,[112] reported this outcome as the number of people, and this was 2 in each arm.

All definitions of exacerbations:

Fritsch 2006,[112] Szefler 2008[113] Verini 2010[114] and Petsky 2010[116] all reported outcomes that were not defined as either major or minor and had different definitions to one another. All studies showed at least a trend in favour of fewer exacerbations in the intervention arm. The only study to report a significant between-group difference was the conference abstract Petsky 2010.[116] Exacerbations were not clearly defined, but occurred in 6/31 participants in the control group (19.4%), and 15/32 in the control group (46.9%, p=0.021). Pijnenburgh 2005[115] did not report this outcome.

ICS use:

With the exception of Petsky 2010[116] all the studies provided some indication of between-group differences in ICS use. Fritsch 2006[112] and Szefler 2008[113] reported statistically significantly higher ICS use in the FeNO group, Pijnenburg 2005[115] reported very similar levels, while the values in the remaining study (in terms of absolute n using ICS) were difficult to interpret. The differences in the effects on ICS useage between studies may be due to the specifics of the step-up/step-down protocols and/or the characteristics of the patients selected. In the case of Pijnenburg 2005[115] where patients may have been generally better controlled at the outset, step-down of ICS may have been more likely than in Szefler 2008[113] where patients were poorly controlled and ICS was perhaps more likely to be stepped up. Having said this, poorly controlled patients may be poorly controlled due to being non-reponsive to corticosteroids, and use of FeNO may actually result in a decrease in ICS use, if low FeNO always indicates a decrease in treatment. However, in the case of Szefler 2008[113] ICS medication was not reduced solely on the basis of low FeNO if symptoms were not reduced, so this type of step-down is not likely. As such, it may be that the increase in ICS usage in Szefler 2008[113] is a function of the population selected as well as the management protocol.

HRQoL:

Only reported in one study in abstract form (Petsky 2010[116]), and using an unknown tool. No conclusions can be confidently drawn from this data.

Asthma control and other medication use:

Four studies provided some data on asthma control, none of which demonstrated any statistically significant effects favouring either intervention or control. With respect to additional medication use, three studies provided data, and there did not appear to be a clear direction of effect within the data.

Adverse events, mortality, compliance and test failure rates

Szefler 2008[113] reported no difference in adverse events between groups and no mortality was observed. The adverse events listed were: Eyes, ears, nose and throat; gastrointestinal disorders; haematology disorders; infections; musculoskeletal symptoms and skin symptoms.

Conclusions for FeNO-guided management in children:

Due to the high levels of heterogeneity in multiple study characteristics and outcome definitions, it was not possible to draw any firm conclusion as to which step-up/step-down protocol or cut-off points offer the best efficacy. However, it would seem possible to conclude that FeNO-guided management of most descriptions is likely, during the first year of management, to result in non-statistically significant trends towards better management overall. It is unclear whether ICS use is likely to

increase or decrease, and this may depend on the details of the step-up step-down protocols or the characteristics of the patients recruited to the trials in terms of control and severity.

5.3.2. Generalisability of results to UK practice

5.3.2.1 Diagnostic review

5.3.2.1.1 Adults

Only studies with some relevance to UK practice were considered. Of these, not all used NIOX MINO or NObreath, and the studies with the highest relevance to UK practice can be broken down into four types:

- **Studies of all patients presenting with symptoms of asthma versus a reference standard that includes the most common tests in the UK pathway**
    - The most relevant in this category are: Schneider 2013,[75] and Schneider 2009,[77,78] which used NIOX MINO and Smith 2005,[87] and Smith 2004[90] which used Niox and an unknown device respectively.
- **Studies recruiting patients who are difficult to diagnose versus test appropriate to the UK practice.**

  These include:
    - Schleich, 2012,[83] which used Niox and Pedrosa 2010[89] which used NIOX MINO
    - Bobolea 2012, [92] which also used NIOX MINO and selected a population of patients who were negative by MCT. The reference standard was adenosine challenge test.
- **Studies recruiting patients with chronic cough who have already undergone other tests.**
    - Hsu 2013[79] Hahn 2007[80] and Prieto 2009.[82]These are useful to demonstrate that FeNO can predict ICS responsiveness in these patients, rather than a diagnosis of asthma. None used NIOX MINO.
- **Studies using FeNO in conjunction with another test versus an appropriate reference standard**
    - Schleich, 2012[83] recruited a difficult to diagnose group, combined FeNO with FEV_1%<101% and used a reference standard of MCT. It could be argued, however, that in the UK only patients negative by FEV_1% would receive a MCT, in which case this combination may not be a useful one.
    - Cordeiro 2011[91] recruited patients presenting with symptoms of asthma and combined FeNO with airway reversibility to administration of bronchodilator. This would be equivalent to FeNO being used to prevent MCT, but some of the included patients would not have received MCT under the UK pathway anyway. As such, it is again unclear how useful this combination of tests would be in UK practice.

5.3.2.1.2 Children

Only studies with some relevance to UK practice were considered. Of these, the studies with the highest relevance to UK practice can be broken down into:

- **Studies of all patients presenting with symptoms of asthma versus a reference standard that includes the most common tests in the UK pathway**
    - Woo 2012[100] which recruited patients in Position A on the pathway and used the NIOX MINO device versus a reference standard that roughly equates to UK practice.
    - Linkosalo 2012[97], which used a Sievers NOA 280 chemiluminescence device for patients in Position A on the pathway with a reference standard of exercise challenge test. Not all presenting patients would receive this test in UK practice
        - FeNO would be positioned before exercise challenge test and could triage patients away from this.
    - Sivan 2009, which used an EcoMedics device in patients at Position A in the pathway versus a reference standard similar to UK practice
        - In this study, FeNO replaces the whole pathway prior to ICS use

5.3.2.1 Management review

5.3.2.1.1 Adults

Generalisability to UK practice is clear-cut in adults as Shaw 2007[108] used UK guidelines as the comparator and was based in the UK. Patients were recruited from primary care, and included mild to severe asthmatics (unless a severe exacerbation had been experienced in the previous 4 months) and atopic patients as well as non-atopic patients. Smokers were excluded, so results may not be generalizable to this group. However, this study offers the best generalisability to UK practice in terms of setting, population and comparator.

However, if management protocols different to that used in Shaw 2007[108] were to be considered for recommendation, other studies may offer some useful data. Input from a clinician (personal communication from Professor Ian Pavord, August 2013) suggests that the management protocol described by Powell 2011,[111] where symptoms control LABA dose and FeNO levels control ICS dose is generally thought to be the best design. This study was conducted in pregnant women only, and its generalisability to the whole asthma population is not assured. The protocol described in Shaw 2007[108] appears to be similar to this, in that FeNO controls ICS and leukotriene receptor antagonist dose, whilst symptoms scored according to the Juniper scale control SABA, LABA and theophylline doses. Importantly, this allows for low FeNO levels to result in a reduction in ICS dose regardless of symptoms

5.3.2.1.2 Children

Generalisability in studies recruiting patients was less clear cut. None were set in the UK and none used UK guidelines in the comparator group. On the basis of reported quality, Szefler 2008[113] was at lowest risk of bias, and for patients who are uncontrolled, this may be the best study to base generalisations on. However, clinical input to the project (personal communication with Professor Ian Pavord, August 2013) indicated that Szefler 2008[113] has been criticised for not allowing step-down of ICS on the basis of low FeNO levels if symptoms are still present. In addition, the patient population in this study was patients who were uncontrolled, which may introduce bias in that patients will be less likely to be indicated for step-down of ICS.

For patients who are mild to moderate asthmatics, Fritsch 2006[112] may be the best study to select as Pijnenburg 2005[115] uses only symptoms to guide asthma management in the control arm, whilst Fritsch 2006[112] uses symptoms, SABA use and lung function tests, which is probably more comparable with UK practice. However, Fritsch 2006[112] used only FeNO to guide management and it would seem more likely that clinicians would use other measures such as symptom control and lung function to guide treatment. This would allow the stepping down of treatment based on FeNO values, but may also be less sensitive than using a combination of factors. Unfortunately, there is not a study in Children that addresses this particular problem by combining FeNO with other indicators in a protocol that allows step-down in the presence of low FeNO regardless of symptomatic control. Fritsch 2006[112] did not report data in a way that allowed calculation of rates of exacerbations per person year, and is of limited use for this reason. Pijnenburg 2005[115] provides the necessary data and was selected for modelling.

6. THE COST-EFFECTIVENESS OF FENO TESTING FOR THE DIAGNOSIS AND MANAGEMENT OF ASTHMA

6.1 Introduction

This chapter presents an assessment of the cost-effectiveness of FeNO testing for the diagnosis and management of asthma. The chapter is comprised of two main sections: (i) a review of existing evidence relating to the cost-effectiveness of FeNO testing in the diagnosis and management of asthma, and (ii) an exposition of the methods and results of two *de novo* health economic models developed by the External Assessment Group (EAG) to evaluate the cost-effectiveness of FeNO testing for the diagnosis and management of asthma.

The chapter is set out as follows. Section 6.2 sets out the aims and objectives of the economic analysis. Section 6.3 presents the methods and results of a critical review of existing economic analysis of FeNO testing for asthma; this includes a critical review of published studies as well and other economic evidence submitted by the manufacturers of the FeNO devices considered in this assessment. The section also includes a summary of methodological problems and concerns associated with undertaking economic analyses of interventions for the diagnosis and management of asthma. Section 6.4 presents the methods and results of the *de novo* economic analyses undertaken by the EAG. Sections 6.5 and 6.6 summarises the main findings of the health economic analyses of FeNO testing and highlights the key uncertainties surrounding the evidence base used to inform the *de novo* analysis.

6.2 Aims and objectives of the health economic assessment of FeNO

The purpose of this chapter is to assess the expected cost-effectiveness of NIOX MINO, NIOX VERO and NObreath versus current standard care for the diagnosis and management of asthma. Importantly, there is uncertainty not only with respect to whether FeNO testing might represent a cost-effective use of health care resources, but also with respect to how FeNO should be used in the most cost-effective manner within existing asthma pathways. Thus, the economic analysis attempts to address the following questions:

1. What is the cost-effectiveness of FeNO testing using NIOX MINO, NIOX VERO and NObreath as compared against current standard tests for the diagnosis of asthma in England and Wales?
    a. Should FeNO testing be used *alongside* existing standard tests for the diagnosis of asthma?
    b. Should FeNO testing be used *in place of* existing standard tests for the diagnosis of asthma?

2. What is the cost-effectiveness of FeNO testing using NIOX MINO, NIOX VERO and NObreath as compared against standard guidelines for the management of asthma in England and Wales?
3. What are the key uncertainties relating to the cost-effectiveness of FeNO testing and how might these be resolved or reduced?

The next section presents the methods and results of a review of existing evidence relating to the cost-effectiveness of FeNO testing.

6.3 Review of existing evidence relating to the cost-effectiveness of FeNO testing for the diagnosis and management of asthma

6.3.1 Purpose of the review

We undertook a systematic review of existing economic analyses of FeNO testing in the diagnosis of asthma and for the management of patients with diagnosed asthma. This also included a focussed review of economic studies of other interventions for the diagnosis and/or management of asthma. The purpose of the review of existing health economic analyses was threefold:

1) To identify existing economic analyses of FeNO testing using NIOX MINO, NIOX VERO or NObreath for the diagnosis and/or management of asthma;
2) To identify existing models which may be used to inform the structure of the *de novo* economic models developed by the EAG;
3) To identify potentially relevant evidence sources to inform parameter values within the *de novo* economic models developed by the EAG.

6.3.2 Review methods

6.3.2.1 Methods used to identify existing economic studies

We undertook systematic searches across a range of electronic databases to identify published studies of FeNO testing for the diagnosis and/or management of asthma. We also searched for other economic studies of interventions for the diagnosis or management of asthma. All searches were undertaken by an Information Specialist (RW) during the period 30th May 2013 to 7th June 2013.

Four separate strands of searching were undertaken; these are detailed below.

Economic search 1: NIOX MINO/NObreath in either the diagnosis or management of asthma (30th May 2013)

This search used free-text terms relating to NIOX MINO and NObreath (including manufacturer names); these terms were combined with a sensitive economic search filter.

Economic search 2: Models of asthma and FeNO (30th May 2013)

This search used the search strategies developed for the management studies in the clinical effectiveness review (see Section 5.1.1) and combined these with a sensitive economic search filter. Studies that were found in the first search would also be retrieved in this search.

Economic search 3: Asthma management models (3rd June 2013)

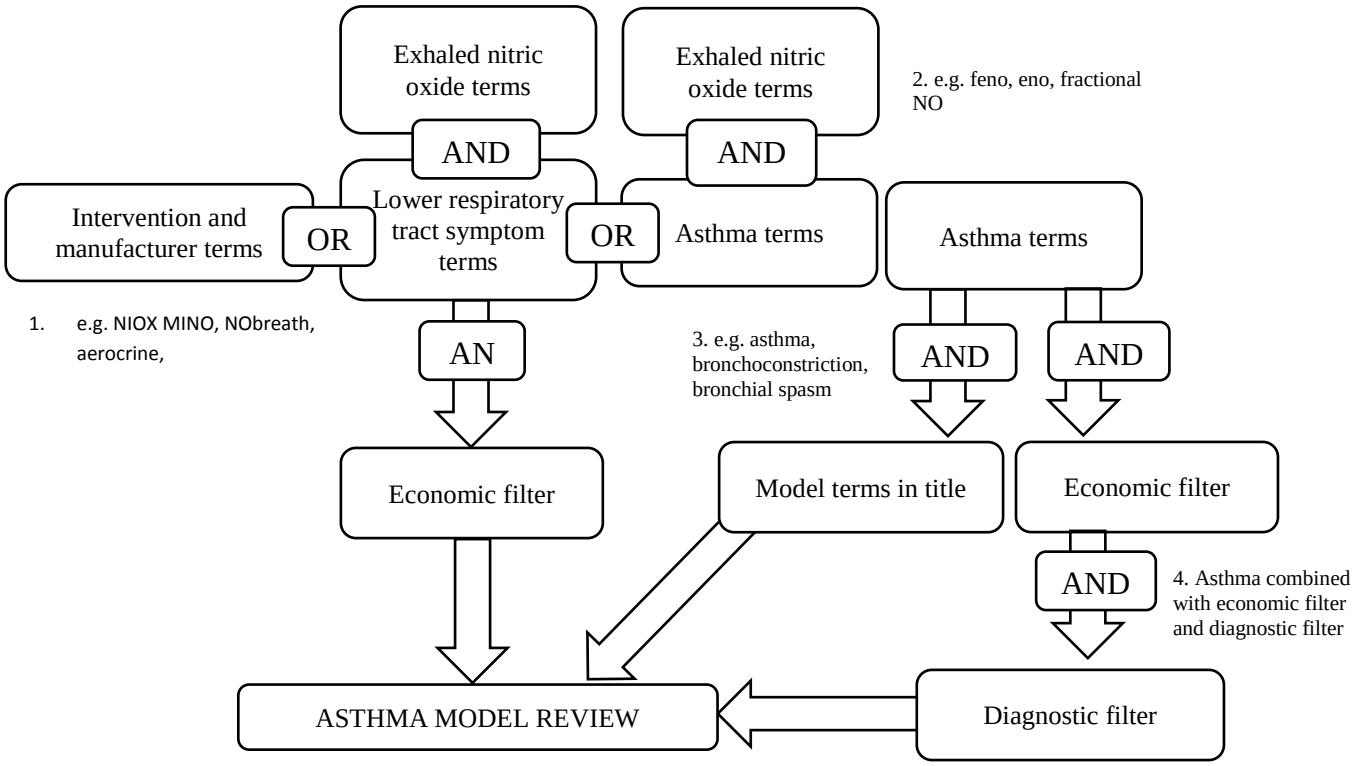
This focussed search used free-text terms for asthma combined with cost terms in the title and the economic model subject heading. A sensitive economic filter was not applied in this search.

Economic search 4: Asthma diagnostic models (7th June 2013)

This focussed search used free-text terms for asthma (as used in the asthma management model search [economic search 3 described above]) combined with a sensitive economic evaluations search filter and a diagnostic search filter.

These four searches are shown diagrammatically in Figure 20.

Figure 20: Diagrammatic representation of search approach



All of the above searches were performed within the following databases:

- MEDLINE and MEDLINE In Process: Ovid. 1948-present
- EMBASE: Ovid. 1974-present
- Cochrane Library: Wiley Interscience.
    - Cochrane Database of Systematic Reviews (CDSR)1996-present
    - Health Technology Assessment Database (HTA) 1995-present
    - NHS Economic Evaluation Database (NHS EED) 1995-present
- Science Citation Index Expanded (SCIE): Web of Science 1899-present
- Conference Proceedings Citation Index – Science (CPCI-S): Web of Science. 1990-present

The economic MEDLINE search strategy is detailed in Appendix 14.

As noted in Sextion 5.1.1.4, an additional separate search was also undertaken to identify evidence relating to NIOX VERO in August 2013.

6.3.2.2 Inclusion and exclusion criteria for the review

Given the anticipated dearth of published economic analyses of studies relating to FeNO, we adopted broad inclusion criteria for the review (see Box 1).

Box 1: Inclusion and exclusion criteria for the review of economic analyses of asthma diagnosis and management

Inclusion criteria

- Economic analyses of costs and consequences of interventions for the diagnosis and/or management of asthma in children and/or adults
- Studies reporting on the cost-effectiveness of NIOX MINO, NIOX VERO or NObreath for the diagnosis or management of asthma.

Exclusion criteria

- Letters, commentaries and editorials
- Economic studies which do not relate to diagnostic or management interventions
- Studies which do not relate to asthma
- Studies which do not involve (i) a model-based analysis, (ii) economic evaluations alongside trials or other forms of empirical clinical study *or* (iii) estimates of the costs and consequences of FeNO testing for the diagnosis of asthma

6.3.2.3 Data sifting

The titles and abstracts of all records identified by the search were reviewed by one member of the research team (JM). The full text of studies considered eligible for inclusion were then retrieved for a more detailed examination.

6.3.2.4 Critical appraisal methods

The identified studies of FeNO were critically appraised using the Drummond checklist for economic evaluations[131] and the NICE Reference Case for diagnostic studies.[132] The identified studies were also informally assessed against current guidelines for the development and reporting of health economic models.[133] Studies of other interventions for the diagnosis and/or management of asthma were not subjected to a formal critical appraisal but were instead used to inform the design and development of the *de novo* health economic analyses (detailed in Section 6.4).

6.3.3 Results of the review of FeNO for asthma diagnosis and management

6.3.3.1 Number and type of studies included in the review

The results of the four economic searches are presented in Table 50. A total of 1,898 potentially relevant citations were identified from the four strands of searches. The full texts of 27 studies were retrieved for further examination. The full text of one of these study could not be retrieved and was excluded. Of the remainder, only two studies were identified which related to FeNO testing in the diagnosis and/or management of asthma.[134,135] The focussed searches did not identify any further cost-utility models of other interventions for the diagnosis of asthma. Sifting of the focussed management model searches identified a further thirteen studies which were used more generally to inform the model structure, although none of these relate to FeNO testing.[136-148] In addition, one additional management study which was detailed in the appendices of a UK HTA report was identified.[13]

Table 50: Summary of results of the economic searches

Database	**Search**			
	1. NIOX MINO NObreath	**2. Asthma and FeNO**	**3. Asthma management**	**4. Asthma diagnostic**
MEDLINE and MEDLINE In Process	2	29	311	338
EMBASE	7	144	420	590
CDSR	0	48	0	69
HTA	4	8	4	0
DARE	0	2	3	14
NHS EED	1	2	119	12
SCIE	5	85	295	457
CPI-S	0	3	15	37
Total unique citations	**14**	**269**	**567**	**1048**

In addition, Aerocrine submitted evidence relating to the cost-effectiveness of NIOX MINO for the diagnosis and management of asthma.[149] This submission included a Microsoft Excel® spreadsheet model and a brief slideset. This submission is included as part of the economic review presented in this chapter. Aerocrine did not submit any economic evidence relating to the cost-effectiveness of the NIOX VERO device. Bedfont Scientific did not submit any evidence relating to either the effectiveness or cost-effectiveness of the NObreath device.

Section 6.3.3.2 presents the findings of the review of asthma diagnosis models; Section 6.2.3.3 presents the findings of the review of asthma management models.

6.3.3.2 Existing economic analyses of FeNO testing for the diagnosis of asthma

Methods and results of included diagnostic studies

The searches included only one UK model-based published economic analysis relating to the diagnosis of asthma;[134] this study assessed the cost-effectiveness of FeNO testing (specifically NIOX MINO) compared to standard diagnostic tests. This model has been published across two papers[134,135] and also forms the basis of the Aerocrine submission to NICE for this appraisal.[149] The general model structure and many of the evidence inputs are the same across these three analyses.

(i) Price et al (2009) - An economic evaluation of NIOX MINO airway inflammation monitor in the United Kingdom – diagnostic model[134]

Description of economic model and analysis

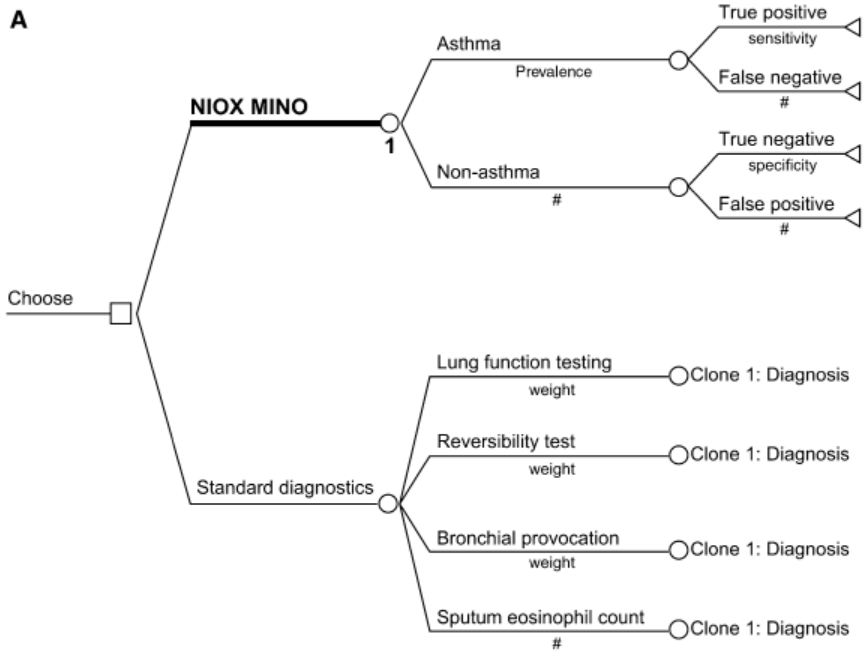
Price *et al* present the methods and results of two economic analyses: (1) a model to assess the cost savings associated with using NIOX MINO for the diagnosis of asthma and (2) a model to assess the cost-effectiveness of NIOX MINO for the management of asthma. The model of asthma management is reviewed in detail later in Section 6.2.3.3.

The conceptual form of the Price *et al* diagnostic model is presented in Figure 21. Within the model, the costs and outcomes of competing diagnostic strategies are modelled using a simple deterministic decision tree based on the true underlying probability of asthma and the operating characteristics of a variety of tests used for the diagnosis of asthma in the NHS. The population under evaluation within the Price diagnostic model is reported to relate to *"non-smoking adult patients with mild to severe asthma as seen in both primary and secondary care."*[134] The intervention is defined in the base case analysis as FeNO testing using NIOX MINO alone, although a secondary analysis is also reported for a joint diagnostic modality comprising of NIOX MINO plus spirometry using FEV_1 testing. The comparator within the base case analysis is a blended comparison of standard diagnostic tests: (1) lung function testing, (2) reversibility test, (3) bronchial provocation, and (4) sputum eosinophil count. The selection of tests included in the analysis was based on the BTS/SIGN asthma guidelines,

although the source for the proportionate weighting of each of these is unclear within the Price *et al* paper. It should also be noted that current BTS/SIGN guidelines state that sputum induction is not in common usage and it currently remains a research tool. In contrast to the published Price model, the Aerocrine submission model does not adopt a blended comparison approach but instead evaluates each individual diagnostic test as a decision option in its own right.

The model structure employs a single decision node whereby the model cohort is assumed to receive a single imperfect diagnostic intervention; those patients who receive an incorrect diagnosis are later assumed to achieve a correct diagnosis of either true asthma or not asthma. The published model estimates the costs associated with NIOX MINO versus the blended comparison of standard diagnostic tests. The analysis takes the form of a comparative cost analysis; health outcomes are not explicitly considered in the published analysis (note that the number of misdiagnoses are not reported within the Price *et al* paper but could be easily calculated from the table of model input parameters). Diagnostic outcomes in terms of true-/false positive/negative are estimated explicitly within the Aerocrine model. Within the Price *et al* paper, costs are valued at 2005 prices. The model time horizon is undefined but relates to the time from presentation to correct diagnosis. No discounting is applied to costs.

Figure 21: Model structure employed within the Price diagnostic model[134]



The Price diagnostic model makes the following structural assumptions:

- NIOX MINO will replace existing diagnostic tests rather than be used alongside them.

- Time is not explicitly considered within the model with respect to the resolution of incorrect diagnoses (false-positives or false-negatives).
- Negative health consequences (QALY losses) associated with incorrect diagnoses are not quantified within the model.
- All incorrect diagnoses are assumed to be corrected at the next outpatient visit.

The parameter values and evidence sources from which these are drawn are reported in Table 51.

Table 51: All parameter values and evidence sources used in the Price diagnostic model[134]

Parameter	Value	Source
Test operating characteristics		
Sensitivity FeNO (flow rate 50ml/s; >20ppb)	0.88	Smith *et al*[90]
Specificity FeNO (flow rate 50ml/s; >20ppb)	0.79	
Sensitivity FeNO (flow rate 50ml/s; >33ppb) + FEV_1<80% predicted	0.94	Smith and Taylor[150]*
Specificity FeNO (flow rate 50 ml/s; >33ppb) + FEV_1<80% predicted	0.93	
Sensitivity PEF A%M >21.6%	0.43	Hunter *et al*[151]
Specificity PEF A%M >21.6%	0.75	
Sensitivity Reversibility test: FEV_1 >2.9% improvement after salbutamol	0.49	
Specificity Reversibility test: FEV_1 >2.9% improvement after salbutamol	0.70	
Sensitivity Bronchial provocation: methacholine PC_{20} <8mg/ml	0.91	
Specificity Bronchial provocation: methacholine PC_{20} <8mg/ml	0.90	
Sensitivity sputum eosinophil count >1%	0.72	
Specificity sputum eosinophil count >1%	0.80	
Disease characteristics parameters		
Asthma prevalence	0.36	Smith *et al*.[90]
Comparator usage (blended comparison weightings)		
Proportion usage PEF charting	0.485	BTS/SIGN guidelines[152]
Proportion usage reversibility testing	0.485	
Proportion usage bronchial provocation	0.025	
Proportion usage sputum eosinophil count	0.005	
Cost parameters		
Cost NIOX MINO	£22.90	Aerocrine AB
Cost peak flow charting (2 visits)	£89.27	NHS Reference Costs 2005[153]
Cost reversibility test	£29.27	
Cost bronchial provocation	£48.50	
Cost sputum eosinophil count	£48.50	
Cost outpatient GP visit	£30.00	PSSRU[154]
Cost outpatient lung practitioner	£44.00	

** note this is a non-systematic review / opinion paper. Whilst Smith and Taylor[150] does state these sensitivity and specificity values and refers to two other empirical studies, neither includes the quoted estimates. The empirical source of the reported values for FeNO plus FEV_1 is unclear.*

The headline results of the economic analysis are presented as a simple cost difference between NIOX MINO and the blended comparison of standard tests for asthma diagnosis. Uncertainty surrounding model input parameters was explored using simple one-way sensitivity analyses. These analyses include varying model parameters describing test sensitivity, true underlying asthma prevalence in the modelled population, the costs of NIOX MINO and other diagnostic tests, the number of additional visits required in order to resolve an initially incorrect diagnosis, a comparison of NIOX MINO versus reversibility testing plus PEF charting and a comparison of NIOX MINO plus FEV_1 versus standard tests.

Diagnostic model results presented by Price et al[134]

The diagnostic model results reported by Price *et al*[134] are summarised in Table 52. In the base case analysis, the authors report that the cost of an asthma diagnosis made using NIOX MINO was £29 per patient, or £43 less than when using standard diagnostic tests (£72 per patient).

Table 52: Summary of cost-minimisation results presented by Price *et al*[134]

Scenario	NIOX MINO	Standard tests	Incremental
Base case	£29	£72	-£43
Variation in test sensitivity -50% (all tests simultaneously)	£35	£76	-£40
Variation in test sensitivity +10% (all tests simultaneously)	£29	£72	-£43
Variation in test sensitivity -50% (bronchial provocation and sputum only)	£39	£81	-£42
Variation in test sensitivity +10% (bronchial provocation and sputum only)	£28	£71	-£43
Asthma prevalence set to 10%	£30	£70	-£40
Asthma prevalence set to 50%	£29	£74	-£45
Asthma prevalence set to 90%	£28	£78	-£50
NIOX MINO cost -50%	£18	£72	-£54
NIOX MINO cost +200%	£75	£72	£3
Cost of standard diagnostic tests +50%	£29	£72	-£43
Cost of standard diagnostic tests +100%	£29	£102	-£72
Cost of standard diagnostic tests +150%	£29	£131	-£102
Cost of standard diagnostic tests +200%	£29	£161	-£131
2 visits for false diagnosis	£36	£86	-£50
4 visits for false diagnosis	£49	£113	-£63
NIOX MINO vs reversibility + PEF charting	£29	£131	-£102
NIOX MINO + FEV_1 vs standard tests	£115	£72	£42

The results indicate that within the base case analysis, NIOX MINO is expected to produce cost savings (£43) when compared against the blended comparison of standard diagnostic tests for asthma. These results do not account for potential health benefits associated with the improved accuracy of diagnosis. The sensitivity analysis indicates that NIOX MINO is expected to produce cost savings in

all scenarios except (i) when the cost of NIOX MINO is increased by 200% and (ii) within the comparison of NIOX MINO plus FEV_1 testing versus the blended comparison of current standard diagnostic tests.

The authors note that *"is it is likely that, in practice, FeNO measurement will be used in conjunction with other tests rather than as their replacement. We examined this scenario and found that the combination of FeNO measurement plus lung function testing increased costs for diagnosing asthma by £42."*[134] Given the authors' interpretation of the likely placement of NIOX MINO, it is unclear why the base case analysis within the paper does not reflect this scenario. Given the proposed placement of FeNO within the existing pathway and the absence of quantified health outcomes within the Price diagnostic model, it is unclear whether the potential additional benefits associated with diagnosis using FeNO testing outweigh the opportunity costs associated with generating them.

The next section briefly outlines the economic analysis of NIOX MINO for asthma diagnosis as presented within the Aerocrine submission to NICE.

(ii) Additional analysis presented within the submitted Aerocrine diagnostic model

As noted above, Aerocrine also submitted a spreadsheet model to NICE as part of the appraisal process.[149] The model was accompanied by a brief Microsoft Powerpoint® slideset although this does not include a description of the intended base case analysis results and little detail is provided supporting the structure, assumptions or choices regarding evidence used to inform the model parameters. The submitted Aerocrine model adopts a very similar structure and assumptions to the diagnostic model reported by Price *et al*.[134] It should be noted that in the absence of a detailed written description of the Aerocrine submission model, it is difficult to provide a full critique of its methods and results. This task was further hindered as the worksheet tabs and many sets of calculations were structurally hidden within the Excel worksheet thus making formula auditing problematic.

The following differences should be noted between the Price diagnostic model[134] and the Aerocrine diagnostic model:[149]

(i) Differences in the specification of diagnostic options:

The Aerocrine model assesses a different set of options to Price *et al*.[134] The Aerocrine model assesses the following diagnostic options:

- Spirometry alone
- Spirometry and (if negative) methacholine challenge
- Spirometry and (if negative) FeNO

- Spirometry and FeNO
- FeNO alone
- Spirometry and (if negative) sputum induction

It should be noted that some of these options include sequences of diagnostic tests; these are implemented within the model by assuming that the probabilities of obtaining a positive or negative results from sequences of tests are uncorrelated with one another. In other words, the use of prior tests in a sequence will remove some candidates from the population, alter the prevalence of true disease in the remaining population and may impact upon diagnostic accuracy of subsequent tests in that sequence. The validity of assuming no correlation between tests is questionable and no evidence is presented to support this. Within the Aerocrine submission model, all standard tests are evaluated as individual comparators in their own right rather than being combined and weighted within a blended comparison.

(ii) Different assumptions relating to the cost impact of misdiagnosis and resolution

The submitted model includes the costs of treating patients who are false-positive using inhaled corticosteroids over a 1-year time horizon; these treatment costs were not included in the published Price diagnostic model.[134] Conversely, the Aerocrine model does not include the assumptions made by Price *et al* regarding the costs of additional visits to resolve misdiagnosis.

(iii) Different parameter values and evidence sources

The Aerocrine diagnostic model includes some different parameter values to the Price diagnostic model. The parameter values and sources employed within the Aerocrine diagnostic model are detailed in Table 53.

Table 53: Key parameter values and evidence sources used in the Aerocrine diagnostic model

Parameter	Value	Source
Test operating characteristics		
Sensitivity spirometry alone	0.29	Schneider *et al*[77]
Specificity PEF A%M >21.6%	0.90	
Sensitivity FeNO+spirometry	0.94	Smith and Taylor[150]*
Specificity FeNO+spirometry	0.93	
Sensitivity FeNO alone	0.88	Smith *et al*[90]
Specificity FeNO alone	0.79	
Sensitivity methacholine challenge	0.91	Hunter *et al*.[151]
Specificity methacholine challenge	0.90	
Sensitivity sputum induction	0.72	
Specificity sputum induction	0.80	
Disease characteristics parameters		
Asthma prevalence	0.36	Smith *et al*[90]
Cost parameters		
Cost spirometry	£1	Source unclear
Cost spirometry plus FeNO	£11	Assumption
Cost FeNO	£10	Assumption
Cost spirometry plus methacholine challenge	£63	2005 NHS Reference Costs (reported in Price *et al*[134]) uplifted to 2012 values
Cost spirometry and sputum induction	£63	2005 NHS Reference Costs (reported in Price *et al*[134]) uplifted to 2012 values
Annual NHS cost for long-acting inhaled corticosteroids (prescribing using standard guidelines)	£138	BNF 51[155] (reported in Price *et al*[134] uplifted to 2012 values)

** note this is a non-systematic review / opinion paper. Whilst Smith and Taylor[150] does state these sensitivity and specificity values and refers to two other empirical studies, neither includes the quoted estimates. The empirical source of the reported values for FeNO plus FEV_1 is unclear.*

It should be noted that the marginal per-test cost for NIOX MINO within the Aerocrine model is assumed to be £10.00; this is substantially lower than that assumed within the Price et al paper (£22.90). Via email correspondence with the EAG, the manufacturer stated that *"The £10 per-use cost is a subjective figure and the model is designed to allow this number to be changed according to local conditions and costs deemed relevant to local payers… The £22.90 figure was one identified by Price et al as an indicative value incorporating all the externality costs in secondary care. Again this is a figure that needs to be adapted to local payer requirements. Most overhead and labour costs are fixed in the NHS and as such would mean that the higher figure may not be realistic hence the lower figure of £10 that has been used."* (Personal communication: Mr David Plotts, Director for Northern Europe & UK Managing Director, Aerocrine).

Summary of results of the Aerocrine diagnostic model

Table 54 presents the results presented within the Aerocrine diagnostic model. The results relate to a population of 840 patients; this population size is not justified within the model.

Table 54: Results estimated within the Aerocrine diagnostic model[149]

Diagnostic option	No. correct diagnoses (true-positive and true-negatives)	No. incorrect diagnoses (false-positive and false-negatives)	Difference	Cost incorrect diagnoses (diagnosis cost only)	No. of false-positive diagnoses	Cost false-positive diagnoses	Cost false-positive steroid use
Spirometry alone	572	268	303	£268	54	£54	£7,419
Spirometry and (if spirometry negative) methacholine challenge	719	121	597	£4,319	102	£3,102	£14,096
Spirometry and (if spirometry negative) FeNO	659	181	478	£1,455	155	£1,171	£21,441
Spirometry and FeNO combined	784	56	728	£614	38	£414	£5,193
FeNO alone	691	149	542	£1,492	113	£1,129	£15,580
Spirometry and (if spirometry negative) sputum induction	629	211	419	£9,938	151	£6,150	£20,773

The Aerocrine diagnostic model suggests that the combination of spirometry plus FeNO testing is expected to result in the greatest number of correct diagnoses and the fewest number of incorrect diagnoses. This is due to the assumed sensitivity and specificity of this combination (sourced from the expert review paper by Smith and Taylor[150]), both of which are higher than the values for all other tests included in the analysis.

Critical appraisal of the Price/Aerocrine diagnostic models

The use of the Price/Aerocrine diagnostic model to inform judgements about the cost-effectiveness of NIOX MINO versus standard diagnostic tests for asthma is subject to a number of problems; these are detailed below.

(i) Deviations from NICE Reference Case

Table 55 shows the extent to which the Price diagnostic model and the Aerocrine diagnostic models adhere to the NICE Reference Case for economic evaluations of diagnostic interventions.[132] Whilst the Price diagnostic model was not originally developed to inform this NICE appraisal, the model

submitted by Aerocrine follows the same general approach and therefore should be interpreted in light of NICE's Reference Case.

Table 55: Adherence of the Price/Aerocrine diagnostic model to the NICE Reference Case

Element of health technology assessment	Reference Case	EAG comments
Defining the decision problem	The scope developed by the Institute	The Price diagnostic model was not developed specifically to inform the NICE diagnostic appraisal of FeNO, yet this same general model approach was employed within the Aerocrine submission. The intervention and comparators are generally in line with the NICE scope. However the economic outcome does not include health consequences quantified in terms of health gains/losses. The population in both models is restricted to non-smoking adults with mild to severe asthma as seen in both primary and secondary care.
Comparator	Therapies routinely used in the NHS, including technologies regarded as current best practice	Comparators include tests commonly used in the NHS for the diagnosis of asthma - bronchial provocation, lung function testing, reversibility testing and sputum eosinophil count. Sputum induction is not widely used in England and Wales. Importantly, the base case analysis is presented as a blended comparison rather than as an incremental analysis between individual options. This is generally inappropriate as it may mask the most effective and/or the most cost-effective diagnostic option. Within the Aerocrine model, options are evaluated as individual diagnostic interventions. These include spirometry alone, spirometry and (if negative) methacholine challenge, spirometry and (if negative) FeNO, spirometry and FeNO, FeNO alone and spirometry and (if negative) sputum induction.
Perspective on costs	NHS and PSS	A payer perspective was adopted by Price *et al*[134] however this is restricted to short-term costs only; treatment costs for diagnosed asthma are not included. The Aerocrine diagnostic model includes costs of diagnostic tests and treatment costs for false-positives. The time horizon for costing is not explicit. Personal Social Services (PSS) costs are not considered in either model.
Perspective on outcomes	All health effects on individuals	Health gains and losses associated with correct/incorrect diagnoses are not reported by Price *et al*. The Aerocrine model reports numbers of true-/false- positive/negatives expected within a cohort of 840 patients.
Type of economic evaluation	Cost-utility analysis	The Price diagnostic analysis represents a cost-comparison; whilst diagnostic outcomes are calculable, these are not reported. The Aerocrine model quantifies numbers of correct/incorrect diagnoses but does not values these in terms of health gains or losses.
Synthesis of evidence on outcomes	Based on systematic review	Price *et al* report that estimates of test sensitivity and specificity are based on three published papers identified by a systematic review of the literature.[90,150,151] The

Element of health technology assessment	Reference Case	EAG comments
		Aerocrine submission does not present any detail regarding methods used to identify or select evidence used to inform its parameters. The full range of empirical evidence relating to the diagnostic accuracy of FeNO used in combination with other tests is not captured in either model.
Measure of health effects	QALYs	Neither the Price diagnostic model nor the Aerocrine diagnostic model measure or value health outcomes associated with correct/incorrect diagnoses.
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	HRQoL is not captured in either model.
Source of preference data for valuation of changes in HRQoL	Representative sample of the public	HRQoL is not captured in either model.
Discount rate	An annual rate of 3.5% on both costs and health effects	In both models, costs and outcomes are not discounted.
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	HRQoL is not captured in either model.

(ii) Absence of quantified health consequences resulting from diagnostic decisions

A key limitation of both the Price diagnostic model and the Aerocrine diagnostic model is that neither model attempts to value the health gains/losses resulting from correct /incorrect diagnoses of asthma. Whilst it may be reasonably inferred that a more sensitive and specific test will result in more correct diagnoses, and hence greater health gains due to the use of that test, these factors are not captured within either model. Consequently, it is difficult to infer whether the health gains associated with a more sensitive and/or specific test outweigh the potential opportunity costs associated with displacing existing treatments and services.

(iii) Use of a blended comparison approach (Price diagnostic model only)

The base case analysis presented within the Price paper adopts a blended comparison approach. Results are not presented as an incremental comparison of the costs and consequences of NIOX MINO versus individual comparator tests. This is misleading - whilst the base case analysis suggests

that NIOX MINO alone is more sensitive than the weighted mix of standard tests used to diagnose asthma, its sensitivity and specificity are both lower than those for bronchial provocation. It would be more appropriate to incrementally compare NIOX MINO versus each individual diagnostic test; this is the approach adopted within the Aerocrine model submitted to NICE.

(iv) Anticipated use of NIOX MINO

Both the Price diagnostic model and the Aerocrine diagnostic model reflect a situation in which NIOX MINO would replace existing standard tests for the diagnosis of asthma. The situation in which NIOX MINO is added to existing tests within the pathway, versus those existing tests, as is suggested to be the more likely use of FeNO testing within the NHS by Price *et al*,[134] is not adequately considered within the analysis. In addition, both models lack clarity with respect to the diagnostic setting in which the choice of diagnostic strategy is made (i.e. primary or secondary care).

(v) Non-specific placement of NIOX MINO within the broader diagnostic pathway for asthma

The published Price analysis crudely compares NIOX MINO versus a comparison of individual diagnostic tests. In reality, some patients may only achieve a positive or negative diagnosis following a sequence of tests. This is undoubtedly an issue relating to the available evidence base at the time of model development, however this limitation should be borne in mind when interpreting the results reported within the Price paper. In contrast, the submitted Aerocrine model includes some test sequences, however these do not reflect potential correlations between each test in the pathway (sensitivity and specificity are assumed to be random and uncorrelated between tests).

(vi) Crude assumptions regarding the resolution of incorrect diagnoses

The Price model assumes that incorrect diagnoses are resolved at the next visit. Conversely, the submitted Aerocrine model does not include the costs of additional visits required to resolve incorrect diagnoses, but instead attempts to capture the costs associated with inhaled corticosteroid use in patients who are false-positive. Both of these factors reflect relevant costs to the NHS and should be included in any economic analysis of FeNO testing. In addition, the time horizon over which incorrect diagnoses prevail is unclear and no discounting is applied to cost estimates.

(vii) Questionable validity of FeNO+FEV_1 operating characteristics

It is noteworthy that the source of the estimates of test sensitivity and specificity for spirometry plus FeNO, the most favourable option within both diagnostic models included in this review,[134,149] appear to have been derived from an expert review paper[150] rather than an empirical study. The expert review paper does make reference to the sensitivity and specificity estimates of 0.93 and 0.94 as used in the model, and does provide an apparent (yet ambiguous) reference to two other empirical studies.[90,156] However, neither the Dupont nor Smith studies[90,156] referenced by Smith and Taylor[150] report these

estimates (or indeed any estimate of the joint sensitivity and specificity of FeNO plus FEV_1). The credibility of these estimates cannot be verified by the EAG and hence the credibility of the Price/Aerocrine model findings should be considered highly questionable.

(viii) Lack of clarity regarding methods to identify and select evidence

Within the Aerocrine model, the sources of the costs of spirometry are unclear and the costs of NIOX MINO appear to be based solely on assumption (see earlier personal communication). The costs of NIOX MINO are substantially different between the two models (£22.90 versus £10). It is unclear whether either estimate would reflect the true costs borne by the NHS. In addition, the methods used to identify and select evidence regarding test operating characteristics are particularly unclear within the Aerocrine submission model.

(ix) Limited consideration of uncertainty

Both versions of the diagnostic model are evaluated deterministically using point estimates of parameters. Probabilistic sensitivity analysis is not reported by Price *et al*[134] and is not included in the submitted Aerocrine model.[149]

It is reasonable to suggest that the existing evidence base relating to the cost-effectiveness of FeNO testing for the diagnosis of asthma is methodologically limited and should be interpreted with caution.

6.2.3.3 Existing economic analyses of FeNO testing for the management of asthma

Methods and results of included management studies

The Price *et al* study[134] detailed above also included the methods and results of a separate model of the cost-effectiveness of FeNO testing using NIOX MINO for the management of asthma. The same model structure was also used in the German economic evaluation of FeNO testing for asthma reported by Berg *et al*.[135] No other published papers were identified which related to FeNO testing for the management of asthma. In addition, the submission by Aerocrine[149] also included an asthma management model based on the analysis published by Price *et al*.[134]

(i) Price et al (2009)- An economic evaluation of NIOX MINO airway inflammation monitor in the United Kingdom[134] – management model

Description of economic model and analysis

The management model as described by Price *et al*[134] uses a decision tree approach to evaluate the cost-effectiveness of NIOX MINO versus standard guidelines for the management of asthma. The model adopts a UK health care payer perspective and costs and outcomes are evaluated over a 1-year time horizon. Results are presented in terms of the incremental cost per QALY gained. Patients within the model were assumed to be non-smokers with mild to severe diagnosed asthma. Patients were

assumed to be at Step 3 and above as per Global Initiative for Asthma[74] and BTS/SIGN guidelines,[152] i.e. receiving ICS and LABA for asthma management. Patients were assumed to visit their GP four times per year to determine appropriate ICS dosage; it is unclear whether this applies to groups or the FeNO management group only.

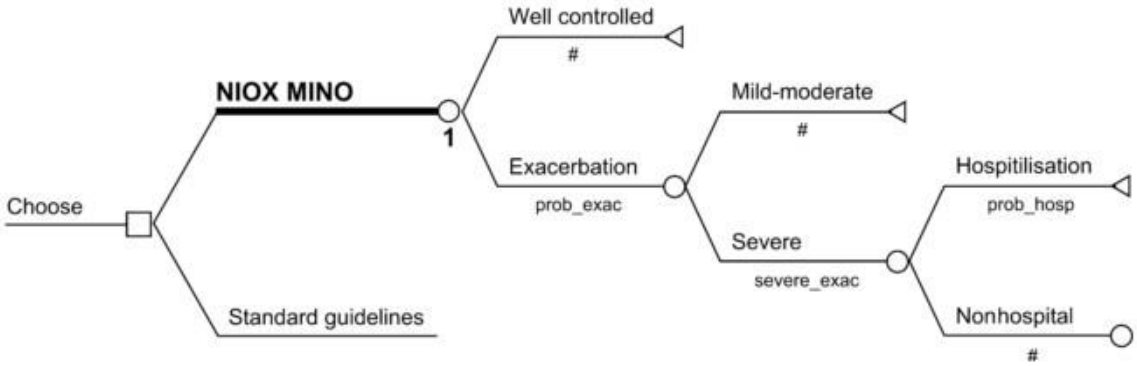
The two management strategies compared within the model were:

- Intervention: ICS dosage titration using NIOX MINO.
- Comparator: ICS dosage titration based on standard guidelines.

The model uses different sources to inform parameters relating to the baseline risks and relative risks of exacerbation and ICS use.[2,107,157] Only one of these three studies[2] involved asthma management according to BTS/SIGN guidelines.

The structure of the model is shown in Figure 22. The model assumes that patients are either well controlled, or experience an exacerbation. Exacerbations are assumed to be either severe or mild-moderate. A proportion of severe exacerbations are assumed to require hospitalisation whilst the remainder are assumed to be manageable on an outpatient basis. A mild to moderate exacerbation was defined as an exacerbation requiring SABA in addition to usual medication; a severe exacerbation was defined as an exacerbation requiring corticosteroids (and in some patients, hospitalisation). The successful control of exacerbations is assumed to be related to improvements in HRQoL and reductions in ICS use.

Figure 22: Model structure employed within the Price management model[134]



The parameter values and evidence sources listed in the Price management model are shown in Table 56.

Table 56: All parameter values and evidence sources used in the Price management model

Parameters	Value	Source
Baseline event probabilities		
Exacerbation risk during 1 year	0.71	Jayaram *et al*[157]
Proportion of exacerbations that are severe	0.23	Jayaram *et al*[157]
Hospitalisation for severe exacerbations	0.23	Green *et al*[2]
Proportion of severe exacerbations requiring outpatient visit	0.75	Andersson *et al*[158]
Mean number of severe exacerbations per year (overall population)	2	Jayaram *et al*;[157] Tattersfield *et al*[159]
Mean number of severe exacerbations per year (moderate to severe asthma)	4	Green *et al*[2]
Impact of FeNO management		
Reduction in ICS dose	0.42	Smith *et al*[107]
Relative risk reduction of exacerbation	0.29	Jayaram *et al*[157]
Relative risk reduction of hospitalisation for severe exacerbation	0.83	Green *et al*[2]
Utility values		
Well controlled asthma	0.93	Szende *et al*[160]
Mild/moderate exacerbation	0.65	Szende *et al*[160]
Severe exacerbation	0.52	Szende *et al*[160]
Resource cost parameters		
Outpatient visit to GP	£30.00	PSSRU[154]
Outpatient visit to lung specialist	£44.00	PSSRU[154]
Hospitalisation for asthma	£2231.45	BNF 51[155]
Maintenance therapy (1 year) with long-acting β2-agonist	£359.84	BNF 51[155]
Maintenance therapy (1 year) with inhaled corticosteroid	£109.00	BNF 51[155]
Rescue therapy (1 week) with short-acting β 2-agonist	£7.38	BNF 51[155]
Rescue therapy (1 week) with oral prednisone	£5.13	BNF 51[155]

In addition to the base case analysis, the authors undertook 18 one-way sensitivity analyses. These include examining the impact of the baseline risk of exacerbations, health utilities, the number of routine visits required per year, ICS dose reductions and the costs of NIOX MINO on the cost-effectiveness of NIOX MINO versus standard guidelines.

Management model results presented by Price et al[134]

The model results reported by Price *et al* are presented in Table 57. For patients with moderate-severe asthma, FeNO monitoring was estimated to result in 0.004 additional QALYs compared to standard guidelines. FeNO was also estimated to result in cost savings of £554 per patient. For patients with mild-moderate asthma, FeNO monitoring was estimated to result in 0.06 additional QALYs compared to standard standard guidelines. FeNO was also estimated to result in cost savings of £341 per patient in this group. Given its lower cost and increased QALY gain, FeNO was expected to dominate standard guidelines in both patient groups. It should be noted that the distinction between mild-moderate and moderate-severe in terms of input parameters is not entirely clear from the Price paper.

The results of the simple sensitivity analyses indicate that for all but one scenario (NIOX MINO in addition to rather than instead of standard lung function tests), NIOX MINO is expected to dominate standard guidelines. Within the last scenario, NIOX MINO in addition to standard lung function tests is expected to cost £279 per QALY gained as compared against standard guidelines.

Table 57: Sensitivity analyses results reported by Price *et al*[134]

	Cost (£)			**QALY**			**Incremental cost-effectiveness ratio (ICER)**
	NIOX MINO	Standard guidelines	Difference	NIOX MINO	Standard guidelines	Difference	
Moderate/Severe Asthma	628	1181	-554	0.730	0.726	0.004	Dominating
1 year baseline risk of exacerbation of 0.35 (base case, 0.71)							
	589	915	-326	0.857	0.83	0.027	Dominating
Utility for moderate control of asthma of 0.76 (base case, 0.65)							
	666	1007	-341	0.835	0.800	0.035	Dominating
Different number of monitoring visits per year for mild/severe asthma (base case, four visits)							
Two visits/ year	620	828	-208	0.785	0.726	0.059	Dominating
Six visits/year	712	1185	-473	0.785	0.726	0.059	Dominating
Different number of monitoring visits per year for moderate/severe asthma (base case, four visits)							
Two visits/ year	582	1003	-421	0.730	0.726	0.004	Dominating
Six visits/year	673	1360	-687	0.730	0.726	0.004	Dominating
Different NIOX MINO cost for mild/severe asthma							
-50%	620	1007	-387	0.785	0.726	0.059	Dominating
+50%	712	1007	-295	0.785	0.726	0.059	Dominating
Different NIOX MINO cost for moderate/severe asthma							
-50%	582	1181	-599	0.730	0.726	0.004	Dominating
+50%	673	1181	-508	0.730	0.726	0.004	Dominating
Different level of ICS dose reduction for mild/severe asthma (base case 42%)							
10%	683	1007	-324	0.785	0.726	0.059	Dominating
80%	645	1007	-362	0.785	0.726	0.059	Dominating
Different level of ICS dose reduction for moderate/severe asthma (base case 42%)							
10%	639	1181	-543	0.730	0.726	0.004	Dominating
80%	616	1181	-565	0.730	0.726	0.004	Dominating
Different relative risk reduction for exacerbation for mild/severe asthma (base case, 29%)							
10%	707	1007	-300	0.747	0.726	0.021	Dominating
50%	621	1007	-386	0.828	0.726	0.102	Dominating
Different relative risk reduction for hospitalisation for moderate/severe exacerbation (base case, 0.83)							
10%	869	1181	-312	0.727	0.726	0.001	Dominating
100%	571	1181	-610	0.731	0.726	0.005	Dominating
NIOX MINO in addition to rather than instead of standard lung function tests (added costs)							
	1023	1007	17	0.785	0.726	0.059	£279 per QALY gained

(ii) Additional analysis presented within the submitted Aerocrine management model[149]

The schematic of the Aerocrine management model is presented in Figure 23.

Figure 23: Management model submitted by Aerocrine

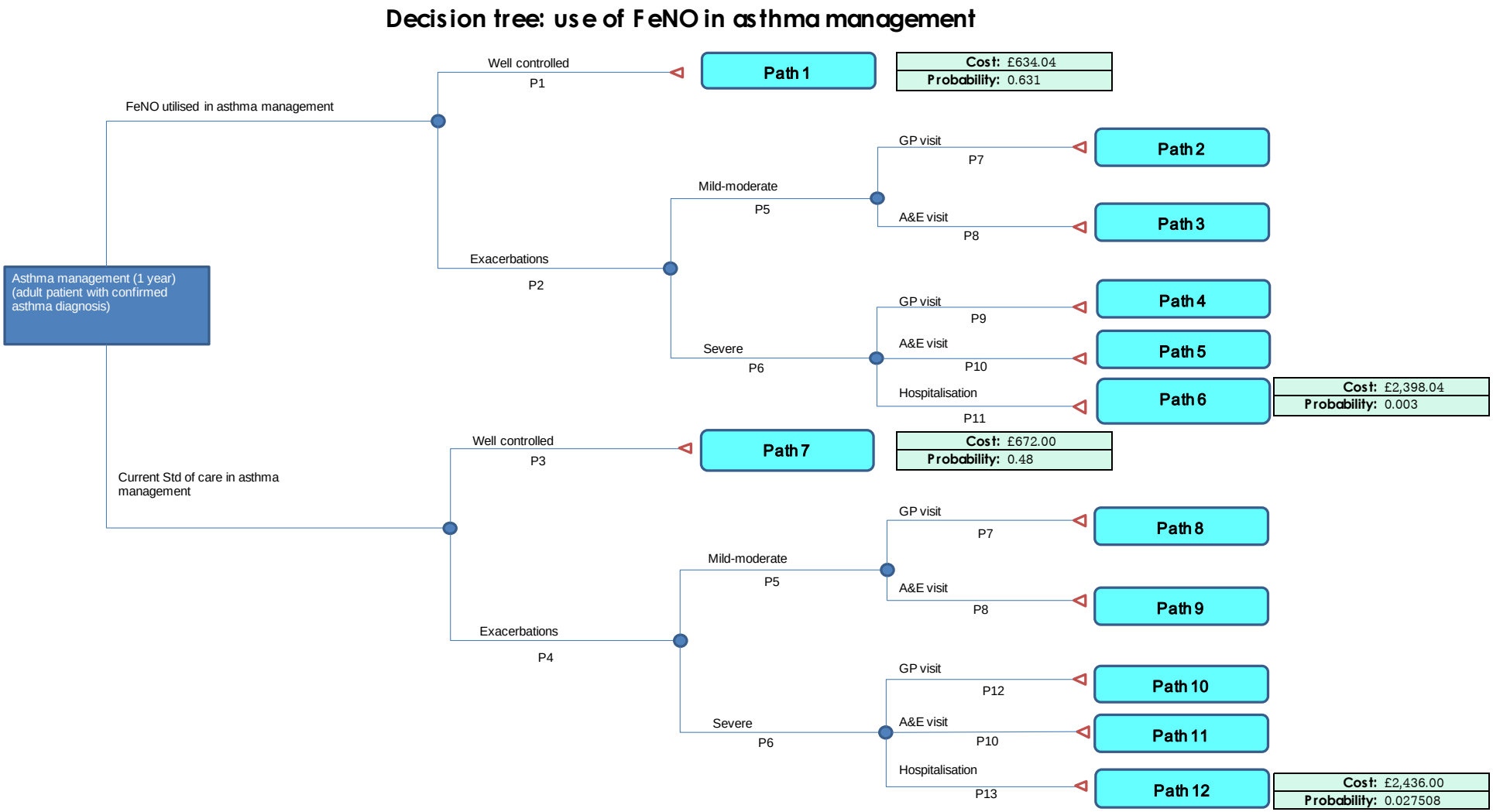


Table 58 presents the parameter values and evidence sources used in the submitted Aerocrine management model; the column on the right hand side indicates whether the source and parameter value is the same as that used in the published Price management model.[134]

Table 58: Parameter values used in the Aerocrine management model

Variable ID	Variable description	Value	Source/justification	Same as Price model[134]
P2	Likelihood of exacerbation using FeNO for management	0.369	Jayaram *et al*[157]	No
P4	Likelihood of exacerbation using Standard Care Guidelines for asthma management	0.520	Akinbami *et al*[161]	No
P6	Likelihood that exacerbations will be Moderate-Severe	0.230	Green *et al*[2]	Yes
P8	Likelihood that Mild-Moderate asthma exacerbations will be treated at an emergency room or urgent care centre	0.500	Expert opinion*	Unclear
P10	Likelihood that FeNO patient experiencing Moderate-Severe asthma exacerbations will be treated at an emergency room centre	0.750	Andersson *et al*[158]	Yes
P13	Likelihood that Standard Care patient experiencing a Moderate-Severe exacerbation will require hospitalization	0.230	Green *et al*[2]	Unclear
F1	Reduction in ICS dose due to FeNO use	0.42	Smith *et al*[107]	Yes
F2	Reduction due to FeNO use in risk of hospitalization for severe exacerbations	0.83	Green *et al*[2]	Yes
F3	Reduction in risk of exacerbations due to FeNO use	0.29	Jayaram *et al*[157]	Yes
C1	Cost of FeNO	£10.00	Assumption	Unclear
C2	Cost of Spirometry	£1.00	Source unclear	Unclear
C3	Annual cost of asthma medications for patients managed with FeNO	£536.04	BNF 51[155] (uplifted to 2012 prices)	Unclear
C4	Annual cost of asthma medications using Standard Guidelines	£594.00	BNF 51[155] (uplifted to 2012 prices)	Unclear
C5	Cost per office visit to General Practitioner	£38.00	PSSRU 2005[154] (uplifted to 2012 prices)	Unclear
C6	Cost per office visit (referral) to Lung Specialist	£144.00	PSSRU 2005[154] (uplifted to 2012 prices)	Unclear
C7	Cost of A&E. visit for asthma exacerbation	£81.00	NHS Reference Costs 2012[162]	Unclear
C8A	Cost of rescue medications for moderate - severe exacerbations	£15.00	BNF 51[155] (uplifted to 2012 prices)	Unclear
C8B	Cost of rescue medications for mild-moderate exacerbations	£9.00	BNF 51[155] (uplifted to 2012 prices)	Unclear
C9	Average hospital cost for asthma	£867.00	Weighted average of HRG's	Unclear

	admission due to exacerbation		DZ15A-F within NHS Reference Costs 2012[162]	
C10	Annual number of check-ups for asthma management	2	Expert opinion†	Yes
C13	Average annual number of exacerbations	2	Jayaram *et al*[157]	Yes
U1	Utility value of asthma patients with good control	0.93	Szende *et al*[160]	Yes
U2	Utility value of asthma patients with mildly reduced control	0.76	Szende *et al*[160]	Yes
U3	Utility value of asthma patients with moderately reduced control	0.65	Szende *et al*[160]	Yes
U4	Utility value of asthma patients with poor control	0.52	Szende *et al*[160]	Yes

**The text in the model states that the model authors were unable to find statistics specific to visits for mild/moderate exacerbations; an assumption was made that half will seek care in the emergency department setting and the other half will visit their doctors surgery (GP or pulmonary specialist)*
†An assumption was made that well controlled asthma will result in 2 office visits per year.

The management model submitted by Aerocrine is similar to the published Price management model in terms of its structure and both models share many common parameter values. However, the two models do not make identical assumptions and hence do not provide identical estimates of incremental costs and effects of FeNO monitoring compared against standard guidelines.

Summary of results of the Aerocrine management model

Whilst the Aerocrine management model does not present the ICER for FeNO monitoring versus standard guidelines in the main results worksheet, elsewhere the model indicates that FeNO monitoring is expected to produce an addition 0.045 QALYs and reduce costs by £103.11 compared to standard care.

Critical appraisal of the Price/Aerocrine management models

The use of the Price/Aerocrine management model to inform judgements about the cost-effectiveness of NIOX MINO is subject to a number of methodological problems, as detailed below.

(i) Deviations from NICE Reference Case

Table 59 shows the extent to which the Price/Aerocrine management models adhere to the NICE Reference Case for economic evaluations of diagnostic interventions.[132]

Table 59: Adherence of the Aerocrine management model to the NICE Reference Case

Element of health technology assessment	Reference Case	EAG comments
Defining the decision problem	The scope developed by the Institute	The patient population is defined in both models as non-smoking adults diagnosed with mild to severe asthma. This population excludes children and smokers. The intervention and comparator are in line with the NICE scope.
Comparator	Therapies routinely used in the NHS, including technologies regarded as current best practice	The comparator is standard care without FeNO monitoring. This is appropriate although it should be noted that the studies used to inform the model parameters did not all use BTS/SIGN guidelines to guide treatment.
Perspective on costs	NHS and PSS	The published Price management model purports to have adopted a payer perspective. It appears that the submitted Aerocrine model adopts the same perspective although this is not explicitly stated in the model workbook.
Perspective on outcomes	All health effects on individuals	Health outcomes reflect those accrued by NHS patients. Health gains are assumed to be influenced only by the level of control achieved which is in turn assumed to be directly related to the incidence of exacerbations.
Type of economic evaluation	Cost-utility analysis	The model takes the form of a decision-tree based cost-utility analysis. This adopts a short time horizon (1-year). Longer-term costs and outcomes associated with FeNO monitoring are not considered within the Aerocrine management model or the published Price management model.
Synthesis of evidence on outcomes	Based on systematic review	Parameter values appear to have been selected in a non-systematic fashion. Estimates of relative reductions of exacerbations are drawn from different sources than estimates of reductions in medication use (the former relates to monitoring using sputum inducation rather than FeNO testing but is assumed to be equivalent).
Measure of health effects	QALYs	The HRQoL impacts of different levels of control were estimated based on estimates from the literature.
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	Health utilities were based on adequacy of asthma control rather than exacerbations *per se*, based on a study reported by Szende *et al*.[160] Within this study, two hundred and twenty-eight consecutive adult outpatients and inpatients at four Hungarian sites completed a variety of HRQoL instruments including the EQ-5D. Utilities related to control were then qualitatively mapped to the incidence of different severities of exacerbation.
Source of preference data for valuation of changes in	Representative sample of the public	Preference-based health utilities appear to have been generated using the UK EQ-5D tariff.[163]

Element of health technology assessment	Reference Case	EAG comments
HRQoL		
Discount rate	An annual rate of 3.5% on both costs and health effects	Owing to the short time horizon costs and outcomes are not discounted.
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	No additional equity weighting was applied.

(ii) Relative risk reduction for exacerbations

Price *et al* argue that the relative risk reduction associated with using FeNO monitoring may be overly conservative as the data used were drawn from a patient population including patients with mild asthma, whereas the relative risk reductions may be greater in patients with more severe asthma. The validity of this statement is unclear and evidence to support this assertion is not presented in the Price paper.

(iii) Impact of FeNO measurement on ICS dosage

The authors also argue that some of the parameters, such as the effect of FeNO measurement on ICS usage, were based on patients in primary care, whereas other parameters, such as impact on exacerbations, were based on patients in secondary care.

(iv) Time horizon

The model adopts a very short time horizon (1-year). The impact of mortality and discounting over a longer horizon may alter the cost-effectiveness estimates presented.

(v) Failure to undertake probabilistic sensitivity analysis

The authors did not undertake probabilistic sensitivity analysis (PSA). Instead, the results are presented based on the point estimates of parameters and uncertainty analysis is restricted to one-way sensitivity analyses. It should be noted that the economic evaluation of NIOX MINO from the German perspective did include a full probabilistic analysis.[135] The reason for the exclusion of PSA in the UK models is unclear.

(vi) Questionable methods for the selection of evidence used to inform the model parameters

The methods used to identify and select evidence to inform the model parameters were not fully described in either the Price *et al* paper or the Aerocrine model. It is unclear whether other evidence

sources exist which indicate different parameter values may be more appropriate. In particular, the model draws estimates of the relative reduction in exacerbations for FeNO from a study which used sputum induction monitoring rather than FeNO monitoring, hence assuming equivalence, despite the fact that exacerbation risk information was reported in the FeNO trial used to estimate reductions in ICS usage.[107]

(vii) Inappropriate sourcing of resource and cost estimates

Several unit cost parameters within the Aerocrine model are based on those presented in Price *et al*[134] and uplifted to 2012 values. For parameters such as drug costs and HRGs, this is inappropriate as the BNF and NHS Reference Costs are updated regularly to reflect current prices. Consequently, several of the cost estimates included in the submitted model may not reflect the prices paid by the NHS.

(viii) Inflated baseline exacerbation rate without monitoring

The Price management model and the Aerocrine management model assume a mean rate of 2 exacerbations per patient per year. It appears that this estimate was based on results of a Phase 2 prospective trial of 117 adults reported by Jayaram *et al*.[157] This study reported that there were a total of 126 exacerbations in 63 patients, hence an average number of approximately 2 exacerbations per patient. However, the trial duration was greater than one year, and the mean number of exacerbations per patient per year was reported by the study authors to be 0.75 in one arm of the trial, and 1.02 in the other arm of the trial.[157] The Price *et al* paper[134] also mentions a second study used to inform this baseline exacerbation rate.[159] In this latter study, the authors observed 425 severe exacerbations in 852 randomised patients over a 12-month period (approximate rate = 0.499 exacerbations per year). Both studies clearly indicate that the baseline exacerbation rate used in the Price/Aerocrine model is substantially overestimated hence the expected benefits of FeNO testing are likely to be artificially inflated.

(ix) Assumption that exacerbation determines HRQoL for entire time horizon

The Price/Aerocrine management models make an assumption that the incidence of exacerbations is directly related to the level of control and applies health utilities according to the incidence of exacerbations. The model applies these health utilities over the modelled time horizon (1-year) rather than to the duration over which the exacerbation occurs (hours to weeks). This is likely to substantially overestimate the health benefits associated with reducing exacerbations through improved dose titration.

(x) Use of expert opinion

The Aerocrine management model includes the use of expert opinion to inform a small number of parameters where the authors could not identify relevant evidence. Whilst expert opinion is a valid

source of evidence in such circumstances, no details are provided with respect to the source of these judgements or the methods used to elicit them. In the absence of a written submission which presents these details, the credibility of such judgements remains unclear.

The existing models of FeNO monitoring for asthma[134,149] indicate that NIOX MINO is expected to dominate standard guidelines. However, given the methodological concerns identified within the critical appraisal, these findings should be interpreted tentatively.

6.2.3.4 Other studies relating to the cost-effectiveness of asthma management strategies

Given the limited number of studies of FeNO testing for the management of asthma, we also reviewed other studies of interventions for asthma management to inform the key disease-specific factors that should, or could, be included in a cost-effectiveness model of FeNO for the management of asthma. The main characteristics of these studies are briefly outlined below.

Economic evaluations using QALYs

Thirteen studies (not related to FeNO testing) were included in the focussed review of economic analyses of asthma management interventions; these studies are briefly summarised in Table 60.

Table 60: Summary of other identified economic analyses of asthma management interventions

Authors	Summary description
Studies reporting QALYs	
Peters *et al* 2002[136]	UK technology assessment report evaluating submissions from several manufacturers of inhaler devices. Most (6/8) of the manufacturers submitted cost-minimisation analyses only. The Assessment Group did not develop a *de novo* model. Instead a QALY-based threshold analysis was performed.
Briggs *et al* 2006[137]	Cost-utility analysis undertaken alongside a clinical trial. The intervention was asthma treatment with salmeterol/fluticasone propionate in combination (SFC); the comparator was fluticasone propionate (FP). Utility values for the model states were mapped from Asthma Quality of Life Questionnaire (AQLQ) scores. Within the Gaining Optimal Asthma controL (GOAL) study, patient treatment could be titrated upwards, up to three times, but not downwards. The amount of titration required was used to define three patient groups by asthma severity (stratum 1, stratum 2, stratum 3). The model states were "totally controlled", "well controlled", '"not well controlled but without an exacerbation", and "exacerbation." The cycle length was one week in duration. A multinomial regression approach using individual patient-level data from the trial was used to estimate the transition probabilities of moving between states over the course of each week.
Doull *et al* 2007[138]	Simple economic model comparing the cost-effectiveness of sulmeterol xinafoate/fluticasone proprionate combination inhalers (SFC) versus non-combination inhalers, for adults and children with chronic asthma treated according to BTS/SIGN guidelines. Clinical effectiveness was estimated from meta-analyses comparing the percentage of symptom-free days for each treatment (%SFD). The definition of SFD was assumed to be consistent with that provided in the GOAL study. The %SFD was assumed to be time-invariant, hence differences in clinical effectiveness between treatment options were assumed to be entirely due to this parameter. A 1-year time horizon was used. QALY gains were estimated from AQLQ data recorded in the GOAL study using a mapping algorithm to the EQ-5D.
Wilson *et al* 2010[139]	Economic evaluation comparing the addition of either LTRA or LABA for patients who were already receiving ICS as part of asthma management, and where a decision to add on additional treatment to improve the condition had been made. The analysis was based on a pragmatic trial involving 53 primary care practices. Patients judged to need add-on therapy were randomly assigned to receive either LTRA or LABA. The trial duration was two years. The patient age range included children and adults. The differences in EQ-5D and ACQ scores between LTRA and LABA groups were reported, together with differences in resource use.
Paggiaro *et al* 2011[140]	A poster which discusses a patient-level Markov model. The decision problem concerns the cost-effectiveness of stepping down treatment according to BTS/SIGN guidelines. Very limited detail on the methods was available.
Studies not based on QALYs	
Booth *et al* 1995[141]	Cost comparison based on an RCT comparing fluticasone proprionate 200 micrograms via a Diskhaler versus budesonide 200 micrograms via a reservoir dry powder device. The study provides estimates for the cost per successfully controlled week.
Barnes *et al* 1999[164]	Poster abstract which summarises a meta-analysis comparing fluticasone proprionate and budesonide for the treatment of asthma. The study appears mainly to be a cost-consequence analysis, as it refers to differences in clinical parameters, such as morning PEF rate, successfully treated weeks and symptom-free days. The poster concludes that fluticasone proprionate is both more clinically effective and cheaper than budesonide.

Andersson *et al* 2000[143,143,165]	Poster abstract which argues that using 800 rather than 200 micrograms of budesonide per day is cost-saving in patients with moderate asthma in the UK. Estimates were based on a survey of 20 physicians from the UK, Sweden and Spain.
Everden *et al* 2002[144]	Economic evaluation in children aged 6-17 years inclusive, alongside a prospective multi-centre open-label parallel group study conducted in primary care in the UK and Republic of Ireland (The FACT study). Most (>95%) patients were at BTS Step 1 or Step 2 with a small proportion at Step 3. The trial duration was 12 weeks. Endpoints were change in SABA use (primary endpoint), PEF, number of poorly controlled days; and quality of life evaluated using the Paediatric Asthma Quality of Life Questionnaire (PACLQ). When the clinical outcome was "symptom-free days with no SABA use", use of eformoterol was estimated to rule-out almeterol by simple dominance, saving approximately 25 pence per patient per day whilst resulting in approximately 10 additional symptom-free days over the 12 week period.
Price and Briggs 2002[145]	Markov model based on a 12-week RCT of patients diagnosed with asthma aged 12 to 70 years (FEV_1 of 40% to 85% predicted). The main clinical outcome was the number of "successfully controlled weeks". The intervention arm was fluticasone propionate whilst the comparator was salmeterol/fluticasone propionate combination. Health states included in the model were "successful control", "hospital-managed exacerbation", "primary care managed exacerbation", "sub optimal control" and "treatment failure."
Buxton *et al* 2004[146]	Economic evaluation based on a 3-year international prospective RCT: the Steroid Treatment as Regular Therapy (START) trial. The trial compared budesonide against placebo combined with usual asthma therapy. The trial included patients from the UK, although all costs were converted to US dollars for comparability. ICERs were calculated for the UK as well as other countries where the measure of health benefit was symptom-free days (SFD). Estimates for UK costs were based on only 39 patients.
Price *et al* 2007[147]	Cost-minimisation analysis based on a 6-month, double-blind RCT. Resource use data were collected prospectively; these included medication costs and non-medication costs such as hospitalisations. The trial was international; patients were recruited from 16 countries. Costs were converted to 2004 UK costs. Patients were recruited if they were aged >12 years at the time of recruitment, and had been diagnosed with asthma at least six months previously and had been using ICS continuously for at least 3 months. Compared with using ICS alone, using budesonide/formoterol maintenance and reliever therapy (SMART) was estimated to save around £90 in costs to the NHS over the six month trial period.
Kemp *et al* 2010[148]	Economic evaluation based on a retrospective analysis of patients recorded in the UK General Practice Research Database (GPRD) from 1997 to 2007. Patients were included in the analysis if they had been registered at the same practice, had a diagnosis of persistent asthma, and had been receiving treatment with ICS. Two patient populations were identified: an initiation population who had started ICS; and a step-up population who had been prescribed an increased ICS dose. Both populations had to have been followed up for at least 12 months at their current regimen. The clinical effectiveness and cost effectiveness of three inhaler technologies were compared for these patient populations. The clinical outcome was "achieving asthma control within one year." Asthma control was defined as a composite measure involving no hospital attendance for asthma; no oral corticosteroid use; and no consultation or hospital admission or attendance related to asthma.

Health outcomes and form of economic evaluation

Of the thirteen studies included in the focussed management review, five reported QALYs gained as the measure of health benefit. One of these studies (Paggiaro *et al*[140]) was published only in the form of a conference poster, and provided very limited detail regarding the model structure. The analysis reported by Peters *et al*[136] only performed threshold analyses, indicating the necessary QALY impact to justify an incremental increase in cost. Of the eight studies which did not report QALYs, four were cost-effectiveness analyses. These studies reported health benefits in terms of:

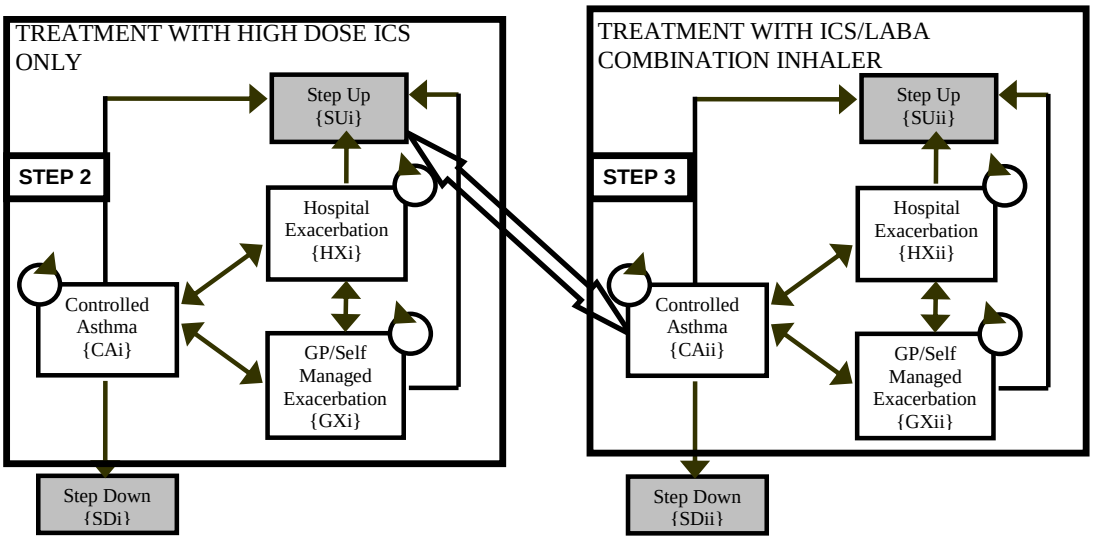
- symptom-free days with no SABA use (Everden 2002[144])
- successfully controlled weeks (Price and Briggs 2002[145])
- symptom-free days (Buxton 2004[146])
- achieving asthma control within one year (Kemp 2010[148]).

Model structures

Several of the included economic evaluations were decision analyses conducted alongside clinical trials and did not explicitly involve the use of evidence synthesis or extrapolation. Three studies used Markov structures.[137,140,145] The Briggs *et al*[137] and Price *et al*[145] studies both used similar methodologies. Each was based primarily on data from a single, though different, study. The Price *et al* model[145] categorised health states into five discrete categories: "successfully controlled", "sub-optimal control", "primary-care managed exacerbation", "hospital managed exacerbation", and "treatment failure." Treatment failure was an absorbing state; patients could transition between any of the other states during any Markov cycle. The cycle length was 1-week, and so the assumption was made that an individual could not have more than one exacerbation within one week. The time horizon of the model was 12-weeks (equal to the duration of the RCT), and the analysis did not extrapolate anticipated lifetime effects of treatment. The model used the number of exacerbation-free weeks as the measure of health benefit; this disease-specific outcome measure is difficult to interpret from a policy context. The Briggs *et al* model[137] was similar, in that it was a model based on individual patient-level data from a single trial. This model adopted four discrete health states: "totally controlled"; "well controlled", "not well controlled" and "exacerbation." These health states differ from those in the Price *et al* model[145] in that there were three non-exacerbation health states, and only one exacerbation health state. This different categorisation implicitly reflects a different set of assumptions about the key factors which influence the clinical effectiveness and cost effectiveness of different treatment options.

In addition to the studies identified by the search strategy described above, the Health Technology Assessment journal was searched from inception onwards for asthma management models. This search identified an additional asthma management model[13] which was similar in structure to the Price *et al* model.[145] The report assessed the comparative effectiveness of different ICS treatments with or without LABA for patients aged 12 years and over who had been diagnosed with chronic asthma. Unlike the Price *et al* model, this model was a cost-utility analysis and therefore measured health benefits in terms of QALYs gained. The intention of this model was to represent clinical practice, as described in the BTS/SIGN guidelines, by including different separate health states to represent dosage levels corresponding to different BTS/SIGN treatment steps. For two steps, corresponding to Step 2 and Step 3 of the BTS/SIGN guidelines, the conceptual model is shown in Figure 24. The cycle length was 1-week in duration, and the time horizon was 5 years. The key disease-specific factors included in the model relate to whether the patient experiences an exacerbation within a model cycle, and if so, the severity of the exacerbation. By allowing transitions between different levels of treatment, however, changes in treatment in response to clinical events were also incorporated.

Figure 24: Conceptual model adopted by Shepherd *et al*[13]



6.3.4 Discussion of available economic evidence on the diagnosis and management of asthma using FeNO and other interventions

The review highlights a dearth of published studies reporting on the cost-effectiveness of FeNO testing for the diagnosis and/or management of asthma. Only one published UK cost-effectiveness model of asthma diagnosis was identified and included in the review,[134] this model estimates the incremental costs of FeNO testing compared against existing standard

tests for asthma. No other cost-effectiveness models of FeNO or other diagnostic tests were identified by the searches. Similarly, the review of economic analyses of asthma management interventions identified only one UK published study of FeNO monitoring.[134] Modified versions of these FeNO management and diagnosis models were submitted to NICE by Aerocrine.[149] No evidence was submitted by Aerocrine with respect to the expected cost-effectiveness of NIOX VERO. Bedfont Scientific did not submit any economic evidence relating to the cost-effectiveness of NObreath.

The wider review of economic analyses of asthma management interventions identified a number of other economic analyses, although few were undertaken within a formal modelling framework involving evidence synthesis and/or extrapolation. These models have the following features in common: (i) the use of a Markov modelling approach with generally short cycle lengths, typically one week in duration; (ii) short time horizons; (iii) separate states for asthma exacerbations. Only two of the model-based studies reported QALYs as the measure of health outcome.

The available economic evidence for FeNO suggests that in the diagnostic setting, NIOX MINO testing may reduce the costs of diagnosis (depending on how it is used) compared to standard tests, whilst in the management setting monitoring using NIOX MINO may dominate standard guidelines. However, this evidence is subject to a number of methodological problems, questionable assumptions and weak evidence. The results of these existing analyses should be interpreted with caution.

6.4 Development of two *de novo* models to estimate the cost-effectiveness of FeNO testing for the diagnosis and management of asthma

6.4.1 Rationale for undertaking developing de novo models

This section describes the *de novo* economic models developed by the EAG to estimate the cost-effectiveness of FeNO testing (specifically using NIOX MINO, NIOX VERO or NObreath) versus standard care for the diagnosis and management of asthma. The EAG analysis involves the development of two models: (1) a *de novo* model to assess the expected cost-effectiveness of FeNO testing in addition to or in place of standard tests for the diagnosis of asthma and (2) a *de novo* model to assess the expected cost-effectiveness of FeNO plus standard guidelines versus standard guidelines for the management of patients with diagnosed asthma. Whilst these models are distinct, they form part of the same overall asthma service pathway, hence they share a number of parameter values and assumptions.

The EAG models were developed in order to attempt to resolve the problems identified with respect to the existing economic analyses of NIOX MINO (see Section 6.3) and to address gaps concerning the absence of evidence relating to the cost-effectiveness of NObreath and NIOX VERO. It should be noted that due to the limitations of the evidence base reviewed in Chapter 5, the structures of both models are necessarily simple.

The decision to develop two models rather than a single model was made because the NICE scope reflects two distinct decision problems. Whilst the FeNO devices are the same in both the diagnostic and management settings, the relevant populations and the way in which FeNO may influence decisions about appropriate clinical options for patients differs between settings; these potential effects are summarised in Table 61.

Table 61: Clinical intent of FeNO testing in diagnostic and management settings

Decision problem	Clinical population	Expected impact of FeNO
Diagnosis	Symptomatic patients with suspected asthma	FeNO testing, alone or in conjunction with other standard tests, may alter the proportion of correct and incorrect diagnoses amongst patients with suspected asthma. Changing the proportion of people with suspected asthma who are correctly/incorrectly diagnosed may then affect the expected downstream costs and health gains/losses.
Management	Patients treated for diagnosed asthma	FeNO may influence the level of medication use and the rates of exacerbations experienced by patients diagnosed with asthma. This will influence the mean costs and health gains accrued by these patients.

At the outset, the EAG had intended to model a scenario in which FeNO testing is used *both* as a diagnostic option and as a management option. This analysis was however not possible due to the necessary differences in the structures of the EAG diagnostic and management models.

6.4.2 Complexity and uncertainty surrounding the economic analysis of FeNO testing for the diagnosis and management of asthma

Given the limitations of the available evidence base (see Chapter 5), evaluating the expected cost-effectiveness of FeNO testing alone or in conjunction with other tests for the diagnosis of asthma is difficult. The BTS/SIGN guidelines for asthma diagnosis and management state that the absence of a gold standard definition of asthma makes it impossible to make evidence-based recommendations on how to make a diagnosis of asthma.[8] Further, differences in patient selection, methodological aspects of study design and the generaliseability of studies to UK practice make the unbiased interpretation of the available diagnostic evidence extremely problematic. The current diagnostic pathway is comprised of a number of tests

which may be used alone or in sequence; there does not exist a standard set of ways in which information from each of these tests should be evaluated and weighted when used together. The evidence base examined within this assessment however mostly relates to studies which estimate the operating characteristics of individual diagnostic tests used at particular points within this broader diagnostic pathway. In addition, the reference standards used within studies to estimate the sensitivities and specificities of other diagnostic tests are not always consistent or optimal, studies relate to different population groups and comparative (head-to-head) studies are few in number. As a consequence, there is considerable uncertainty surrounding the true diagnostic accuracy of FeNO and every other test used within the diagnostic pathway.

The uncertainty in the clinical evidence base is further compounded by the lack of available economic analyses. The review presented in Section 6.3 identified only one economic model of options for asthma diagnosis[134] (note - the same general model was used in the German economic evaluation reported by Berg *et al*[135]). Within this study, the authors highlight a key limitation in the scope of their analysis, that is, the analysis considers FeNO testing as a *replacement* for existing diagnostic tests; this limitation is masked somewhat by the inappropriate use of a blended comparison of multiple diagnostic tests, the absence of quantified health losses associated with misdiagnosis and the absence of a full incremental analysis. If all diagnostic tests can be substituted for one another for all patients with symptoms of asthma, as is implied by the design of the economic comparisons presented in the Price paper,[134] then the most clinically effective option will be the diagnostic test with the greatest sensitivity and specificity (depending on the balance of health losses avoided by obtaining true-positive and true-negative diagnoses). Subject to the per-test costs and the costs and consequences of downstream tests used to correct misdiagnoses, this may or may not also represent the most cost-effective option. As noted in Section 6.3, downstream costs, sequences of diagnostic pathways and consequences of incorrect diagnoses are not fully addressed by the Price/Aerocrine diagnostic model. The existing economic evidence base does not provide any information on the *additional value* of FeNO testing in conjunction with current standard tests for asthma diagnosis.

As noted by Price *et al*,[134] in reality, FeNO is likely to play a role as an adjunct to existing tests currently used within the diagnostic pathway. Whilst Price *et al* attempts to consider the combination of FeNO plus FEV_1, this is compared against the blended comparison of standard diagnostic tests thus it still represents a replacement option. Current pathways for asthma diagnosis in adults and children are complex;[8] within the Price diagnostic model, this complexity is avoided by the neat assumption that all misdiagnoses are resolved at some later

point in time with one subsequent test (i.e. following misdiagnosis, the subsequent test is assumed to have perfect sensitivity and specificity thereby correcting all previously incorrect diagnostic decisions). This is a substantial simplification. In reality, there may be a number of potential places in the existing pathway in which FeNO may provide additional diagnostic information to improve the diagnostic accuracy of current standard tests (see Figures 7 and 8) and misdiagnoses may prevail for months, years, or in some patients, indefinitely. These misdiagnoses may incur unnecessary treatment costs and health losses. The published Price/Aerocrine diagnostic models do not fully address these issues, but instead ask the question *"what is the least expensive test for the diagnosis of asthma?"*

An alternative and more sophisticated approach to evaluating the cost-effectiveness of FeNO in the diagnostic setting would involve assessing the diagnostic accuracy and cost-effectiveness of FeNO *in addition* to existing tests within the pathway. Such an analysis would address the question *"whereabouts in the existing sequences of tests, if anywhere, should FeNO be added to provide the most cost-effective diagnostic pathway for patients with symptoms of asthma?"* This would require either: (1) a similar model structure to that employed by Price *et al*[134] populated using studies which assess the accuracy of the whole diagnostic pathway for children and adults with and without FeNO testing, or (2) the development of a model which estimates the diagnostic outcomes of sequences of tests at each point in the pathway, which simulates the impact of changes in the true underlying prevalence of asthma conditional on the results of each test undertaken, and which fully takes into account the impact of potential correlations between tests which may result in non-random test outcomes in particular patients (e.g. *if Test A is negative would Test B also be negative in Patient C?*). In order for the former approach to be reliable, one would require studies which have assessed FeNO plus other tests versus a reference standard as well as the standard tests (without FeNO) versus the same reference standard, either via direct comparisons within the same study or via indirect comparisons across multiple studies with similar populations and study protocols. Price *et al* note that such data simply do not exist.[134] The review presented in Chapter 5 did however identify several studies in which FeNO was used in conjunction with other tests within *part* of the diagnostic pathway (see Sections 5.2.2.1.3 and 5.2.2.2.2). This evidence is however somewhat patchy. Interestingly, the Aerocrine diagnostic model does attempt to reflect sequential options, the latter modelling approach described above, despite the problems with the available evidence previously highlighted by Price *et al*.[134] The Aerocrine model thus assumes that sequential test outcomes are random and uncorrelated between tests. This may represent a strong assumption which could lead to biased estimates of the cost-effectiveness of FeNO; the magnitude and direction of this likely bias is unclear.

These are important limitations relating to the evidence base which constrain what can be achieved through the development of any economic model of asthma diagnosis. It would be unfair to heavily criticise any model when the main limitations of that model are principally sourced from the weaknesses in the evidence used to inform it. Such weaknesses do however limit the confidence that can and should be placed on the results of the Price/Aerocrine diagnostic models. In light of these issues, the *de novo* EAG diagnostic model attempts to resolve those weaknesses in the Price diagnostic model which can be resolved. Problems relating to the heterogeneity in the evidence base cannot be resolved by the EAG, hence the results of the *de novo* model should also be interpreted tentatively. Insofar as the available evidence allows, the EAG *de novo* diagnostic model attempts to simultaneously address the following two questions:

1. As a replacement test - *is FeNO expected to be more cost-effective than other existing tests used for the diagnosis of asthma?*
2. As an adjunctive test – *is the use of FeNO in conjunction with existing tests expected to be more cost-effective than using existing tests alone?*

The economic analysis of asthma management is subject to fewer complexities due to the availability of more robust direct evidence sourced from RCTs. There do however remain a number of methodological and evidence issues. The most notable of these relate to differences in the frequency of FeNO monitoring between the trials, uncertainty regarding the longer-term benefits of FeNO monitoring over standard care, differences between studies in terms of the step-up/step-down treatment protocols used and associated issues relating to the generalisability of non-UK treatment guidelines and symptom management strategies to UK clinical practice. The economic analysis of FeNO monitoring addresses the following question:

1. *What is the cost-effectiveness of FeNO monitoring versus standard guidelines in the management of asthma*?

6.4.3 The EAG asthma diagnostic model

6.4.3.1 Logic underpinning the diagnostic model structure

The EAG diagnostic model is hinged upon the expected costs and health losses associated with the misdiagnosis of asthma. If a patient has been misdiagnosed, this means that their treatment will not be clinically optimal until their misdiagnosis has been corrected. Misdiagnosis has different implications for those patients who are false-negative and for those patients who are false-positive. For patients who are false-positive, suboptimal treatment means receiving treatment with asthma medication which will provide no health benefit to the patient (because they do not have the underlying disease). This means there is an additional

cost to the NHS without additional health benefit for patients. Furthermore, a proportion of patients with a false-positive diagnosis of asthma may have other more serious pathology which goes undetected (e.g. cancer or tuberculosis) due to an incorrect diagnosis of asthma. Conversely, for patients who are false-negative, suboptimal treatment means not receiving treatment with asthma medication, when in reality the patient would have benefited from the treatment. Until this misdiagnosis is corrected, the patient may suffer from poor asthma control and hence lower HRQoL due to asthma symptoms without adequate treatment. Poor asthma control can a patient's HRQoL during the time they spend without experiencing an exacerbation and also by increasing the proportion of the time that a patient experiences an exacerbation. Clinically significant exacerbations are costly to the NHS, and in the case of exacerbations requiring hospitalisation these costs may be substantial; hence, a patient with undiagnosed asthma may on balance be more costly to the NHS than a patient who is correctly treated for asthma. These patients may also go on to receive expensive and unnecessary tests such as imaging and referrals to specialists until their misdiagnosis is corrected.

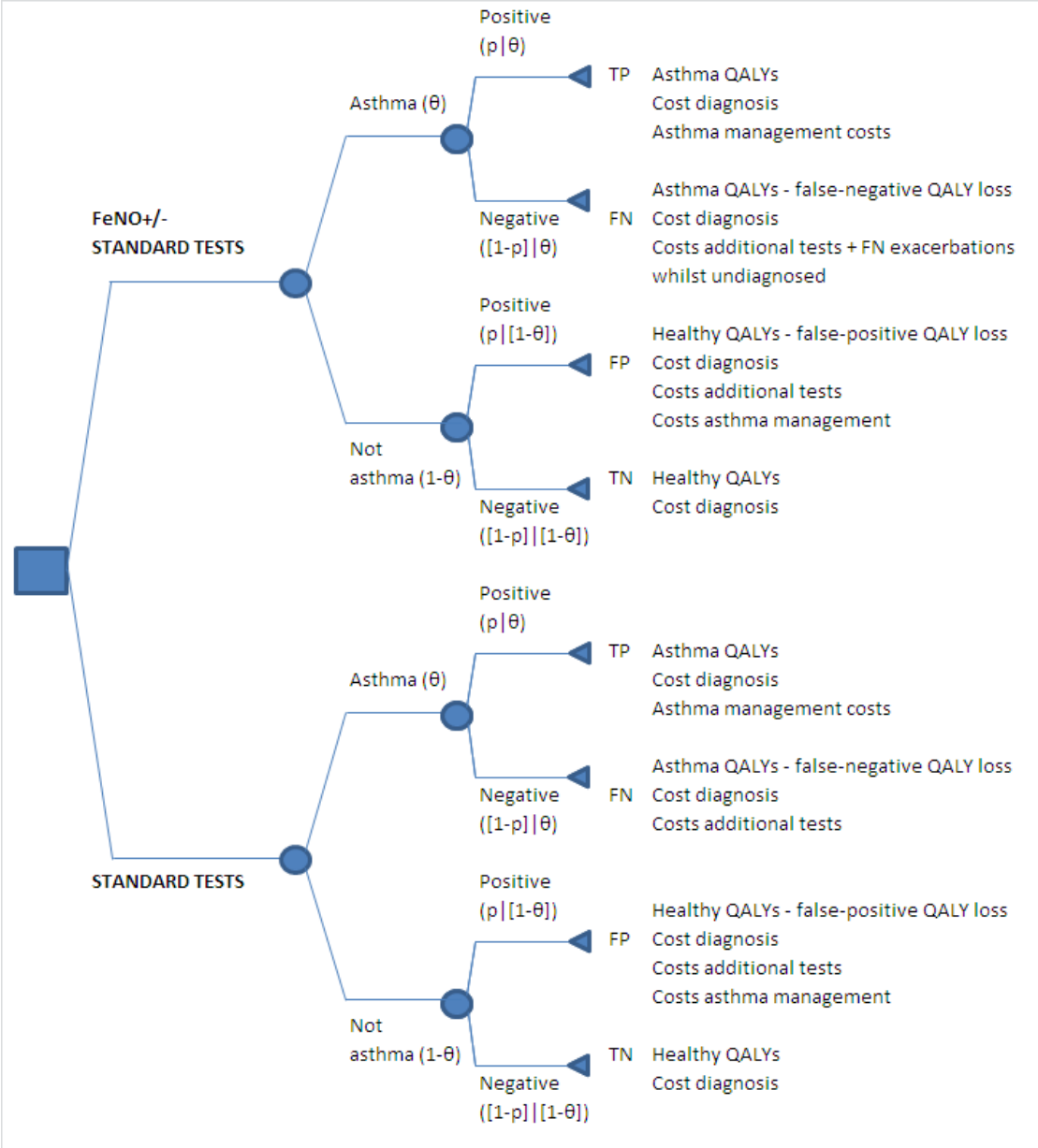
An incorrect false-negative diagnosis may be corrected later following an asthma exacerbation, due to continued asthma-related symptoms which trigger subsequent appointments and investigation, or even due to reconsideration of asthma after tests for other conditions produce negative findings. Similarly, an incorrect false-positive diagnosis may be corrected later due to the continued non-occurrence of exacerbations, a generally high level of HRQoL at very low treatment dosages thus indicating that medications currently being taken by the patient may be unnecessary, or due to continued deterioriation due to other more serious underlying pathology. Whilst it should be expected that the aggregate health consequences resulting from correct decisions should be better than those resulting from incorrect decisions, the implications for HRQoL and costs for false-positive and false-negative diagnosis outcomes are not identical. Because of this, a diagnostic strategy which maximises the area under the curve on a Receiver Operating Curve (ROC) may not necessarily yield the most cost-effective strategy. The EAG diagnostic model is therefore intended to reflect the implications of test sensitivity and specificity on subsequent costs and health consequences for the full range of diagnostic options within the available evidence base.

6.4.3.2 Model structure and assumptions

Figure 25 presents the structure of the EAG diagnostic model. The model is implemented as a simple decision tree. The population under consideration may or may not have true underlying asthma (denoted θ in Figure 25). The model then uses estimates of sensitivity and

specificity associated with each diagnostic test, or combination of tests, to estimate the expected probability that a patient will be diagnosed as having asthma or not having asthma. Therefore, the model estimates the probability that a patient with asthma will be correctly or incorrectly diagnosed as true-positive or false-negative respectively, and the probability that a patient without asthma will be correctly or incorrectly diagnosed as true-negative or false-positive respectively. The model makes the simplifying assumption that incorrect diagnoses (false-negatives and false-positives) are resolved by subsequent tests after some period of time (see Secton 6.4.5). Unnecessary treatment costs and health losses resulting from misdiagnosis are explicitly captured in the model.

Figure 25: Conceptual form of the EAG diagnostic model structure



The diagnostic model estimates costs and health outcomes for each diagnostic option across four groups:

- Patients who are true-positive (test sensitivity x prevalence) are assumed to require the initial diagnostic test(s) with no subsequent tests and are assumed to have their asthma controlled using inhaled corticosteroids (ICS) plus long-acting beta-agonists (LABA).
- Patients who are true-negative (test specificity x [1-prevalence]) are assumed to incur the cost of the initial test(s) with no subsequent tests and are assumed to have normal (general population) health status for the remainder of the model time horizon.
- Patients who are false-positive ([1-test sensitivity] x [1-prevalence]) are assumed to incur the cost of the initial test(s) with subsequent tests to correct their initial misdiagnosis. These patients and are assumed to incur a reduction in health status and incur the costs of ICS and LABA until their misdiagnosis is corrected.
- Patients who are false-negative ([1-test sensitivity] x prevalence) are assumed to incur the cost of the initial test(s) with subsequent tests to correct their initial misdiagnosis. These patients and are assumed to lose health due to poor control until their asthma is correctly diagnosed. These patients are assumed to incur asthma management costs after their asthma is diagnosed for the remainder of the model time horizon. These patients also accrue costs associated with an increased rate of exacerbations until their misdiagnosis is corrected.

The diagnostic model makes the following key structural assumptions:

- All misdiagnoses are eventually corrected within the patient's lifetime. This assumption will bias against those options with greater diagnostic accuracy. The time to correct a false-positive diagnosis may be different to the time to correct a false-negative diagnosis.
- The model time horizon for the analysis is set at 5-years. This exceeds the maximum time to correct misdiagnosis in the base case analysis (see Section 6.4.5). In effect, this reflects a lifetime horizon due to the assumption that all misdiagnoses are corrected. Health benefits gained and costs accrued after the resolution of incorrect diagnoses will be the same between all competing diagnostic decision options.
- False-negatives at initial diagnosis enjoy the same level of HRQoL after their misdiagnosis is corrected as patients who are initially correctly diagnosed as true-positives.
- False-positive results incur health losses until their misdiagnosis is corrected.

- The health consequences of other serious conditions which may be mistaken for the symptoms of asthma (e.g. lung cancer, tuberculosis, COPD etc.) are not reflected in the model.
- Patients who are false-negative may experience an increased rate of exacerbations (compared to true-positives) whilst their asthma remains uncontrolled.
- Improved diagnostic accuracy has no impact on mortality.
- All FeNO tests (NIOX MINO, NIOX VERO and NObreath) are assumed to have equivalent diagnostic accuracy.
- FeNO, spirometry and reversibility testing can be undertaken in primary care. Airway hyperresponsiveness testing (MCT) and sputum induction are undertaken in secondary care.
- Tests undertaken in primary care will involve two GP consultations and a nurse visit. Tests undertaken in secondary care will involve two attendances and a laboratory visit as well as a primary care visit for referral.
- One additional primary care visit, one laboratory visit and two additional secondary care visits are required to achieve resolution of an incorrect diagnosis.
- Owing to a lack of evidence relating to the diagnostic accuracy of each test at each point in the pathway by patient age group, the model structure is "blunt" in that differences between the diagnostic pathways for children and adults are not reflected.

Some of these assumptions are fairly strong and lack evidence to substantiate them. They are however relevant elements of the decision problem and thus require quantification. The impact of these assumptions is tested extensively in the sensitivity analysis (see Section 6.5).

Table 62 summarises the calculations underpinning the expected costs and QALY gains associated with each terminal node within the model.

Table 62: Summary of calculations of expected costs and health outcomes for each test outcome

Diagnostic test outcome	Expected cost	Expected QALY gain
True-positive	Diagnostic test costs + (time horizon.cost asthma management)	Time horizon . utility_asthma
False-positive	Diagnostic test costs + additional tests + (time to correct FP diagnosis . costs of asthma management)	((Timehorizon - time to correct FP diagnosis) . (utility_healthy – disutility_asthma)) + (Time to correct FP diagnosis . utility_ healthy)
True-negative	Diagnostic test costs	Timehorizon . utility_healthy
False-negative	Diagnostic test costs + additional tests + (time to correct FN diagnosis . cost of increased severe exacerbations) + ((timehorizon - time to correct FN diagnosis) . (costs of asthma management))	((Timehorizon-time to correct FN diagnosis) . (utility_asthma) + (time to correct FN diagnosis . disutility poor control)

6.4.3.3 Scope of the EAG diagnostic model analysis

The model is intended to reflect a population of patients with symptoms of asthma as seen in primary and secondary care in England and Wales. Table 63 details the options included in the EAG diagnostic model analysis and the setting in which these tests are assumed to be undertaken.

Table 63: Options included in the EAG diagnostic model

Test(s)	Setting
FeNO >25ppb (NIOX MINO, NIOX VERO or NObreath)	Primary care
FeNO 34ppb (using NIOX MINO, NIOX VERO or NObreath) +FEV_1	Primary care
FeNO 19ppb (using NIOX MINO, NIOX VERO or NObreath) +sputum induction	Secondary care
FeNO >27ppb (using NIOX MINO, NIOX VERO or NObreath) + bronchodilator reversibility	Primary care
FEV_1/FVC	Primary care
PEF monitoring	Primary care
Bronchodilator reversibility	Primary care
Airway hyperresponsiveness (MCT)	Secondary care
Sputum induction	Secondary care

Ppb=parts per billion; PEF=peak expiratory flow

All options are compared within a full incremental analysis. In line with the NICE Reference Case for diagnostic interventions,[132] all costs and health outcomes are discounted at a rate of 3.5%. All costs are valued at 2012/13 prices. No subgroup analyses were conducted due to evidence limitations (a narrative review of subgroup analyses within the FeNO studies is presented in Section 5.2.2.3). The base case analyses are drawn from the results of the probabilistic model and hence reflect the expectation of the mean. Further sensitivity analyses

were undertaken deterministically using point estimates of parameters. Probabilistic sensitivity analysis was used to generate information on the likelihood that each test is expected to produce the greatest net benefit over a range of willingness-to-pay thresholds (λ).

It should be noted that originally the model included an additional combination of diagnostic options - FeNO+FEV_1+bronchodilator reversibility, based on the study reported by Fortuna *et al*.[76] However, as Fortuna *et al* reported that diagnostic accuracy was not improved compared to FeNO alone, incorporating this option into the model would result in a situation whereby it has the same modelled effectiveness and same modelled cost as FeNO alone. In reality, the use of spirometry, reversibility testing and FeNO would result in a small additional cost associated with consumables and/or minor drug costs compared with FeNO alone. Consequently, given the assumption of equivalence with FeNO alone and the expectation that test costs would be marginally higher than for FeNO alone, FeNO+FEV_1+bronchodilator reversibility would always be dominated, hence this option was excluded from the final economic analysis.

6.4.4 The EAG asthma management model

6.4.4.1 Logic underpinning the management model structure

The EAG asthma management model is principally concerned with the potential benefits associated with using FeNO monitoring to enable better disease control in patients who have been diagnosed with asthma. Patients with diagnosed asthma may receive inhaled corticosteroids, LABA and other pharmacological treatments to maintain control of symptoms, to minimise the impact of the disease on patients' HRQoL and to reduce the risk of serious complications of asthma.[8] Treatment in the UK follows a stepped approach, with escalation of medication until control is reached. The incidence of exacerbations generally indicates poor asthma control; these exacerbations also impact upon patient's HRQoL and may be expensive to manage. Monitoring of FeNO levels may provide information to allow for the better control of asthma, thereby reducing unnecessary medication use in patients who do not require such treatment, maintaining medication levels where appropriate, and increasing medication use in patients with poor disease control to avoid the health losses and costs associated with exacerbations.

6.4.4.2 Model structure

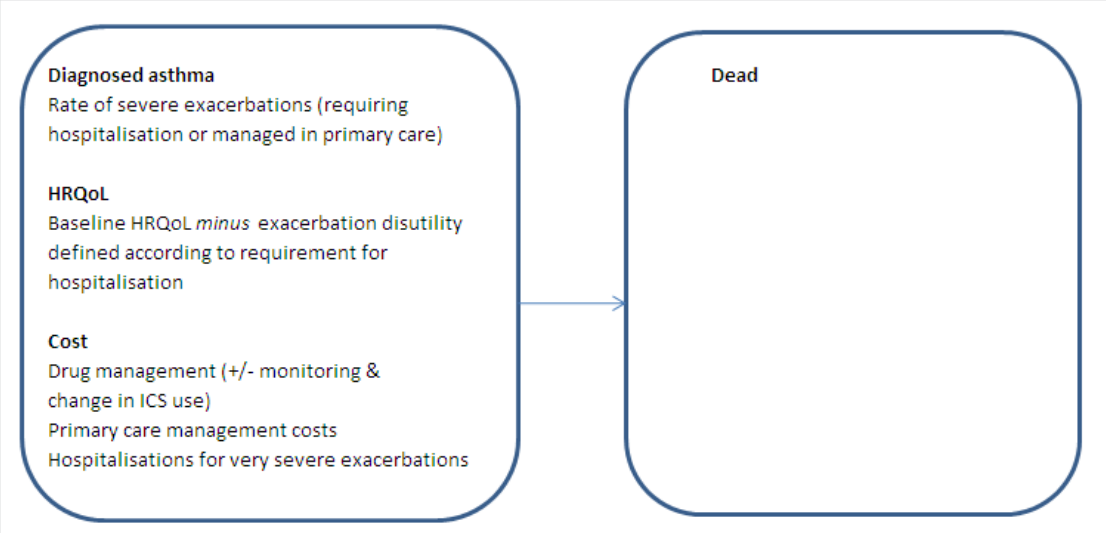
Figure 26 presents the structure of the *de novo* EAG management model. The model adopts a simple Markov framework with two states: (1) alive with diagnosed asthma and (2) dead. The model assumes that differences in HRQoL between treatment groups in the alive state are driven by the incidence of exacerbations, whilst cost differences are influenced by the

exacerbation rate and the mean level of medication use in each treatment group. Each exacerbation is associated with a reduction in HRQoL and a cost of management. Exacerbations which require hospitalisation are assumed to have a greater impact on HRQoL loss and are assumed to be more expensive to treat compared to other less severe exacerbations. Within each treatment group, the rate of exacerbations is modelled together with an estimate of required medication over time.

The management model makes the following key structural assumptions:

- Short-term impacts on exacerbations and medication use observed in the empirical studies associated with FeNO monitoring are assumed to be maintained in the longer-term (indefinitely in the base case). Given the clinical evidence used to inform the analysis, this is a strong assumption which will favour FeNO.
- Impacts of FeNO monitoring on costs and health outcomes occur only during the period in which FeNO is used (this applies only to the sensitivity analysis).
- Exacerbations are associated with a short-term reduction in HRQoL.
- The use of FeNO monitoring leads to impacts on exacerbations.
- A proportion of severe exacerbations may require hospitalisation whilst the remainder may be managed in a primary care setting. Other less severe exacerbations may be managed at home.
- Improved asthma management has no impact on mortality.
- All FeNO devices (NIOX MINO, NIOX VERO and NObreath) are assumed to have equivalent impacts on dose titration decisions in the management setting.

Figure 26: Conceptual form of the EAG asthma management model



6.4.4.3 Scope of the EAG management model analysis

The management model analysis compares the incremental costs of four options: FeNO monitoring using (i) NIOX MINO, (ii) NIOX VERO and (iii) NObreath against (iv) standard guidelines in (a) children and (b) adults. It should be noted that each FeNO option also includes the use of guidelines, as determined by the clinical evidence used to inform the exacerbation rate and ICS use parameters. The starting age for the children subgroup is assumed to be 5 years whilst the starting age for the adult group is assumed to be 18 years. The adult subgroup analysis also includes a separate subgroup analysis of FeNO monitoring in women who are pregnant. No further subgroup analyses were undertaken. The model adopts a lifetime horizon. All costs and health outcomes are discounted at 3.5%. All costs are valued at 2012/13 prices. The base case analyses are drawn from the results of the probabilistic model and hence reflect the expectation of the mean. Further sensitivity analyses were undertaken deterministically using point estimates of parameters. Probabilistic sensitivity analysis was used to generate information on the likelihood that each option is expected to produce the greatest net benefit over a range of willingness-to-pay thresholds (λ).

6.4.5 Evidence used to inform the EAG diagnostic and management model parameters

Table 64 presents the parameter values, distributions and evidence sources used to inform the two models. These are described in more detail below.

Table 64: Parameters, distributions and evidence sources used in the *de novo* EAG models

Parameter	Distribution	Mean	Param1	Param2	Source
Diagnostic model parameters					
Diagnostic accuracy					
FeNO - sensitivity	beta	0.49	75.00	79.00	Schneider *et al*[75]
FeNO - specificity	beta	0.75	180.00	59.00	
FeNO+FEV_1 - sensitivity	beta	0.24	20.00	62.00	Schleich *et al*[83]
FeNO+FEV_1 - specificity	beta	0.99	91.00	1.00	
FeNO+sputum induction - sensitivity	beta	0.87	98.31	14.69	Sivan *et al*[99]
FeNO+ sputum induction – specificity	beta	0.86	24.03	4.07	
FeNO+bronchodilator reversibility – sensitivity	beta	0.87	36.54	5.46	Cordeiro *et al*[91]
FeNO+bronchodilator reversibility – specificity	beta	0.90	64.80	7.20	
FEV_1/FVC - sensitivity	beta	0.61	41.54	26.56	Hunter *et al*[151]
FEV_1/FVC - specificity	beta	0.60	11.37	7.58	
PEF - sensitivity	beta	0.43	29.15	38.64	
PEF - specificity	beta	0.75	14.21	4.74	
Bronchodilator reversibility – sensitivity	beta	0.49	33.29	34.65	
Bronchodilator reversibility – specificity	beta	0.70	13.28	5.69	
Airway hyperresponsiveness (MCT) – sensitivity	beta	0.91	61.01	6.03	
Airway hyperresponsiveness (MCT) – specificity	beta	0.90	22.43	2.49	
Sputum induction – sensitivity	beta	0.72	48.91	19.02	
Sputum induction - specificity	beta	0.80	15.26	3.81	
Disease and population parameters					
Prevalence of true asthma	beta	0.47	412.00	469.00	Schleich *et al*,[83] Sivan *et al*,[99] Fortuna *et al*,[76] Cordeiro *et al*,[91] Schneider *et al*[75]
Probability patient is male (children)	n/a	0.55	-	-	Sivan *et al*[99]
Probability patient is male (adults)	n/a	0.40	-	-	Schneider *et al*[75]
Patient age at diagnosis (children)	n/a	5	-	-	Assumption to reflect decision problem
Patient age at diagnosis (adults)	n/a	18	-	-	
Resource cost parameters					
NIOX MINO - marginal per-test cost	n/a	£7.07	-	-	Based on information provided by Bedfont Scientific and Aerocrine
NIOX VERO - marginal per-test cost	n/a	£6.36	-	-	
NObreath - marginal per-test cost	n/a	£4.82	-	-	
Primary care GP visit	normal	£43.00	£43.00	£4.30†	PSSRU 2012[166]

Parameter	Distribution	Mean	Param1	Param2	Source
Primary care practice nurse visit	normal	£13.69	£13.69	£1.39†	
Secondary care respiratory medicine outpatient visit	normal	£204.29	£204.29	£30.64	NHS Reference Costs 2012[162]
Secondary care laboratory visit	normal	£203.29	£203.29	£30.49	NHS Reference Costs 2012[162]
Number additional primary care tests – false-positive	n/a	1.00	-	-	Structural assumptions based on expert opinion
Number additional secondary care tests - false-positive	n/a	2.00	-	-	
Number additional laboratory visits - false-positive	n/a	1.00	-	-	
Number additional primary care tests – false-negative	n/a	1.00	-	-	
Number additional secondary care tests - false-negative	n/a	2.00	-	-	
Number additional laboratory visits - false-negative	n/a	1.00	-	-	
Annual rate of additional exacerbations in uncontrolled false-negatives	normal	1.02	1.02	0.10†	Assumption based on Jayaram *et al*[157]
Annual asthma drug management costs (children)	normal	£201.00	£10.00	-	Main *et al*[167]
Annual asthma drug management costs (adults)	normal	£231.00	£10.00		Shepherd *et al*[13]
Diagnosis QALY gain/loss parameters					
Time until correct diagnosis (years) - false-positive	normal	1.50	1.50	0.26	Expert opinion
Disutility false-positive	Assumed to be equal to asthma disutility (see below)				
Time until correct diagnosis (years) - false-negative	normal	0.67	0.67	0.17	Expert opinion
Disutility poor asthma control	beta	0.04	1.39	33.35	McTaggart-Cowan *et al*[168]
Disutility asthma	beta	0.05	49.92	1027.40	Sullivan *et al*[169]
HRQoL non-asthma population	multivariate normal	0.96	-	-	Ara and Brazier[170]
Management model parameters					
Exacerbation rate parameters					
Duration of FeNO monitoring benefit (years)	n/a	Lifetime	-	-	Assumption
FeNO annual exacerbation rate (children)	lognormal	0.36	0.36	0.00	Szefler *et al*[113]
FeNO annual exacerbation rate (adults)	lognormal	0.33	0.33	0.09	Shaw *et al*[108]
Guidelines annual exacerbation rate (children)	lognormal	0.47	0.47	0.00	Szefler *et al*[113]
Guidelines annual exacerbation rate (adults)	lognormal	0.42	0.42	0.10	Shaw *et al*[108]
Management HRQoL parameters					
Disutility severe hospitalised exacerbation*	beta	0.56	1.21	3.84	Lloyd *et al*[171]
Disutility severe non-hospitalised exacerbation	beta	0.32	12.06	25.62	

Parameter	Distribution	Mean	Param1	Param2	Source
Duration severe hospitalised exacerbation (years)	gamma	0.08	15.62	0.00	Expert opinion
Duration severe non-hospitalised exacerbation (years)	gamma	0.01	12.23	0.00	
Management resource cost parameters					
Additional FeNO monitoring visits year 1	n/a	4	-	-	Assumption based on BTS/SIGN recommendations
Additional FeNO monitoring visits subsequent years	n/a	4	-	-	
r.d.i ICS use year 1 FeNO (children)	normal	0.98	0.98	0.05†	Szefler *et al*[113]
r.d.i ICS use year 2+ FeNO (children)	normal	0.97	0.97	0.05†	
r.d.i ICS use year 1 guidelines (children)	normal	0.87	0.87	0.05†	
r.d.i ICS use year 2+ guidelines (children)	normal	0.78	0.78	0.05†	
r.d.i ICS use year 1 FeNO (adults)	normal	1.20	1.20	0.05†	Shaw *et al*[108]
r.d.i ICS use year 2+ FeNO (adults)	normal	0.77	0.77	0.05†	
r.d.i ICS use year 1 guidelines (adults)	normal	1.06	1.06	0.05†	
r.d.i ICS use year 2+ guidelines (adults)	normal	1.27	1.27	0.05†	
Cost severe non-hospitalised exacerbation	normal	£44.73	£44.73	-	PSSRU[166] and BNF[172]
Cost severe hospitalised exacerbation	normal	£1,267	£1,267	£253.34	NHS Reference Costs 2012[162]

** Mean reflects additive disutility for severe non-hospitalised + severe hospitalised*
† Standard error determined subjectively
Param=parameter; FEV_1=forced expiratory volume in 1 second; FVC=forced vital capacity; r.d.i=relative dose intensity
Normal distribution – param1=mean, param2=SE; lognormal distribution – param1= mean, param2=SE; beta distribution – param1=alpha, param2=beta; gamma distribution – param1=alpha, param2 = beta; multivariate normal distribution – variance-covariance matrix not shown

Diagnostic test accuracy

Estimates of test accuracy for diagnostic tests were drawn from a number of separate studies[75,76,83,91,99,151] based on the results of the systematic review (see Chapter 5). Tables 65 and 66 summarise the sources from which these estimates were derived and the actual values selected. As far as the evidence allows, the economic analysis included studies which presented estimates of the sensitivity and specificity of individual tests as well as combinations of FeNO plus other standard tests.

The study reported by Schneider *et al*[75] was used to inform estimates of the sensitivity and specificity of FeNO alone; this study was selected due to its broad population, high study quality and because the reference standard broadly reflects the BTS/SIGN guidelines. This study used the NIOX MINO device. The study reported sensitivity and specificity across a range of cutoffs. The cutoff of >25ppb had the highest sum sensitivity and specificity (0.49

and 0.75 respectively), hence this estimate was used in the model. Additional diagnostic interventions involving FeNO plus other standard tests were included according to their availability;[76,83,91,99] the FeNO cutoff values used from these studies were driven by availability of reported estimates and were not based on choices made by the EAG. As noted above, the combination of FeNO+FEV_1bronchodilator reversibility was excluded from the final model as based on Fortuna *et al.*,[76] and the model costing assumptions, it will always be dominated by FeNO alone. Estimates of the operating characteristics of other standard tests for asthma diagnosis were drawn from Hunter *et al.*[151] This is consistent with the manufacturer's economic analysis, although it should be noted that this study may introduce bias through the use of a case-control design. Sensitivity analyses were undertaken to examine whether the use of alternative estimates[90] of the sensitivity and specificity of existing tests alters the cost-effectiveness of FeNO (see Section 6.5).

Across all diagnostic options, test operating characteristics were derived directly from data reported in the study publications. Uncertainty surrounding sensitivity and specificity estimates was modelled using independent beta distributions based on patient numbers reported in the studies.

Owing to the limitations in the evidence base, the model necessarily draws a number of unadjusted (naïve) indirect comparisons between the included studies. As a consequence, the results of the health economic analysis may be subject to bias and confounding due to differences between studies in terms of study design, recruited populations and reference standards. This same limitation is evident in the Price/Aerocrine diagnostic models and is unavoidable given the available evidence base. As the included studies did not provide sufficient information that would allow us to meaningfully discriminate between the sensitivity and specificity of all tests across population subgroups, we assumed that test operating characteristics were common to all patient populations. This assumption may not hold in reality.

Table 65: Summary of studies used to inform test accuracy parameters

Study	Study design	Population	Setting	Reference standard	Age range
Hunter *et al* (2002)[151]	Cross-sectional case control	69 asthma, 20 pseudoasthma patients and 21 healthy subjects	secondary care - single centre (UK)	Subjects with asthma had consistent clinical features, were symptomatic at the time of the investigations, had FEV_1 values of >65% of predicted, and had one or more of the following conditions: a provocative concentration of a substance (methacholine) causing more than a 20% fall in FEV_1 (PC_{20}) of < 8mg/mL; a >15% increase in FEV_1 10 min after receiving 200µg inhaled salbuterol; or a >20% maximum within-day variability of PEF when measured twice daily for >14 days.	mean age 44 years, range 15-70 years in asthma group
Smith *et al* (2004)[90]*	Prospective cohort study	47 consecutive patients referred by the GP to pulmonary function laboratory for investigation of possible asthma	secondary care (New Zealand)	Diagnosis of asthma made on basis of the following: relevant symptom history (present in all patients) using the American Thoracic Society (ATS) criteria and a positive test for bronchial hyperresponsiveness and/or a positive response to a bronchodilator.	range 8-75 years
Schleich *et al* (2012)[83]	Prospective cohort study	174 steroid naive patients with respiratory symptoms	secondary care (Belgium)	Asthma was diagnosed based on airway hyperresponsiveness demonstrated by inhaled concentration of methacholine provoking a 20% fall in FEV_1 of less than 16mg⁄ml.	range 20-59 years
Sivan *et al* (2009)[99]	Prospective, consecutive patients	150 consecutive children referred for evaluation of possible asthma	secondary care (Israel)	Patient's history of 2 or more clinical exacerbations of wheezing documented by a physician, dyspnea, or cough relieved by bronchodilators, documented variability in FEV1≥15% in response to bronchodilators at any time during the follow-up period (reversibility), or documented variability in FEV_1 ≥15% over time with or without controller medications: ICS or montelukast. Results of provocation tests were included when available. Children in whom asthma did not manifest within 18 months of follow-up were considered as not having asthma.	range 5-18 years
Cordeiro *et al* (2011)[91]	retrospective (analysis of	114 patients referred to	Secondary care (Netherlands)	History of typical respiratory symptoms and FEV_1% improvement of >12% and >200 ml, or PC_{20} histamine of ≤8 mg/ml, according to GINA guidelines	included age 7 to 83

	prospective data base)	general outpatient allergy clinic			
Schneider *et al* (2013)[75]	Prospective, consecutive cohort study	393 adults with symptoms suggestive of asthma	Germany, private practice run by 5 pneumologists	Measurements including spirometry were performed according to standard protocols (American Association for Respiratory Care, 2001) and reference values were adapted to sex, age, and height. Patients with FEV_1 < 80% predicted received salbutamol with an additional WBP investigation 20 min later. An obstructive airway disease was diagnosed if FEV_1/VC was ≤0.70. It was classified as asthma if clinical symptoms and history fitted and the change in FEV_1 was ≥12% compared to baseline and ≥200 mL and lung function returned to the predicted normal range. An incomplete bronchodilator response was stated if the response was ≤12% compared to baseline and ≤200 mL and lung volumes remained below predicted. It was classified as COPD, if clinical symptoms and history fitted and the bronchodilator response of FEV_1 after salbutamol was <12% compared to baseline and <200 mL. If there was no bronchial obstruction, bronchial provocation was performed to determine bronchial hyperresponsiveness to methacholine according to the 1- concentration-4-step dosimeter protocol. This yields similar results as the ATS multi-concentration protocol but offers advantages in clinical practice. An "asthma" diagnosis required a 20% fall in FEV_1 from baseline after inhaling methacholine stepwise until the maximum concentration (16 mg/mL), alternatively a doubling of airway resistance (Raw) and its increase to ≥2.0 kPa*s. The responsible pneumologist was blinded to the FeNO results and made the diagnostic decision only on basis of medical history, physical examination, spirometry, WBP and bronchial provocation results.	unclear (adults)

**this study is used in the sensitivity analysis only*

Table 66: Summary of test operating characteristics used in EAG model

Author	Test(s)	FeNO cutoff	Sensitivity	Specificity
Schneider *et al* 2013[75]	FeNO	>25ppb	0.49	0.75
Schleich *et al* 2012[83]	FeNO+FEV_1	>34ppb	0.24	0.99
Sivan *et al* 2009[99]	FeNO+sputum induction	>19ppb	0.87	0.89
Cordeiro *et al* 2011[91]	FeNO+bronchodilator reversibility	>27ppb	0.87	0.90
Hunter *et al* 2002[151]	FEV_1/FVC	n/a	0.61	0.60
	PEF	n/a	0.43	0.75
	Bronchial reversibility	n/a	0.49	0.70
	Airway hyperresponsiveness (MCT)	n/a	0.91	0.90
	Sputum induction	n/a	0.72	0.80

n/a=not applicable; ppb=parts per billion; TP=true positive; TN=true-negative; FP=false-positive; FN=false-negative

Disease and population parameters

The true pre-test probability of asthma in undiagnosed patients was estimated as a weighted mean of the number of cases of asthma and non-asthma diagnosed in the studies used to inform the diagnostic test accuracy parameters.[75,76,83,90,91,99] We did not include the Hunter *et al*[151] study in this calculation as it did not recruit consecutive patients due to its study design. Across these studies, 412 of 881 patients were diagnosed with asthma (probability = 0.47).

We estimated the probability that a patient is male using two studies.[75,99] These values are used only to estimate the baseline HRQoL without asthma and thus do not impact upon the model results.

Non-asthma utility

Preference-based HRQoL values for patients without asthma were estimated using a general population EQ-5D regression model reported by Ara and Brazier (*Modelled EQ-5D = 0.9508566 + 0.0212126*male -0.0002587*age - 0.0000332*age*^{2}).[170] Uncertainty surrounding this regression equation was modelled using a multivariate normal distribution. As this parameter is common to all diagnostic comparator groups, it has no effect on the estimates of incremental health gain for the diagnostic tests included in the economic analysis.

Disutility associated with asthma

The disutility associated with asthma was taken from the Catalogue of EQ-5D values reported by Sullivan *et al*.[169] Within this study, community-based UK preferences were applied to EQ-5D descriptive questionnaire responses in the US-based Medical Expenditure Panel Survey (MEPS). Sullivant *et al.* used regression models to estimate the marginal disutility associated

with a variety of diseases and conditions, assuming an additive model. Based on these models, the estimated disutility for asthma was estimated to be -0.0463. Uncertainty surrounding this parameter was modelled using a beta distribution using bootstrapped confidence intervals provided in the supplementary appendices to the paper (available from: http://mdm.sagepub.com/content/31/6/800/suppl/DC1).

This disutility is applied indefinitely to all patients with asthma and to patients without asthma who test false-positive until their misdiagnosis is corrected. It should be noted that this disutility is unlikely to fully reflect health losses associated with the delayed diagnosis of more serious pathology such as cancer or tuberculosis.

Disutility associated with poor asthma control

The impact of poor asthma control on HRQoL was informed by a recent systematic review of studies which reported the use of the EQ-5D in patients with asthma (Davis *et al*[173]). Within this review, two studies were identified which reported the impact of loss of control on patients' health status.[160,168] Within the study reported by Szende *et al*,[160] 228 consecutive adult outpatients and inpatients at four sites in Hungary completed the EQ-5D, the SF-36, the St George's Respiratory Questionnaire and a direct time trade off question. The patients' level of asthma control was determined by physicians. EQ-5D estimates are reported for four health states: "good control", "mildly reduced control, "moderately reduced control" and "poor control." EQ-5D estimates ranged from 0.93 for good control to 0.52 for poor control. Within the study reported by McTaggart-Cowan *et al*,[168] 157 asthma patients completed the Health Utilities Index Mark 3 (HUI-3), the EQ-5D and the SF-6D. The degree of asthma control was self-reported by patients. McTaggart-Cowan *et al* reported EQ-5D values for four health states: "very well controlled", "well controlled", "adequately controlled" and "not controlled." EQ-5D estimates ranged from 0.90 for very well controlled to 0.80 for not controlled. The impact of loss of control is markedly different between these two studies. As Szende *et al* recruited inpatients and outpatients, it is very likely that a number of study subjects were identified because they were experiencing an exacerbation at the time at which they completed the questionnaire; this may over-estimate their valuations of HRQoL. For this reason, we derived disutilities from McTaggart-Cowan *et al* study. We assumed that the health loss associated with poor control due to a false-negative diagnosis relates to the difference between the "well controlled" state and the "not controlled" state (mean disutility = -0.04). Uncertainty surrounding this parameter was modelled using a beta distribution based on the mean difference between the two health states; this method ensures that the notionally better health state always has a monotonically better valuation than that for the notionally worse health state.

This disutility is applied to all patients with asthma who test false-negative until their misdiagnosis is corrected.

Time to resolution of incorrect diagnoses

There is a dearth of empirical evidence relating to the time required to resolve incorrect diagnoses (false-positives and false-positives); indeed such studies would be difficult, if not impossible, to undertake prospectively. However, the time to resolve incorrect diagnosis is of direct relevance to the decision problem and must be quantified in order to evaluate the cost-effectiveness of alternative diagnostic options for asthma. Given the lack of empirical evidence relating to these parameters, we attempted to elicit these quantities from clinical experts. We asked six clinical experts (see Acknowledgements) the following questions:

1. *"For someone who has been incorrectly diagnosed as 'not asthmatic', how long on average do you think it will take for this incorrect diagnosis to be corrected? What is your 95% confidence interval around this average?"*
2. *"For someone who has been incorrectly diagnosed as 'asthmatic', how long on average do you think it will take for this incorrect diagnosis to be corrected? What is your 95% confidence interval around this average?"*

A total of four experts provided responses. One expert suggested, with considerable uncertainty, that the time to resolve a false-negative diagnosis may be in the region of 4-12 months whilst the time to resolve a false-positive diagnosis may be in the region of 12-months or longer. The expert indicated substantial uncertainty around these estimates.

The second expert stated that for false-negatives, time to correct misdiagnoses will *"mainly depend on chronicity and persistence of asthma: (a) In those with chronic persistent asthma (BTS step 2 or higher); [the] mean will only be a few weeks with relatively tight c.i., as the patient will presumably not be given treatment, will become symptomatic and demands further investigations/treatment where the true diagnosis will be revealed by other methods i.e. lung function etc. (b) In those with mild intermittent/infection induced exacerbation, it may take much longer (mean [may be] months or even year or two with [a] wide CI) as they may not get regular symptoms so the diagnosis (no asthma) may seem correct until they are exposed to the trigger and become symptomatic or get an exacerbation."*

With respect to false-positives, the second expert stated that: *"…this is even more difficult to estimate but here the means and c.i. may be in years. With an incorrect diagnosis of asthma, patients are put on treatment and they may become asymptomatic (for other reasons e.g. placebo effect,) and it is presumed that they are better because of treatment and hence*

continued on it. There is a reluctance to reduce treatment if patient[s] are doing well. This was one of the argument of using eNO (to monitor, not to diagnose), that by titrating asthma treatment with eNO you can manage airway inflammation better with lower doses of inhaled steroids."

The third expert stated that these questions were *"impossible questions to answer"* but indicated that *"misdiagnosis may never be corrected [for] both false-positive and false-negative."* In addition, the third expert stated that *"patients may make the decision themselves and just stop going back to the doctor"* and that *"asthmatic symptoms may come and go."* The expert also stated that a patient who has had asthmatic symptoms and becomes asymptomatic might be considered an asymptomatic asthmatic or may be said to have had an incorrect diagnosis of asthma by someone who sees them when well. The expert also stated that these problems are due to the absence of a reliable diagnostic test for asthma.

The fourth expert simply stated that these quantities are *"unknowable"* but did suggest that the values quantified by the first expert above were not unreasonable.

The fifth and sixth experts were not able to provide quantitative estimates.

Based on these responses, we assumed that the time to resolve a false-negative diagnosis has a mean of 8 months with a 95% confidence interval of 4-12 months. We also assumed that the time to resolve a false-positive diagnosis has a mean of 18-months with a 95% confidence interval of 12-24 months. Uncertainty surrounding these quantities was modelled using normal distributions. These estimates should be considered to be highly uncertain and are tested extensively in the sensitivity analysis.

Resource costs

Test costs

Calculating the likely marginal per-test cost for NIOX MINO, NIOX VERO and NObreath is somewhat complicated as the devices each have different lifetimes and test kits for each device are available at a lower marginal costs according higher volumes of kits purchased. The lifetimes of the NIOX MINO and NIOX VERO devices are determined either by time or by the number of tests undertaken (whichever limit is reached first).

The NIOX MINO device (Aerocrine) has a unit cost of £2,100 and an effective unit lifetime of 3-years or 3,000 tests (whichever comes first). The NIOX VERO device (Aerocrine) has a unit cost of £2,310 and an effective unit lifetime of 5-years or 5,000 tests (whichever comes

first). The NObreath device (Bedfont Scientific) costs £1,995 and, according to the manufacturer, an unlimited unit lifetime.

Maintenance for NObreath is provided free of charge. No maintenance is required for NIOX MINO or NIOX VERO.

Test kits for NIOX MINO are available in packs of 300 at a price of £1,350, 500 at a price of £2,100, or 1,000 at a price of £3,950. Test kits for NIOX VERO are available in packs of 300 at a price of £1,500, 500 at a price of £2,200 or 1,000 at a price of £4,200. Mouthpieces for NObreath are available in packs of 50, 100, 300 or 1,000 at prices of £195, £365, £995 and £2,995 respectively.

The NObreath device requires replacement of the sensor unit every 2-years at a cost of £295. Besides test kits, NIOX MINO and NIOX VERO do not require any further consumables or replacement costs.

Based on information provided by Bedfont Scientific and Aerocrine, Table 67 presents the estimated annuatised marginal per-test costs assuming a usage of 300 tests per device per year (this estimate is based on estimates of mean usage provided by Aerocrine). All calculations are based on the lifetime of the specific device and the lowest cost estimates for the required number of test kits at the assumed level of throughput and lifetime of the device. We assumed that whilst the NObreath device has an unlimited life, advances in technology would lead to replacement of the device within 10-years. Annuatisation was undertaken assuming a rate of 3.5%.

Table 67: Marginal per-test costs for FeNO devices

	NIOX MINO	**NIOX VERO**	**NObreath**
Lifetime (years)	3	5	10
Total tests assumed/year	300	300	300
Equipment	£2,100.00	£2,310.00	£1,995.00
Test kits (1000 mouthpieces)	£3,950.00	£4,200.00	£2,995.00
Test kits (500 mouthpieces)	£2,100.00	£2,200.00	n/a
Test kits (300 mouthpieces)	£1,350.00	£1,500.00	£995.00
Test kits (100 mouthpieces)	n/a	n/a	£365.00
Sensor replacements	n/a	n/a	£295.00
Total cost over device lifetime	£6,150.00	£8,910.00	£12,455.00
Annuatisation factor for specific device lifetime	2.90	4.67	8.61
Annuatised marginal per-test cost	**£7.07**	**£6.36**	**£4.82**

It should be noted that these marginal per-test costs do not include any costs associated with education and training that may be required to teach NHS staff how to instruct patients how to use the device correctly in order to minimise test failure rates (see Chapter 7).

We assumed that spirometry, reversibility testing and FeNO can be done in primary care and would require two GP visits and one nurse visit. We assumed that sputum induction and airway hyperresponsiveness (MCT) would be undertaken in secondary care and would require two secondary care visits, one laboratory visit as well as an initial GP visit for referral (*personal communication: Dr John White, 17th July 2013*).

The unit cost of a GP visit was taken from Curtis *et al*;[166] the economic analyses uses an estimate of £43 which reflects the cost of an appointment lasting 11.7 minutes including direct staff costs and qualifications. The cost of a GP practice nurse visit was assumed to be £13.69 assuming a visit duration of 15.5 minutes based on the same source.[166] Secondary care attendance costs were based on the HRG for respiratory medicine attendances (cost = £204.29).[162] The cost of a laboratory visit was based on the HRG for simple bronchodilator studies (cost =£203.29). We assumed standard errors around these estimates were normally distributed with a standard error equal to 15% of the mean.

As HRGs are calculated using full economic costing, we assumed that all visit costs include the costs associated with capital, training, staff costs and procedure costs associated with all existing diagnostic tests for asthma. For the strategies which include FeNO testing, the marginal per-test cost of FeNO was added to these visit costs (see Table 67).

Costs associated with resolving misdiagnoses

We assumed that incorrect diagnoses would be resolved at a later point in time. We crudely assumed that one additional primary care attendance, two additional secondary care attendances and one laboratory visit would be required to correctly diagnose false-positive results and false-negative results. This is an assumption and should be interpreted with some caution.

Costs associated with loss of control for false-negatives

The model assumes that patients with asthma who initially test negative experience an increased rate of exacerbations compared to true-positive patients who are correctly diagnosed and receive treatment. It is likely that ethical implications associated with the design of an empirical research study to collect this information would be prohibitive. We assumed that false-negative patients would experience one exacerbation each year in which

they remain undiagnosed; this was loosely based on the higher absolute exacerbation estimate for diagnosed patients reported by Jayaram *et al*.[157] The model assumes that a proportion of these exacerbations will require hospitalisation (see below).

Costs of asthma management

We assumed that, on average, patients would be at Step 3 in the BTS/SIGN asthma guidelines. Current technology appraisal guidance from NICE on the use of ICS for children and adults recommends that the least expensive option is used and does not differentiate between drugs in terms of effectiveness. We derived estimates of the annual cost of combined inhalers from two previous health technology assessment reports.[13,167] Main *et al* estimated the least expensive annual cost for combined inhalers to be £201 for children (Symbicort Turbohaler).[167] Shepherd *et al* estimated the least expensive annual cost of combined inhalers to be £231 for adults[13] (Symbicort Turbohaler). Scrutiny of the current version of the BNF indicates that the annual cost of these inhalers has not changed since the original HTA reports were published.

Additional management model parameters

General population mortality

The probability of dying from all causes was taken directly from current interim life tables and was applied according to the proportionate split of males to females with asthma.[174]

Duration of NIOX benefit

In the base case analysis we assumed that the impact of FeNO monitoring on dose titration and exacerbations would be retained indefinitely over the patient's lifetime. Whilst this is plausible, there is no long-term RCT evidence to support or refute this assumption. We examine the impact of this assumption within the sensitivity analysis.

Annual exacerbation rates with FeNO or standard care

Annual exacerbation rates with and without FeNO testing were derived from the RCT reported by Szefler *et al*[113] for children and from the RCT reported by Shaw *et al*[108] for adults. Changes in ICS use with/without FeNO monitoring for children and adult subgroups were also drawn from these trials.

The RCT reported by Shaw *et al* was selected for use in the adult subgroup as it was the only UK-based study included in the systematic review for adults (see Section 5.2.3.1), because it reflects BTS/SIGN guidelines,[8] and because it reported data on severe exacerbation rates and changes in ICS use (the relevant parameters for the model). The population within this RCT

relates to adult non-smokers and never smokers who were deemed to be compliant with medication and who had not experienced a severe exacerbation within 4-weeks of study entry. This allowed for the inclusion of a broader range of severity compared with the other studies. Patients were aged between 20 and 81 years of age and were treated and followed up for 12-months.

Of the studies included in the systematic review for children (see Section 5.2.3.2), the study reported by Szefler *et al*[113] appears to most closely reflect current UK practice, hence this study was selected to inform the exacerbation rates and ICS use parameters for the children's subgroup. Within this study, patients were either on long-term control treatment with symptoms of persistent asthma or evidence of uncontrolled disease, or not on long-term control treatment with symptoms of persistent asthma and evidence of uncontrolled disease. Patients were treated and followed up for 46 weeks. This trial was undertaken in the US.

Szefler *et al*[113] reported that 32.1% (s.d.=4.67) of 276 patients in the FeNO group and 42.0% (s.d.=4.94) of 270 patients in the control group received ≥1 courses of prednisone over the 46 week study period; this was taken as a proxy for severe exacerbations. The authors also report that 3.3% (s.d.=1.78) of patients in the FeNO group and 4.1% (s.d.=1.98) of patients in the control group were hospitalised at least once. We used these data to estimate the annual rate of exacerbations for each arm (0.36 for FeNO and 0.47 for standard care). It should be noted that the data available in the paper relate to the number of patients experiencing exacerbations events rather than the number of exacerbations events. We calculated the probability that an exacerbation required hospitalisation by pooling the exacerbation and hospitalisation data for the two study arms (probability = 0.04).

Shaw *et al*[108] reported 18 exacerbations in 12 patients in the 58 FeNO group patients and 26 exacerbations in 19 patients in the 60 control group patients over 42 weeks. This corresponds to an annual exacerbation rate per patient of 0.33 (s.d.=0.69) for the FeNO group and 0.42 (s.d.=0.79) for the control group. Shaw *et al* did not report the proportion of severe exacerbations requiring hospitalisation and so this probability was assumed to be the same as that observed in the Szefler *et al* study detailed above.

Exacerbation rates were assumed to follow a lognormal distribution. The probability that an exacerbation requires hospitalisation was modelled using a beta distribution.

Impacts of exacerbation on HRQoL

The impact of exacerbations was based on a valuation study reported by Lloyd *et al*;[171] this study was identified from the systematic review reported by Davis *et al*.[173] Lloyd *et al* report the impact of exacerbations on HRQoL in patients with moderate to severe asthma (BTS levels 4 and 5) in the UK. Within this study, 112 patients completed a variety of health status questionnaires including the EQ-5D. Disutilities associated with severe non-hospitalised and severe hospitalised exacerbations were calculated based on the differences between the valuations for the three states. Uncertainty surrounding these parameters was modelled using beta distributions based on the difference between the two health states; this method ensure that the notionally better health state always has a monotonically better valuation than that for the notionally worse health state. The disutility of a severe exacerbation resulting in hospitalisation (compared to no exacerbation) was estimated to be -0.56 whilst the impact of other exacerbations which do not result in hospitalisation (compared to no exacerbation) was estimated to be -0.32. Disutilities are assumed to be additive and are therefore not influenced by the baseline level of HRQoL.

Severe exacerbations not resulting in hospitalisation were assumed to last for 4 days whilst major exacerbations resulting in hospitalisation were assumed to last for 4 weeks. These quantities were based on subjective estimates provided by experts. These durations were assumed to follow gamma distributions with standard errors fitted to capture the range of estimates elicited (2-6 weeks for exacerbations requiring hospitalisation and 3-7 days for other severe exacerbations).

Resource costs

Additional costs of FeNO monitoring

We assumed that FeNO monitoring would be undertaken during routine GP visits and would require one additional nurse visit once every 3 months.[8] The marginal cost of FeNO monitoring was applied as the per-test cost plus the cost of a primary care nurse appointment.

Changes in medication (ICS) use over time

We derived estimates of change in ICS use with and without FeNO monitoring in children from the RCT reported by Szefler *et al*[113] and in adults from Shaw *et al*.[108] We assumed that during the period for which ICS use was observed in each study (12-months in Shaw and 46-weeks in Szefler), ICS use would reflect the observed mean, with relative dose intensity (r.d.i.) calculated as the mean over the observed period divided by the baseline ICS dosage for each study arm. Beyond this point, we assumed that ICS use would remain constant at the

level of the last observation for each study arm for the remainder of the duration over which FeNO impacts on exacerbations and titration decisions (see Table 68).

Table 68: Estimated ICS dose (relative to baseline)

Parameter	FeNO	Guidelines
Children (Szefler et al[113]		
Mean RDI first 39-weeks	0.98	0.87
Mean RDI subsequent	0.97	0.78
Adults (Shaw et al[108])		
Mean RDI first 12-months	1.20	1.06
Mean RDI subsequent	0.77	1.27

Costs of managing exacerbations

We assumed that a proportion of exacerbations would require hospitalisation whilst the remainder could be managed in primary care. We assumed that severe exacerbations which do not require hospitalisation would require one GP attendance (cost=£43.00) plus oral steroids for 5 days (cost = £1.73) based on an earlier HTA report.[167] We derived the cost of asthma hospitalisation from current NHS Reference Costs (cost = £1,266.72).

6.4.6 Model evaluation

The model was evaluated probabilistically using standard Monte Carlo sampling techniques over 5,000 random samples. Central estimates of cost-effectiveness are presented based on upon the expectation of the mean. Headline results are presented as incremental cost-effectiveness ratios (ICERs), cost-effectiveness planes and cost-effectiveness acceptability curves (CEACs). In addition, a large number of deterministic sensitivity analyses (DSA) were undertaken; these analyses are detailed below. All incremental analyses were calculated using an automated tabular algorithm developed by one of the study authors (PT).

Deterministic sensitivity analysis undertaken using the diagnostic model

DSA scenario D1 – point estimates of parameters

The model was evaluated using point estimates of parameters rather than the expectation of the mean.

DSA scenarios D2 and D3 – alternative discount rates

The model was evaluated with discount rates of 0% for costs and QALYs (DSA scenario D2) and 6% for costs and QALYs (DSA scenario D3).

DSA scenario D4 – all tests undertaken in secondary care

The model was run assuming that all tests are undertaken in a secondary care setting.

DSA scenarios D5 and D6 – alternative asthma control disutilities for false-negatives

The model was run assuming different disutilities for patients who are false-negative. In DSA scenario D5, the most extreme disutility from McTaggart-Cowan *et al*[168] was assumed ("very well controlled" to "not controlled" state disutility=-0.10). In DSA scenario D6, the most extreme disutility from Szende *et al*[160] was assumed ("good control" to "poor control" state disutility=-0.41).

DSA scenarios D7 and D8 – alternative disutilities for false-positives

The model was run assuming different disutilities for patients who are false-positive. In DSA scenario D7, the base case disutility applied to false-positives was doubled, whilst in DSA scenario D8, this disutility was halved.

DSA scenarios D9 and D10 – FeNO test costs

The model was evaluated assuming that the marginal per-test costs for all FeNO devices (NIOX MINO, NIOX VERO and NObreath) are double (DSA scenario D9) or half (DSA scenario D10) those assumed in the base case analysis.

DSA scenarios D11-D13 – Alternative assumptions concerning NObreath device lifetime

Within the base case analysis, the NObreath device is assumed to have a fixed lifetime of 10-years (for costing purposes). In DSA scenario D11, the analysis is repeated assuming a maximum lifetime for NObreath of 3-years (equal to the maximum lifetime of the NIOX MINO device). In DSA scenario D12, the analysis is repeated assuming a maximum lifetime for NObreath of 5-years (equal to the maximum lifetime of the NIOX VERO device). In DSA scenario D13, the analysis is repeated assuming a maximum lifetime for NObreath of 20-years (double that assumed in the base case). These alternative assumptions result in marginal per-test costs for NObreath of £14.32, £8.88 and £2.32 for DSA scenarios D11, D12 and D13, respectively.

DSA scenarios D14 and D15 – test visit costs

The model was evaluated assuming that all primary and secondary care visit costs are double (DSA scenario D14) or half (DSA scenario D15) those assumed in the base case analysis. This includes the costs of initial visits and subsequent visits to resolve misdiagnosis.

DSA scenarios 16 and 17 – false-negative exacerbation rate

The model was evaluated assuming that the base case incremental exacerbation rate for false-negatives is double (DSA scenario D16) or half (DSA scenario D17) that assumed within the base case analysis.

DSA scenarios D18 and D19 – asthma treatment costs

The model was evaluated assuming that asthma treatment costs are double (DSA scenario D18) or half (DSA scenario D19) those assumed in the base case analysis.

DSA scenarios D20 to D25 – time to resolve misdiagnosis

The model was evaluated assuming a range of different assumptions regarding the time to resolve initial misdiagnoses (both false-negatives and false-positives). The model was evaluated assuming 2x, 3x, 4x, 5x, 10x and 0.5x the base case time to correct diagnosis parameters in DSA scenarios D20 to D25 respectively. In these analyses, the time horizon was set equal to the maximum time to resolve false-positive and false-negative results (note that this does not affect the incremental model results).

DSA scenarios D26 and D27 – alternative sources for diagnostic accuracy of FeNO alone

The base case analysis used estimates of the diagnostic accuracy for FeNO from Schneider *et al*.[75] In DSA scenario D26, the model was evaluated using alternative estimates from Schleich *et al*;[83] at a cutoff of 34ppb, the sensitivity and specificity of FeNO was 0.35 and 0.95 respectively. In DSA scenario D27, the model was evaluated using alternative estimates from Pedrosa *et al;*[89] at a cutoff of 40ppb, the sensitivity and specificity of FeNO was 0.74 and 0.73 respectively. Both of these studies reflect a difficult to diagnose population, although it should be noted that this is not the case for the other comparators in this scenario analysis.

DSA scenario D28 – alternative source for diagnostic accuracy of non-FeNO comparators

The base case analysis draws estimates of sensitivity and specificity for individual comparators from the study reported by Hunter *et al*.[151] In DSA scenario D28, the model was evaluated using estimates for FEV_1/FVC, PEF and sputum induction from the comparative diagnostic study reported by Smith *et al*.[90]

DSA Scenarios D29-D31 – "rule-out" diagnostic decision approach

In DSA scenarios D29 to D31, the model was evaluated assuming a "rule-out" diagnostic approach for all diagnostic tests. In these scenarios, any patient who tests negative is "ruled-out" and is treated as being not asthmatic (as per the base case structure), whilst any patient testing positive is assumed to immediately undergo further tests to confirm their diagnosis. As

a consequence, no patient loses health due to initially testing false-positive. The "rule-out" approach was evaluated over three scenarios: DSA scenario D29 – base case test characteristics for FeNO options; DSA scenario 30 – best sensitivity for FeNO options; DSA scenario 31 – best specificity for FeNO options.

DSA scenarios D32-D34 – "rule-in" diagnostic decision approach

In DSA scenarios D32 to D34, the model was evaluated assuming a "rule-in" diagnostic approach for all tests. In this scenario, any patient who tests positive is "ruled-in" and is treated as being asthmatic (as per the base case structure), whilst any patient testing negative is assumed to immediately undergo further tests to confirm their diagnosis. As a consequence, no patient loses health due to initially testing false-negative. The "rule-in" approach was evaluated over three scenarios: DSA scenario D32 – base case test characteristics for FeNO options; DSA scenario 33 – best sensitivity for FeNO options; DSA scenario 34 – best specificity for FeNO options.

Deterministic sensitivity analysis undertaken using the management model

DSA scenario M1 – point estimates of parameters

The model was evaluated using point estimates of parameters rather than the expectation of the mean.

DSA scenarios M2 and M3 – alternative discount rates

The model was evaluated with discount rates of 0% for costs and QALYs (DSA scenario M2) and 6% for costs and QALYs (DSA scenario M3).

DSA scenario M4 – pregnant women subgroup analysis

In this scenario the model was evaluated specifically for a subgroup comprising women who are pregnant. This analysis was based on the RCT reported by Powell *et al.*[111] We estimated annual exacerbation rates of 0.58 and 1.26 for the FeNO and guidelines groups respectively. Mean ICS use over the study period was estimated to be approximately by 77% of the baseline dose for the FeNO group and 102% of the baseline dose for the guidelines group. These estimates were assumed to apply for the first 5 months (the observed period in the trial). The final observations of 73% and 105% for FeNO and guidelines respectively were assumed to be carried forward over the remainder of the time horizon.

DSA scenario M5 – alternative source of exacerbation rates and ICS use for children

In DSA scenario M5, the model was evaluated using alternative estimates of exacerbation rates and ICS use over time, based on the RCT reported by Pijnenburg *et al.*[115] We estimated

exacerbation rates of 0.18 and 0.39 for the FeNO and guidelines groups respectively. ICS use over the 1-year follow-up period was similar in both groups; relative dose intensity versus the baseline dose was estimated to be 1.16 in both groups. Beyond the first-year, r.d.i. was estimated to be 1.23 for the FeNO group and 1.22 for the guidelines group.

DSA scenario M6 and M7 – alternative source of exacerbation rates and ICS use for adults

In DSA scenarios M6 and M7, the model was evaluated using alternative estimates of exacerbation rates and ICS use over time, based on the RCTs reported by Smith *et al*[125] (DSA scenario M6) and Syk *et al*[109] (DSA scenario M7).

Using Smith *et al*,[125] we estimated severe exacerbation rates of 0.16 and 0,17 for the FeNO and guidelines groups respectively. ICS use in the first year, relative to the first observation, was estimated to be 0.85 and 1.08 for FeNO and standard care respectively. ICS use based on the last observation was estimated to be 0.90 and 1.30 for the FeNO group and guidelines group respectively.

Using Syk *et al*,[109] we estimated exacerbation rates of 0.09 and 0.07 for the FeNO and guidelines groups respectively. ICS use in the first year, relative to the first observation, was estimated to be 0.97 and 0.96 for the FeNO group and the guidelines group respectively. ICS use based on the last observation was estimated to be 0.88 and 0.99 for the FeNO group and the guidelines group respectively.

DSA scenarios M8-M17 – Alternative assumptions regarding duration of FeNO impact

A number of scenarios were undertaken to examine the impact of assuming alternative durations over which FeNO would impact upon ICS use and exacerbations. We examined the following durations - 1 year, 2 years, 3 years, 4 years, 5 years, 10 years, 15 years, 20 years, 30 years and 40 years.

DSA scenarios M18 and M19

The model was evaluated assuming that the marginal per-test costs for all FeNO devices (NIOX MINO, NIOX VERO and NObreath) are double (DSA scenario M18) or half (DSA scenario M19) those assumed in the base case analysis.

DSA scenarios M20-M22 – Alternative assumptions concerning NObreath device lifetime

In DSA scenario M20, the model was evaluated assuming a maximum lifetime for NObreath of 3-years (equal to the maximum lifetime of the NIOX MINO device). In DSA scenario M21, the analysis is repeated assuming a maximum lifetime for NObreath of 5-years (equal to

the maximum lifetime of the NIOX VERO device). In DSA scenario M22, the analysis is repeated assuming a maximum lifetime for NObreath of 20-years (double that assumed in the base case). These result in marginal per-test costs for NObreath of £14.32, £8.88 and £2.32 for DSA scenarios M20, M21 and M22, respectively.

DSA scenarios M23 and M24

The model was evaluated assuming that the number of nurse visits for the FeNO group was double (DSA scenario M23) or half (DSA scenario M24) the number applied in the base case analysis.

DSA scenarios M25 and M26 – alternative assumptions regarding exacerbation rates

The model was evaluated assuming that the exacerbation rates for the FeNO and standard care groups are equal to double (DSA scenario M25) and half (DSA scenario M26) those rates assumed in the base case analysis.

DSA scenarios M27 and M28 – alternative assumptions regarding exacerbation disutility

The model was evaluated assuming that the exacerbation disutilities for the FeNO and standard care groups are equal to double (DSA scenario M27) or half (DSA scenario M28) those disutilities assumed in the base case analysis.

DSA scenario M29 – mean observed ICS use projected forward

The model was evaluated assuming that mean ICS use observed within the clinical trials is maintained over the remainder of the model time horizon.

DSA scenarios M30 and M31– alternative assumptions regarding ICS change over time

The model was evaluated assuming that the mean RDI for ICS in the FeNO and standard care groups is equal to double (DSA scenario M30) and half (DSA scenario M31) that assumed in the base case analysis.

With the exception of DSA scenarios M4-M7, all deterministic sensitivity analyses within the management model were undertaken in both the children and adult subgroups.

6.4.7 Model validation methods

We took a number of measures to ensure the credibility of the models and their results. The conceptual models were discussed extensively amongst the EAG prior to implementation. The lead modeller (PT) checked the integrity of all model calculations and VBA programming whilst developing the model. The models were re-checked once they were

complete. PT also re-built deterministic versions of both models in a more disaggregated form to ensure that all calculations were implemented as intended; these replicated models gave exactly the same results as the full models. All model input parameters and pre-model analyses were checked and inputted values were compared against the sources from which they were derived. The results of the models were compared against our *a priori* expectations, given the model structures and input parameters, and any discrepancies were investigated. A large number of sensitivity analyses and black-box tests were undertaken to ensure that the models were behaving as expected. Finally, the assessment report was peer reviewed by clinical experts, other researchers within ScHARR and by NICE (see Acknowledgements).

6.5 *De novo* EAG model results

6.5.1 Diagnostic model results (all patients)

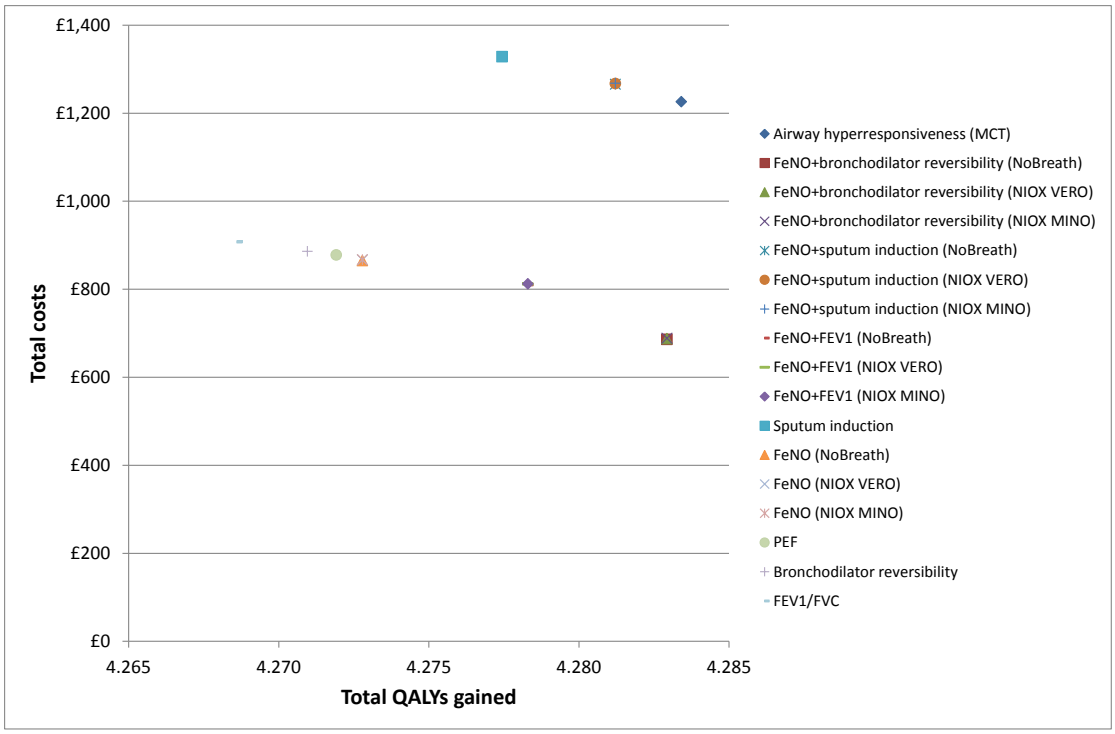
6.5.1.1 Central estimates of cost-effectiveness – diagnosis (all patients)

Table 69 presents the central estimates of cost-effectiveness based on the probabilistic version of the diagnostic model. The results suggest that across the 17 diagnostic options included in the economic analysis, the expected difference in QALY gains is likely to be very small. NIOX MINO and NIOX VERO, alone or in combination with other tests, are expected to be dominated as their marginal per-test cost is higher than that for NObreath. Airway hyperresponsiveness (MCT) is expected to produce the greatest QALY gain; this is because this option has both the highest sensitivity and specificity of all tests included in the economic analysis. With the exception of FeNO+bronchodilator reversibility, all other options are expected to be ruled out by simple dominance. The incremental cost-effectiveness of airway hyperresonsiveness versus FeNO+bronchodilator reversibility is expected to be approximately £1.125million per QALY gained. This information is presented on the absolute cost-effectiveness plane presented in Figure 27.

Table 69: Central estimates of cost-effectiveness - diagnosis

Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
Airway hyperresponsiveness (MCT)	4.2834	£1,226.00	0.0005	£539.92	**£1,125,074**
FeNO+bronchodilator reversibility (NoBreath)	4.2829	£686.08	-	-	-
FeNO+bronchodilator reversibility (NIOX VERO)	4.2829	£687.61	-	-	**dominated**
FeNO+bronchodilator reversibility (NIOX MINO)	4.2829	£688.33	-	-	**dominated**
FeNO+sputum induction (NoBreath)	4.2812	£1,265.78	-	-	**dominated**
FeNO+sputum induction (NIOX VERO)	4.2812	£1,267.32	-	-	**dominated**
FeNO+sputum induction (NIOX MINO)	4.2812	£1,268.03	-	-	**dominated**
FeNO+FEV1 (NoBreath)	4.2783	£810.14	-	-	**dominated**
FeNO+FEV1 (NIOX VERO)	4.2783	£811.67	-	-	**dominated**
FeNO+FEV1 (NIOX MINO)	4.2783	£812.38	-	-	**dominated**
Sputum induction	4.2774	£1,328.28	-	-	**dominated**
FeNO (NoBreath)	4.2728	£865.06	-	-	**dominated**
FeNO (NIOX VERO)	4.2728	£866.60	-	-	**dominated**
FeNO (NIOX MINO)	4.2728	£867.31	-	-	**dominated**
PEF	4.2719	£877.91	-	-	**dominated**
Bronchodilator reversibility	4.2710	£886.27	-	-	**dominated**
FEV1/FVC	4.2686	£907.71	-	-	**dominated**

Figure 27: Cost-effectiveness plane – diagnosis (all patients)



6.5.1.2 Uncertainty analysis – diagnosis (all patients)

Figure 28 presents CEACs for the diagnostic options. Assuming a willingness-to-pay threshold of £20,000 per QALY gained, FeNO+bronchodilator reversibility (using NObreath) has the highest probability of producing the greatest amount of net benefit (probability = 0.98). Assuming a willingness to pay threshold of £30,000 per QALY gained, FeNO+bronchodilator reversibility (using NObreath) also has the highest probability of producing the greatest amount of net benefit (probability = 0.95). These results are also summarised in Table 70.

Figure 28: Cost-effectiveness acceptability curves – diagnosis (all patients)

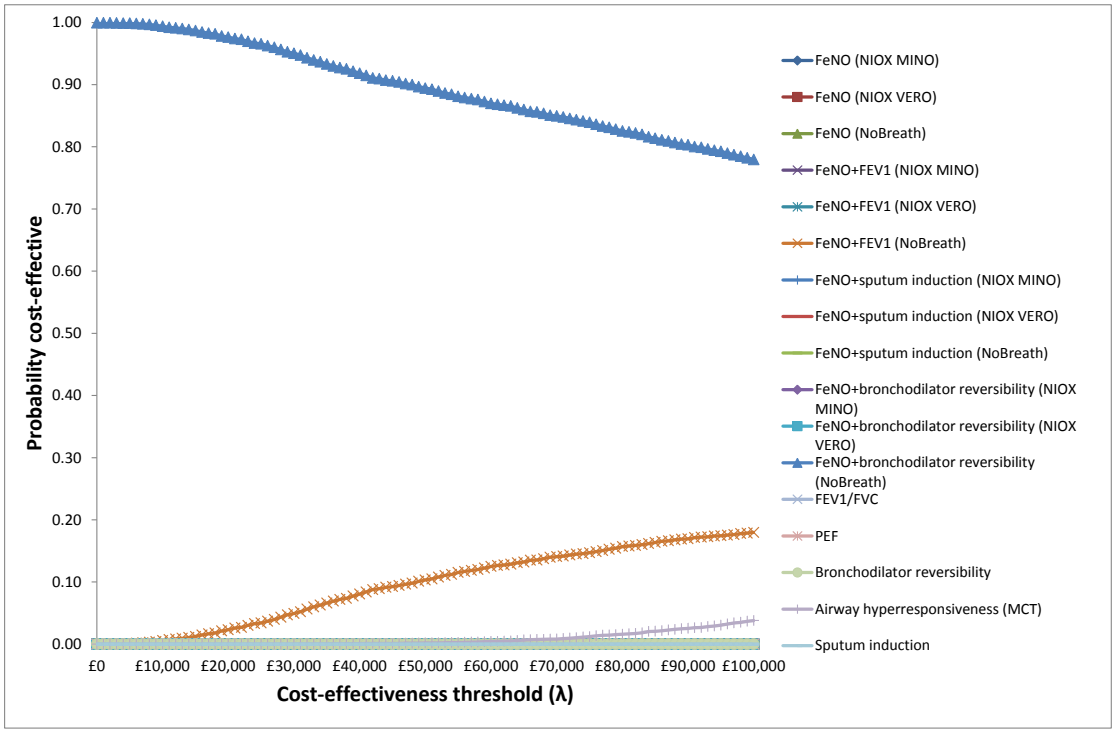


Table 70: Probability of optimality – diagnosis (all patients)

Option	Probability optimal λ=£20,000/QALY gained	Probability optimal λ=£30,000/QALY gained
FeNO (NIOX MINO)	0.00	0.00
FeNO (NIOX VERO)	0.00	0.00
FeNO (NObreath)	0.00	0.00
FeNO+FEV_1 (NIOX MINO)	0.00	0.00
FeNO+FEV_1 (NIOX VERO)	0.00	0.00
FeNO+FEV_1 (NObreath)	0.02	0.05
FeNO+sputum induction (NIOX MINO)	0.00	0.00
FeNO+ sputum induction (NIOX VERO)	0.00	0.00
FeNO+ sputum induction (NObreath)	0.00	0.00
FeNO+bronchodilator reversibility (NIOX MINO)	0.00	0.00
FeNO+bronchodilator reversibility (NIOX VERO)	0.00	0.00
FeNO+bronchodilator reversibility (NObreath)	**0.98**	**0.95**
FEV_1/FVC	0.00	0.00
PEF	0.00	0.00
Bronchial reversibility	0.00	0.00
Airway hyperresponsiveness (MCT)	0.00	0.00
Sputum induction	0.00	0.00

6.5.1.3 Deterministic sensitivity analysis – diagnosis (all patients)

Tables 71 to 76 present the results of the deterministic sensitivity analysis. In all analyses, the rank ordering of non-dominated options is maintained except where indicated by parentheses [ ].

Table 71: Deterministic sensitivity analysis – diagnosis (all patients) scenarios D1-D6

DSA scenario	D1 Point estimates of parameters	D2 Undiscounted costs and outcomes	D3 Discount rate=6%	D4 All tests in secondary care	D5 Disutility from McTaggart Cowan	D6 Disutility from Szende
Airway hyperresponsiveness (MCT)	£1,094,325	£1,081,089	£1,103,827	dominating	£437,730	£106,763
FeNO+bronchodilator reversibility (NoBreath)	-	-	-	dominated	-	-
FeNO+bronchodilator reversibility (NIOX VERO)	dominatedinated	dominated	dominated	dominated	dominated	dominated
FeNO+bronchodilator reversibility (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NoBreath)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NoBreath)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
Sputum induction	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NoBreath)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
PEF	dominated	dominated	dominated	dominated	dominated	dominated
Bronchodilator reversibility	dominated	dominated	dominated	dominated	dominated	dominated
FEV_1/FVC	dominated	dominated	dominated	dominated	dominated	dominated

Table 72: Deterministic sensitivity analysis – diagnosis (all patients) scenarios D7-D12

DSA scenario	D7	D8	D9	D10	D11	D12
	FP disutility doubled	FP disutility halved	FeNO marginal per-test cost doubled	FeNO marginal per-test cost halved	NObreath lifetime=3years	NObreath lifetime=5years
Airway hyperresponsiveness (MCT)	£1,094,325	£1,094,325	£1,084,543	£1,099,216	£1,091,217	£1,091,217
FeNO+bronchodilator reversibility (NoBreath)	-	-	-	-	dominated	dominated
FeNO+bronchodilator reversibility (NIOX VERO)	dominated	dominated	dominated	dominated	-	-
FeNO+bronchodilator reversibility (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NoBreath)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NoBreath)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
Sputum induction	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NoBreath)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
PEF	dominated	dominated	dominated	dominated	dominated	dominated
Bronchodilator reversibility	dominated	dominated	dominated	dominated	dominated	dominated
FEV_1/FVC	dominated	dominated	dominated	dominated	dominated	dominated

FP=false-positive

Table 73: Deterministic sensitivity analysis – diagnosis (all patients) scenarios D13-D18

DSA scenario	D13 NObreath lifetime=20years	D14 Visit costs doubled	D15 Visit costs halved	D16 FN exacerbation rate doubled	D17 FN exacerbation rate halved	D18 Asthma treatment costs doubled
Airway hyperresponsiveness (MCT)	£1,098,383	£2,196,057	£543,459	£1,090,925	£1,096,025	£1,100,100
FeNO+bronchodilator reversibility (NoBreath)	-	-	-	-	-	-
FeNO+bronchodilator reversibility (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+bronchodilator reversibility (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NoBreath)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NoBreath)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
Sputum induction	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NoBreath)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
PEF	dominated	dominated	dominated	dominated	dominated	dominated
Bronchodilator reversibility	dominated	dominated	dominated	dominated	dominated	dominated
FEV_1/FVC	dominated	dominated	dominated	dominated	dominated	dominated

FN=false-negative

Table 74: Deterministic sensitivity analysis – diagnosis (all patients) scenarios D19-D24

DSA scenario	D19	D20	D21	D22	D23	D24
	Asthma treatment costs halved	Misdiagnosis correction times x2	Misdiagnosis correction times x3	Misdiagnosis correction times x4	Misdiagnosis correction times x5	Misdiagnosis correction times x10
Airway hyperresponsiveness (MCT)	£1,091,438	£556,717	£377,547	£287,986	£234,270	£126,982
FeNO+bronchodilator reversibility (NoBreath)	-	-	-	£3,201	£5,111	£8,523
FeNO+bronchodilator reversibility (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+bronchodilator reversibility (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NoBreath)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NoBreath)	dominated	dominated	dominated	-	-	-
FeNO+FEV_1 (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
Sputum induction	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NoBreath)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
PEF	dominated	dominated	dominated	dominated	dominated	dominated
Bronchodilator reversibility	dominated	dominated	dominated	dominated	dominated	dominated
FEV_1/FVC	dominated	dominated	dominated	dominated	dominated	dominated

Table 75: Deterministic sensitivity analysis – diagnosis (all patients) scenarios D25-D30

DSA scenario	D25	D26	D27	D28	D29	D30
	Misdiagnosis correction times halved	FeNO operating characteristics from Schleich *et al*	FeNO operating characteristics from Pedrosa *et al*	Other test FeNO operating characteristics from Smith *et al*	Rule-out decision approach (base case)	Rule-out decision approach (best sensitivity for FeNO)
Airway hyperresponsiveness (MCT)	£2,169,614	£1,094,325	£1,094,325	£1,094,325	£1,119,170	dominated
FeNO+bronchodilator reversibility (NoBreath)	-	-	-	-	£6,965	£6,965 [rank 2]
FeNO+bronchodilator reversibility (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+bronchodilator reversibility (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NoBreath)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+sputum induction (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NoBreath)	dominated	dominated	dominated	dominated	-	- [rank 3]
FeNO+FEV_1 (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
Sputum induction	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NoBreath)	dominated	dominated	dominated	dominated	dominated	£244,914 [rank 1]
FeNO (NIOX VERO)	dominated	dominated	dominated	dominated	dominated	dominated
FeNO (NIOX MINO)	dominated	dominated	dominated	dominated	dominated	dominated
PEF	dominated	dominated	dominated	dominated	dominated	dominated
Bronchodilator reversibility	dominated	dominated	dominated	dominated	dominated	dominated
FEV_1/FVC	dominated	dominated	dominated	dominated	dominated	dominated

Table 76: Deterministic sensitivity analysis – diagnosis (all patients) scenarios D31-D34

DSA scenario	D31	D32	D33	D34
	Rule-out decision approach (best specificity for FeNO)	Rule-in decision approach all options (base case)	Rule-out decision approach all options (best sensitivity for FeNO)	Rule-in decision approach all options (best specificity for FeNO)
Airway hyperresponsiveness (MCT)	£1,119,170	dominated	dominated	dominated
FeNO+bronchodilator reversibility (NoBreath)	£6,965	- [rank 2]	- [rank 2]	- [rank 2]
FeNO+bronchodilator reversibility (NIOX VERO)	dominated	dominated	dominated	dominated
FeNO+bronchodilator reversibility (NIOX MINO)	dominated	dominated	dominated	dominated
FeNO+sputum induction (NoBreath)	dominated	dominated	dominated	dominated
FeNO+sputum induction (NIOX VERO)	dominated	dominated	dominated	dominated
FeNO+sputum induction (NIOX MINO)	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NoBreath)	-	£63,533 [rank 1]	£63,533 [rank 1]	£63,533 [rank 1]
FeNO+FEV_1 (NIOX VERO)	dominated	dominated	dominated	dominated
FeNO+FEV_1 (NIOX MINO)	dominated	dominated	dominated	dominated
Sputum induction	dominated	dominated	dominated	dominated
FeNO (NoBreath)	dominated	dominated	dominated	dominated
FeNO (NIOX VERO)	dominated	dominated	dominated	dominated
FeNO (NIOX MINO)	dominated	dominated	dominated	dominated
PEF	dominated	dominated	dominated	dominated
Bronchodilator reversibility	dominated	dominated	dominated	dominated
FEV_1/FVC	dominated	dominated	dominated	dominated

The deterministic sensitivity analyses indicate the following:

- Across the majority of scenarios, the cost-effectiveness frontier presented in the base case analysis (which includes only airway hyperresponsiveness and FeNO plus bronchodilator reversibility) is maintained. In most scenarios the majority of options are expected to be ruled out due to simple dominance.
- The results based on the point estimates of parameters are similar to the results of the probabilistic analysis.
- Discounting does not have a substantial effect on the cost-effectiveness of the non-dominated diagnostic options.
- The disutility associated with loss of control in false-negatives has a substantial impact upon the incremental cost-effectiveness of airway hyperresponsiveness versus FeNO plus bronchodilator reversibility.
- The false-positive exacerbation rate has no impact on the results as both non-dominated options have the same specificity.
- The cost of the various FeNO devices influences which options are dominated but has only a negligible impact upon the cost-effectiveness results for non-dominated options.
- Longer misdiagnosis correction times substantially improve the cost-effectiveness of airway hyperresponsiveness (MCT) versus FeNO plus bronchodilator reversibility.
- The use of other sources for the operating characteristics of FeNO and standard tests does not impact upon the cost-effectiveness of non-dominated options.
- The use of a "rule-out" decision approach may improve the comparative effectiveness and cost-effectiveness of FeNO alone.
- The use of a "rule-in" decision approach may improve the effectiveness of FeNO plus FEV_1, however the ICER for this option (versus FeNO plus bronchodilator reversibility) is in excess of £63,000 per QALY gained.

6.5.2 Management model results (children)

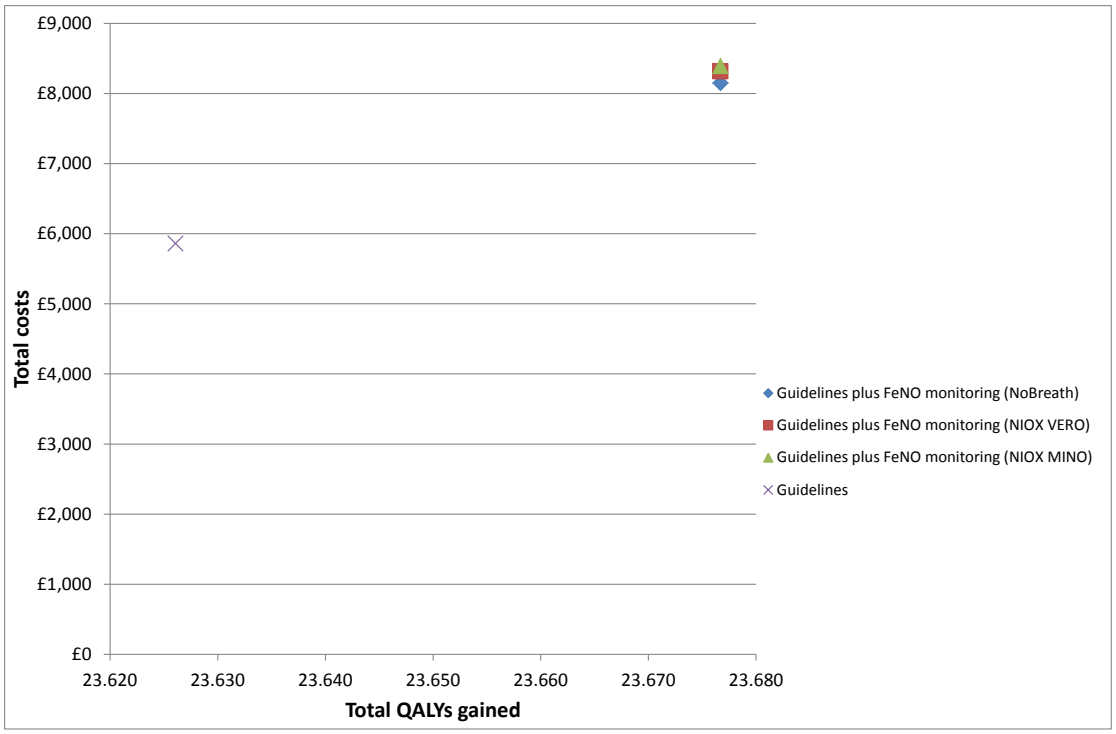
6.5.2.1 Central estimates of cost-effectiveness – management (children)

Table 77 presents the central estimates of cost-effectiveness based on the probabilistic version of the children management model. The results suggest that FeNO testing is expected to produce a small health benefit compared to guidelines (0.05 QALYs). FeNO testing is also expected to be more expensive than guidelines alone; this is due to projected ICS use for the FeNO groups. The results also indicate, as expected, that NIOX MINO and NIOX VERO are expected to be dominated by NObreath due to their slightly higher marginal per-test cost. The incremental cost-effectiveness of NObreath versus guidelines is expected to be approximately £45,213 per QALY gained. This information is presented on the absolute cost-effectiveness plane presented in Figure 29.

Table 77: Central estimates of cost-effectiveness – management (children)

Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
Guidelines plus FeNO monitoring (NoBreath)	23.6767	£8,148.59	0.0506	£2,288.53	**£45,213**
Guidelines plus FeNO monitoring (NIOX VERO)	23.6767	£8,314.30	-	-	**dominated**
Guidelines plus FeNO monitoring (NIOX MINO)	23.6767	£8,391.53	-	-	**dominated**
Guidelines	23.6261	£5,860.06	-	-	-

Figure 29: Cost-effectiveness plane – management (children)



6.5.2.2 Uncertainty analysis – management (children)

Figure 30 presents CEACs for the management options in the children subgroup. These data are also summarised in Table 78. Assuming a willingness to pay threshold of £20,000 per QALY gained, guidelines alone has the highest probability of producing the greatest amount of net benefit (probability = 0.99). Assuming a willingness to pay threshold of £30,000 per QALY gained, guidelines has the highest probability of producing the greatest amount of net benefit (probability = 0.91).

Figure 30: Cost-effectiveness acceptability curves – management (children)

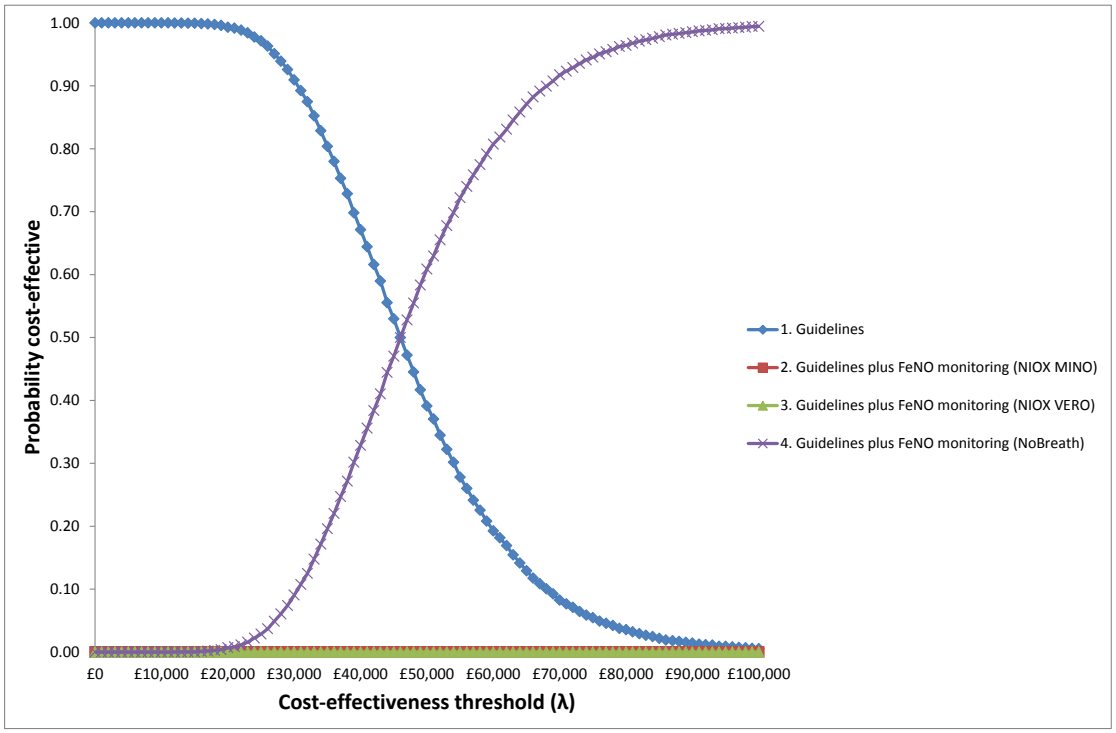


Table 78: Probability of optimality – management (children)

Option	Probability optimal λ=£20,000/QALY gained	Probability optimal λ=£30,000/QALY gained
1. Standard guidelines	**0.99**	**0.91**
2. Standard guidelines plus FeNO monitoring (NIOX MINO)	0.00	0.00
3. Standard guidelines plus FeNO monitoring (NIOX VERO)	0.00	0.00
4. Standard guidelines plus FeNO monitoring (NObreath)	0.01	0.09

6.5.2.3 Deterministic sensitivity analysis results

Table 79 presents the results of the deterministic sensitivity analysis.

Table 79: Deterministic sensitivity analysis – management (children)

Scenario		Guidelines plus FeNO (NObreath)	Guidelines plus FeNO (NIOX VERO)	Guidelines plus FeNO (NIOX MINO)	Guidelines
M1	Point estimates of parameters	£45,138	dominated	dominated	-
M2	Undiscounted costs and outcomes	£46,894	dominated	dominated	-
M3	Discount rate=6%	£44,555	dominated	dominated	-
M5	Analysis based on Pijnenburg *et al*	£18,963	dominated	dominated	-
M8	FeNO impact=1-year	dominating	dominated	dominated	dominated
M9	FeNO impact=2-years	dominating	dominated	dominated	dominated
M10	FeNO impact=3-years	dominating	dominated	dominated	dominated
M11	FeNO impact=4-years	dominating	dominated	dominated	dominated
M12	FeNO impact=5-years	£7,598	dominated	dominated	-
M13	FeNO impact=10-years	£27,660	dominated	dominated	-
M14	FeNO impact=15-years	£34,337	dominated	dominated	-
M15	FeNO impact=20-years	£37,674	dominated	dominated	-
M16	FeNO impact=30-years	£41,025	dominated	dominated	-
M17	FeNO impact=40-years	£42,721	dominated	dominated	-
M18	Marginal per-test FeNO cost doubled	£55,409	dominated	dominated	-
M19	Marginal per-test FeNO cost halved	£40,003	dominated	dominated	-
M20	NObreath lifetime=3years	£45,138	dominated	dominated	-
M21	NObreath lifetime=5years	£45,138	dominated	dominated	-
M22	NObreath lifetime=20 years	£40,878	dominated	dominated	-
M23	FeNO nurse visits doubled	£84,564	dominated	dominated	-
M24	FeNO nurse visits halved	£25,425	dominated	dominated	-
M25	Exacerbation rates doubled	£19,891	dominated	dominated	-
M26	Exacerbation rates halved	£95,632	dominated	dominated	-
M27	Exacerbation disutility doubled	£31,479	dominated	dominated	-
M28	Exacerbation disutility halved	£52,844	dominated	dominated	-
M29	ICS observed mean carried forward	£37,452	dominated	dominated	-
M30	ICS change doubled	£56,206	dominated	dominated	-
M31	ICS change halved	£39,604	dominated	dominated	-

The deterministic sensitivity analyses indicate the following:

- The results of the analysis using point estimates of parameters is similar to those produced using the probabilistic model

- NIOX MINO and NIOX VERO are expected to be consistently dominated by NObreath due to their higher marginal per-test cost.
- Whilst the marginal per-test cost influences which device would be preferred, it does not have a substantial impact on the overall cost-effectiveness of FeNO versus guidelines.
- Discounting has little impact upon the cost-effectiveness of FeNO monitoring.
- The duration over which FeNO monitoring is assumed to impact upon exacerbations and ICS use is a key parameter within the children subgroup. Shorter durations of impact improve the cost-effectiveness of FeNO monitoring.
- The analysis based on Pijnenburg *et al* suggests a considerably more favourable ICER for FeNO versus guidelines in children. This may be explained by the fact that the Szefler *et al* study was undertaken in uncontrolled patients and the study protocol did not allow therapy to be stepped down on the basis of low FeNO alone. This may in part explain why ICS use was higher for FeNO than guidelines alone.
- The model is sensitive to the rate of exacerbations (and associated health loss) and assumptions regarding the number of monitoring visits in which FeNO is used.

6.5.3 Management model results (adults)

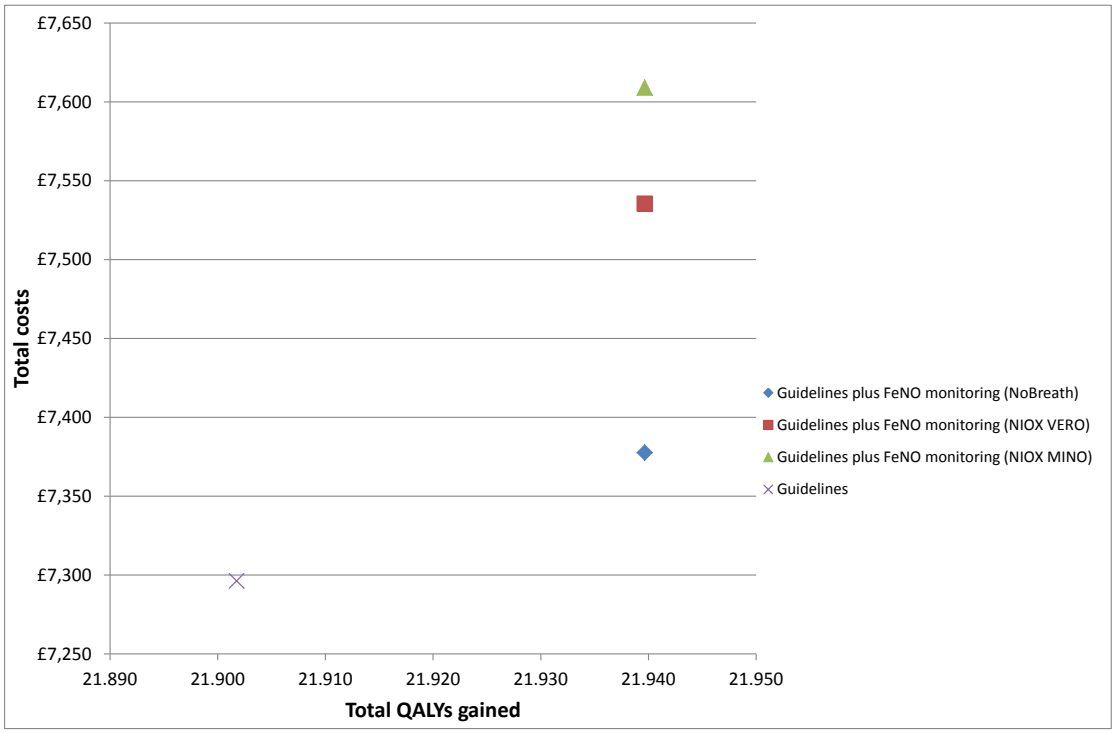
6.5.3.1 Central estimates of cost-effectiveness – management (adults)

Table 80 presents the central estimates of cost-effectiveness based on the probabilistic version of the adult management model. FeNO testing is expected to produce a small incremental health gain compared to standard guidelines (0.04 QALYs). The results also suggest that NIOX MINO and NIOX VERO are expected to be dominated by NObreath (again, this is due to the slightly lower marginal per-test cost for this device). In this population subgroup, the NObreath plus guidelines versus guidelines alone is expected to cost approximately £2,146 per QALY gained. This information is presented on the absolute cost-effectiveness plane presented in Figure 31.

Table 80: Central estimates of cost-effectiveness – management (adults)

Option	QALYs	Costs	Incremental QALYs	Incremental costs	Incremental cost per QALY gained
Guidelines plus FeNO monitoring (NoBreath)	21.9397	£7,377.61	0.0379	£81.31	**£2,146**
Guidelines plus FeNO monitoring (NIOX VERO)	21.9397	£7,535.43	-	-	**dominated**
Guidelines plus FeNO monitoring (NIOX MINO)	21.9397	£7,608.99	-	-	**dominated**
Guidelines	21.9018	£7,296.30	-	-	**-**

Figure 31: Cost-effectiveness plane – management (adults)



6.5.3.2 Uncertainty analysis – management (adults)

Figure 32 presents CEACs for the management options in the adult subgroup. Assuming a willingness to pay threshold of £20,000 per QALY gained, FeNO monitoring using NObreath plus guidelines has the highest probability of producing the greatest amount of net benefit (probability = 0.82). Assuming a willingness to pay threshold of £30,000 per QALY gained, FeNO monitoring using NObreath plus guidelines has the highest probability of producing the greatest amount of net benefit (probability = 0.87). These results are summarised in Table 81.

Figure 32: Cost-effectiveness acceptability curves – management (adults)

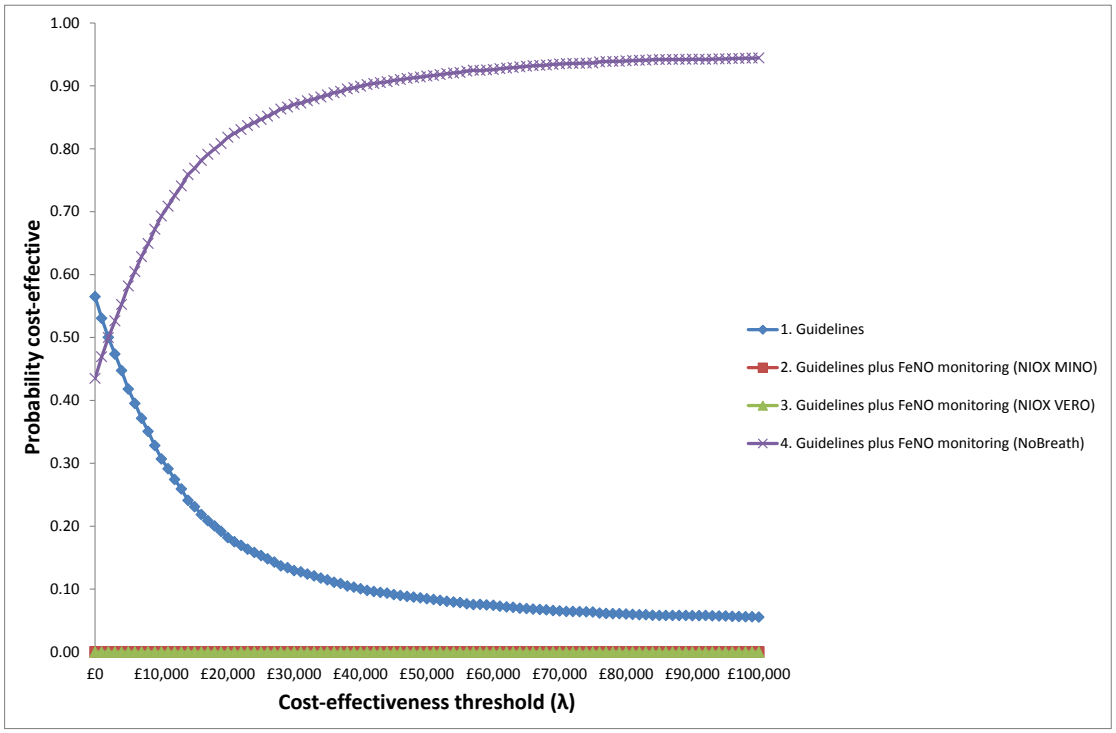


Table 81: Probability of optimality – management (adults)

Option	Probability optimal λ=£20,000/QALY gained	Probability optimal λ=£30,000/QALY gained
1. Guidelines	0.18	0.13
2. Guidelines plus FeNO monitoring (NIOX MINO)	0.00	0.00
3. Guidelines plus FeNO monitoring (NIOX VERO)	0.00	0.00
4. Guidelines plus FeNO monitoring (NObreath)	**0.82**	**0.87**

6.5.3.3 Deterministic sensitivity analyses

Table 82 presents the results of the deterministic sensitivity analysis.

Table 82: Deterministic sensitivity analysis – management (adults)

Scenario		*Guidelines plus FeNO (NObreath)*	*Guidelines plus FeNO (NIOX VERO)*	*Guidelines plus FeNO (NIOX MINIO)*	*Guidelines*
M1	Point estimates of parameters	£2,248	dominated	dominated	-
M2	Undiscounted costs and outcomes	£740	dominated	dominated	-
M3	Discount rate=6%	£3,534	dominated	dominated	-
M4	Pregnant women subgroup	dominating	dominated	dominated	-
M6	Analysis based on Smith *et al*	£184,095	dominated	dominated	-
M7	Analysis based on Syk *et al*	dominated	dominated	dominated	dominating
M8	FeNO impact=1-year	£885,451	dominated	dominated	-
M9	FeNO impact=2-years	£434,284	dominated	dominated	-
M10	FeNO impact=3-years	£283,954	dominated	dominated	-
M11	FeNO impact=4-years	£208,833	dominated	dominated	-
M12	FeNO impact=5-years	£163,795	dominated	dominated	-
M13	FeNO impact=10-years	£73,975	dominated	dominated	-
M14	FeNO impact=15-years	£44,320	dominated	dominated	-
M15	FeNO impact=20-years	£29,707	dominated	dominated	-
M16	FeNO impact=30-years	£15,531	dominated	dominated	-
M17	FeNO impact=40-years	£8,898	dominated	dominated	-
M18	Marginal per-test FeNO cost doubled	£15,273	dominated	dominated	-
M19	Marginal per-test FeNO cost halved	dominating	dominated	dominated	-
M20	NObreath lifetime=3years	£2,248	dominated	dominated	-
M21	NObreath lifetime=5years	£2,248	dominated	dominated	-
M22	NObreath lifetime=20 years	dominating	dominated	dominated	-
M23	FeNO nurse visits doubled	£52,246	dominated	dominated	-
M24	FeNO nurse visits halved	dominating	dominated	dominated	-
M25	Exacerbation rates doubled	dominating	dominated	dominated	-
M26	Exacerbation rates halved	£9,958	dominated	dominated	-
M27	Exacerbation disutility doubled	£1,563	dominated	dominated	-
M28	Exacerbation disutility halved	£2,634	dominated	dominated	-
M29	ICS observed mean carried forward	£66,453	dominated	dominated	-
M30	ICS change doubled	dominating	dominated	dominated	-
M31	ICS change halved	£23,392	dominated	dominated	-

The deterministic sensitivity analyses indicate the following:

- The results of the analysis using point estimates of parameters are very similar to those produced using the probabilistic version of the model.
- NIOX MINO and NIOX VERO are expected to be consistently dominated by NObreath due to their higher marginal per-test cost.
- FeNO monitoring using NObreath is expected to dominate standard guidelines in the subgroup of women who are pregnant.
- Discounting has little impact upon the cost-effectiveness of FeNO monitoring.

- Whilst the marginal per-test cost influences which device would be preferred, it does not have a substantial impact on the overall cost-effectiveness of FeNO versus guidelines.
- The use of exacerbation rates from Syk *et al* and Smith *et al* have a substantial negative impact upon the cost-effectiveness of FeNO monitoring (ICER ranges from £184,000 per QALY gained to dominated).
- The duration over which FeNO monitoring is assumed to impact upon exacerbations and ICS is a key driver of cost-effectiveness. It is noteworthy that in the adult subgroup, cost-effectiveness improves over longer time horizons – the opposite is true within the children subgroup whereby cost-effectiveness worsens over longer time horizons. This is driven entirely by the observed differences in relative ICS use for FeNO and guidelines at the last observed timepoint in the trials.
- The cost-effectiveness of FeNO monitoring is markedly less favourable when projected ICS use is modelled according to the mean ICS use observed in the trial reported by Shaw *et al*.[108]

6.6 Discussion

6.6.1 Summary of cost-effectiveness evidence

There is very limited available evidence concerining the cost-effectiveness of FeNO testing for the diagnosis and/or management of asthma. The systematic review presented in this chapter identified one published UK model of FeNO testing in the diagnostic setting and one published model of FeNO testing in the management setting.[134] These models were published within the same paper. Aerocrine submitted a model of FeNO testing for diagnosis and a model of FeNO testing for management; these models were similar to, but not the same as, the published Price *et al* models.

The Price *et al* diagnostic model indicates that NIOX MINO is likely to be cost saving in comparison to other tests routinely used in the diagnosis of asthma. The model analysis presented by Price *et al*[134] also suggests that NIOX MINO is expected to be more expensive than standard diagnostic tests when used in conjunction with other tests. The EAG critique of this model highlighted a number of problems including the use of a blended comparison, the questionable selection of evidence used to inform the model's parameters and the absence of any quantified health consequences associated with diagnostic test outcomes. The Aerocrine diagnostic model is similar in structure to the published version but does not use a blended comparison approach and includes some updated parameter values. However, the Aerocrine model also fails to reflect the health consequences associated with correct or incorrect diagnostic outcomes. Owing to their limited scope, these diagnostic models do not provide any information regarding the economic trade-off between potential additional health gains resulting from the more accurate diagnosis of asthma and the health loss associated with displacing existing services.

The Price *et al*[134] management model compares FeNO monitoring using NIOX MINO versus guidelines. This model was evaluated within a cost-utility framework and indicates that NIOX MINO produces more health gain at a lower cost than guidelines; in other words, NIOX MINO dominates management using guidelines alone. Aerocrine submitted a similar management model which included some different data and assumptions but ultimately produced the same conclusion as the published analysis reported by Price *et al*.[134] The EAG critique of these management models highlighted a number of problems including the use of a short time horizon, the selective use of efficacy evidence, assumptions regarding equivalence between sputum count monitoring and FeNO, and invalid assumptions regarding the health losses associated with exacerbations. No economic evidence was submitted by the manufacturers for either NIOX VERO or NObreath. The EAG takes the view that neither the published Price *et al* models nor the submitted Aerocrine models represent a suitable basis for informing decision-making about the use of FeNO testing for the diagnosis or management of asthma.

In light of the problems with the available evidence, the EAG developed two *de novo* models:

(i) a model to assess the cost-effectiveness of FeNO testing using NIOX MINO, NIOX VERO and NObreath in addition to, or in place of existing tests, versus other diagnostic options commonly used in the diagnosis of asthma, and;

(ii) a model to assess the cost-effectiveness of NIOX MINO, NIOX VERO and NObreath plus guidelines versus guidelines alone for the management of asthma.

The EAG diagnostic model suggests that across the diagnostic options included in the economic analysis, the expected difference in QALY gains is likely to be very small. Airway hyperresponsiveness (MCT) is expected to produce the greatest QALY gain; this is because this option has the highest sensitivity of all tests included in the economic analysis. All options which incude NIOX MINO or NIOX VERO are expected to be dominated as their marginal per-test cost is higher than that for NObreath (assuming a device lifetime of 10-years). In the base case analysis, all options except airway hyperresponsiveness (MCT) and FeNO plus bronchodilator reversibility testing are expected to be ruled out by simple dominance. The incremental cost-effectiveness of airway hyperresponsivenss (MCT) versus FeNO (NObreath) plus bronchodilator reversibility is expected to be £1.125million per QALY gained. The results of the analysis are particularly sensitive to assumptions about the duration of time required to resolve misdiagnoses, assumptions about health losses incurred by patients who are false-negative, the costs of asthma management and the use of "rule-in" and "rule-out" diagnostic decision rules.

The EAG management model was evaluated across two subgroups – (i) children and (ii) adults. Within both the children and adult subgroup base case analyses, FeNO testing is expected to produce

a small incremental QALY gain compared to guidelines alone. In both subgroups, NIOX MINO and NIOX VERO are expected to be dominated as their marginal per-test cost is higher than that for NObreath. Within the children subgroup, the incremental cost-effectiveness of guidelines plus FeNO monitoring using NObreath versus guidelines alone is expected to be approximately £45,200 per QALY gained. This ICER is influenced considerably by the assumed change in ICS use which is applied over a lifetime horizon. Within the adult subgroup, FeNO monitoring using NObreath versus guidelines alone is expected to cost approximately £2,100 per QALY gained. A similarly favourable result was produced within a further analysis based on a subgroup of women who are pregnant.[111] Importantly, these positive results are not held when alternative trials are used to inform the analysis.[109,125] The results in the children and adult subgroups are particularly sensitive to assumptions regarding changes in ICS use over time, the number of nurse visits for FeNO monitoring and the duration over which FeNO monitoring is assumed to impact upon exacerbations and ICS use.

6.6.2 Limitations of the EAG model

Whilst the EAG models presented here do resolve many of the problems identified within the Price/Aerocrine models, the results drawn from these models remain subject to considerable uncertainty. These are briefly discussed below.

6.6.2.1 Limitations of the diagnostic model

The following represent the key limitations and uncertainties within the EAG diagnostic model:

- *Use of naïve indirect comparisons.* Limitations in the diagnostic evidence base means that the model must use naïve indirect comparisons across studies which assess the diagnostic accuracy of different tests. The review presented in Chapter 5 highlighted considerable heterogeneity between these studies. As such, the results of the economic analysis of FeNO in the diagnostic setting should be interpreted with caution.
- *Non-systematic approach to including non-FeNO comparators.* We did not undertake a systematic review of evidence concerning the diagnostic accuracy of existing tests used in the diagnosis of asthma but instead relied on studies picked up by our systematic review of FeNO studies.[90,151] Whilst a formal review of other tests (excluding FeNO) would be valuable, this was beyond the scope of the assessment and the time and resource available to the EAG precluded this work. It is likely that other studies exist and it is possible that these could be considered more relevant than the studies used in the EAG model.
- *Use of a "blunt" model structure.* We adopted a similar model structure to Price *et al*[134] which assesses options at a particular point in the diagnostic pathway rather than attempting to simulate the entire sequence of tests used throughout the pathway. This model development decision was taken due to limitations in the available evidence.

- *Uncertainty surrounding health losses associated with misdiagnosis.* We crudely elicited estimates of the duration required to resolve a false-negative/false-positive diagnosis. Only one of our experts was able to tentatively quantify the likely values of these parameters. These estimates are very uncertain. There is also uncertainty surrounding the magnitude of the HRQoL loss as well as the duration over which this loss is incurred. It is possible that health losses associated with false-positive diagnoses in patients with more serious underlying pathology are underestimated. It is not clear how this uncertainty could be resolved empirically.

6.6.2.2 Limitations of the management model

The following represent the key limitations and uncertainties within the EAG management model:

- *Use of effectiveness evidence.* The model uses individual studies within the children and adult subgroups. These studies were deemed by the EAG to most closely reflect asthma management in England and Wales. However, the Szefler *et al* study[113] was undertaken in the US and does not fully match BTS/SIGN guidelines on dose titration. Only the use of guidelines in the comparator group within the Shaw study can be considered as "standard" within the UK.
- *Uncertainty surrounding the duration over which FeNO impacts on dose titration.* In line with the NICE Reference Case, the EAG model reflects a lifetime horizon. There is however considerable uncertainty with respect to the duration over which FeNO monitoring would result in different exacerbation rates and ICS use compared to guidelines alone. Within the base case analysis we assumed that this impact would be sustained indefinitely. The sensitivity analysis shows that this parameter is a key driver of cost-effectiveness.

Both the EAG diagnostic model and the EAG management model assume that all FeNO devices have the same diagnostic properties in terms of absolute FeNO measurements and how this translates into sensitivity and specificity. This necessary assumption may not hold in reality.

6.6.3 Areas for further research

Further research would be valuable to reduce some of the uncertainties detailed above. In particular, comparative studies which include FeNO alongside the range of existing standard tests with a common population and common reference standard of long term follow up of at least a year would be useful in assessing the comparative accuracies of these alternative diagnostic strategies. In addition, longer-term studies of FeNO monitoring, in combination with standard UK management guidelines, would be beneficial to better understand the long-term impacts on asthma medication use and exacerbations.

7. ASSESSMENT OF FACTORS RELEVANT TO THE NHS AND OTHER PARTIES

Beyond its likely clinical and cost-effectiveness, a number of other factors relating to the implementation of FeNO testing in the NHS require consideration.

7.1 Training and education

The introduction of FeNO testing for the diagnosis and/or management of asthma has implications for training and education, in terms of teaching NHS staff how to instruct patients how to use the device correctly in order to minimise test failure rates. Repeatability and accuracy of the devices are not dependent on patient performance as the devices will not produce a measurement if flow rate and length of exhalation limits are not met. The precise training and education requirements associated with introducing FeNO testing are dependent on whether it is routinely recommended and if so, the setting that such recommendations relate to. Training may be required for primary care nurses and GPs or for secondary care staff, or for both. It should be noted that these additional costs are not reflected in the marginal per-test costs used within the economic analysis presented in Chapter 6.

7.2 Purchasing of equipment and consumables

The diffusion of FeNO testing into routine NHS practice would involve the purchasing of additional equipment either for GP surgeries or hospitals. Equipment costs include the devices, replacement parts (NObreath only) and other consumables (test kit mouthpieces). The NIOX MINO and NIOX VERO devices both have a finite lifetime and would need to be replaced at a maximum of 3-years and 5-years respectively. The NObreath device does not have a finite lifetime but does require replacement sensor cells every two years. Each FeNO device requires the purchase of test kit mouthpieces; the volume purchased and the number of tests undertaken will influence the overall marginal per-test cost of each device for GP surgeries and trusts. The cost of maintenance of the NObreath device is expected to be free of charge to the NHS. Aerocrine did not mention the cost of maintenance in their cost estimates.

7.3 Replacement of the NIOX MINO device with the newer NIOX VERO device

It is anticipated that the NIOX MINO device will soon be replaced with the newer NIOX VERO device. Both FeNO devices will be available for some time, but in the long-term, NIOX MINO will eventually become redundant. It is likely that the NIOX VERO device will be less expensive (per test) than NIOX MINO hence the justification for purchasing the NIOX MINO device is unclear.

7.4 Impact on demand for current standard diagnostic tests

The introduction of FeNO testing in a diagnostic setting will likely have impacts on the demand for other existing standard tests currently used in the diagnosis of asthma. This change in the level of demand for existing standard tests will be dependent how FeNO is incorporated into the existing asthma diagnosis pathway.

7.5 FeNO testing in children

The diagnostic and clinical evidence considered in this assessment is restricted to patients aged 5-years and over. The potential diagnostic/management benefit of FeNO use in younger children is unknown.

7.6 FeNO testing in older adults

FeNO does not seem to be a useful test in the diagnosis or management of older adults with asthma. In this population, other current standard tests and management approaches may be more applicable.

7.7 Patients with respiratory tract infections

Most studies included in this assessment (see Chapter 5) purposefully excluded patients with recent respiratory tract infections. The diagnostic utility of FeNO testing in these patients is unclear. It may be more appropriate either to use standard diagnostic tests in these patients or to allow a period of recovery before using FeNO.

8. DISCUSSION

8.1 Statement of principal findings

8.1.1 Equivalence of devices

Whilst there was often good correlation between FeNO measurement devices, equivalence of readings could not necessarily be assumed in all situations. Whilst many studies concluded that the comparability of measurements between devices was within clinically acceptable limits, others went on to produce correction equations to correct for systematic bias in measurements. There was also no common justified definition of clinically acceptable differences, and 95% limits of agreement were sometimes very wide (around 20pbb). There seemed to be a generally consistent observation of poorer equivalence between FeNO devices at higher FeNO levels. The direction of disagreement varied between studies and comparator devices.

However, as there is mostly a high degree of correlation between measurements across all devices, estimates of sensitivity and specificity are likely to be a reasonable indication of potential diagnostic accuracy of using FeNO to guide diagnosis and management, but the derived cut-off points are not likely to be interchangeable between devices. As such, for the purpose of this assessment, sensitivities and specificities will be assumed to be interchangeable, but it cannot be assumed that the cut-off points that should be used to achieve them will be the same in each device, and there is still some doubt as to whether the same diagnostic accuracy would be achievable in all devices. This is an important issue that should be considered in the interpretation of the diagnostic accuracy review and the findings of the health economic analysis assessment presented within this report.

Test failure rates were generally low in all devices in adults, the highest reported rate being 3.3%. In children, there may be some problems using the NIOX MINO device in younger children, with failure rates ranging from 5.5% to 27%. One study used NObreath with children and reported no test failures.

8.1.2 Diagnostic accuracy review

This review identified several groups of studies that were similar to one another in terms of the position of the patients in the UK pathway and the reference standards used. These groups were: adults presenting with symptoms of asthma versus most of or the entire UK pathway; a subset of adults presenting with symptoms of asthma versus airway hyper-responsiveness; difficult to diagnose patients versus airway hyper-responsiveness; patients with chronic cough who were difficult to diagnose, versus ICS responsiveness; children with symptoms of asthma versus various reference standards.

No meta-analysis was conducted in any group as clinical heterogeneity between studies was generally extremely high. Estimates of sensitivity and specificity were not consistent within groups and ranged widely when used as a rule-in test, rule-out test and as a diagnostic test. Table 83 summarises the results across studies and groups of studies. Given the wide ranging estimates of sensitivity and specificity, together with heterogeneous cut-off points, it is difficult to draw any firm conclusions as to the diagnostic accuracy of FeNO in any situation and at any given cut-off point. Interestingly, there did not appear to be an obvious difference in the diagnostic accuracy of FeNO versus the whole or parts of the UK pathway in patients who present with symptoms of asthma compared to the diagnostic accuracy of FeNO versus airway hyper-responsiveness in patients who are difficult to diagnose. The large variation in estimates within groups may obscure any true underlying differences in the accuracy of FeNO between groups and versus different reference standards.

However, some limited observations can be made. It would appear that FeNO was more often able to reach 100% specificity than 100% sensitivity, and that ranges of specificity were generally tighter. This may indicate it has best potential for use as a rule-in test. It would also appear that FeNO cut-off points should probably be lower in children than in adults.

In addition to the above, two studies were found that reported results for FeNO in conjunction with another test in adults, one in those difficult to diagnose,[83] and one in patients of all ages with symptoms of asthma.[91] In both cases, the addition of another test to the diagnostic protocol resulted in a change in diagnostic accuracy, but as this involved the usual trade off between sensitivity and specificity it is difficult to tell if this represents an increase or decrease in clinical and cost-effectiveness.

Evidence was limited in the subgroups defined *a priori*, namely pregnancy, the elderly and smokers/environmental tobacco exposure.

- Smokers: FeNO appeared to be able to distinguish between asthmatics and non-asthmatics in adult smokers with similar accuracy as in non-smokers and ex-smokers. It would seem likely that FeNO is generally lower in smokers, and it may be useful to consider a patient's smoking status when interpreting results, or to select lower cut-off points for smokers. Limited data in children support the same conclusion as for adults.
- The elderly: A case control study indicated that FeNO is unlikely to be a useful test in the diagnosis of asthma in the elderly.
- Pregnant women: A cross-sectional study indicated that pregnancy does not alter FeNO levels in asthmatics or non-asthmatics, and that FeNO can distinguish between asthmatic and non-asthmatic pregnant and healthy women.

8.1.3 FeNO-guided management in asthma

Four studies in adults were identified. There were high levels of heterogeneity in multiple study characteristics and outcome definitions, and as such it was not possible to draw any firm conclusions as to which step-up/step-down protocol or cut-off points offer the best efficacy. All studies reported a fall in exacerbation rates per person year, though it appeared that this was mostly driven by mild and moderate exacerbations and this was only statistically significant in one study.[109] Pooled analysis showed a non-significant trend in favour of the intervention group for severe exacerbations and a statistically significant decrease in exacerbations in the intervention groups when considering the composite outcomes of any severity of exacerbation. The effects on ICS use were heterogeneous, with two studies showing statistically significant decreases in ICS use in the FeNO-guided management groups, one study showing a minor increase, and another showing very similar levels of use in each arm. This may indicate that some step-up step-down protocols were better at decreasing ICS use than others, or may be due to the characteristics of the study populations. Pooled analysis showed a statistically non-significant trend towards decreased ICS use. HRQoL was infrequently reported; the two studies used versions of the AQLQ to measure quality of life and both showed no effect in the global score, but one investigated domains and found a statistically significant difference in the symptoms score.

Despite the heterogeneity in results, and the lack of statistically significant findings, it would seem possible to conclude that, in adults, FeNO-guided management of some or most designs is likely, during the first year of management, to result in a non-significant trend towards better management overall with either a small or zero reduction in ICS use. There was no evidence relating to whether these effects would be maintained over a longer time period.

Five studies in children were identified. One study appeared to recruit a group of patients who were well controlled[115] whilst two others recruited patients who appeared to be poorly controlled.[113,116] Both reported fewer severe exacerbations in the intervention arm, but not statistically significantly so. All studies reported a decrease in exacerbations (however defined) in the intervention arm, but only one reported a statistically significant reduction.[116] The effects on ICS use were heterogeneous, with two studies showing a statistically significant increase in ICS use,[112,113] one showing no difference,[115] one being difficult to interpret[114] and one further study not reporting this outcome.[116] HRQoL was only reported within one study,[116] although insufficient details were reported.

Due to the high levels of heterogeneity in multiple study characteristics and outcome definitions, it was not possible to draw any firm conclusion as to which step-up/step-down protocol or cut-off points offer the best efficacy for management. However, it would seem possible to conclude that FeNO-

guided management of most descriptions is likely, during the first year of management, to result in non-statistically significant trends towards better management overall. It is unclear whether ICS use is likely to increase or decrease, and this may depend on the details of the step-up step-down protocols or the characteristics of the patients recruited to the trials in terms of control and severity. There was no evidence relating to whether these effects would be maintained over a longer time period.

Table 83: Summary of diagnostic accuracy

			Highest sum of sensitivity and specificity			Rule-out scenario			Rule-in scenario		
Patients	Reference standard	Number of studies	Range of cut-offs	Range of sensitivity values	Range of specificity values	Range of cut-offs	Range of sensitivity values	Range of specificity values	Range of cut-offs	Range of sensitivity values	Range of specificity values
Adults with symptoms of asthma	Part or whole of UK pathway	4	20 ppb to 47 ppb	32% to 88%	75% to 93%	9 ppb to 15 ppb	85% to 96%	13% to 48%.	47 ppb to 76 ppb	55.6% to 13%	88.2% to 100%
Subset of patients at Position A	Airway reversibility OR airway hyper-responsiveness	2	27 ppb to 36 ppb	77.8% to 87%	60% to 92%	25 ppb	100%	46.7%	100 ppb	27.8%	100%
Difficult to diagnose patients	Versus airway hyper-responsiveness	3	34 ppb to 40 ppb	24.4% to 74.3%	72.5% to 98.9%	NR	NR	NR	NR	NR	NR
Patients with chronic cough, difficult to diagnose	ICS responsiveness	3	20 ppb to 38 ppb	53% to 94.7%	63% to 85%	NR	NR	NR	NR	NR	NR
Children with symptoms of asthma	Various	4	19 to 21 ppb	49% to 86%	76% to 89%,	5 to 20ppb	89% to 94%	14.1% to 70%	30 to 50ppb	20% to 50%	92% to 100%

8.1.4 Independent assessment of cost-effectiveness

The EAG developed two *de novo* models. The first model assesses the cost-effectiveness of FeNO testing using NIOX MINO, NIOX VERO and NObreath in addition to, or in place of existing tests, as compared against other diagnostic options commonly used in the diagnosis of asthma. The second model assesses the cost-effectiveness of NIOX MINO, NIOX VERO and NObreath plus guidelines versus guidelines alone for the management of asthma.

The EAG diagnostic model suggests that across the 17 options included in the analysis, airway hyperresponsiveness (MCT) is expected to produce the greatest QALY gain. All options which include NIOX MINO or NIOX VERO are expected to be dominated as their marginal per-test cost is higher than that for NObreath. The incremental cost-effectiveness of airway hyperresponsivenss (MCT) versus FeNO (NObreath) plus bronchodilator reversibility is expected to be £1.125million per QALY gained. All other options are ruled out of the analysis due to simple dominance. The results of the analysis are particularly sensitive to assumptions about the duration of time required to resolve misdiagnoses, assumptions about health losses incurred by patients who are false-negative, the costs of asthma management and the use of "rule-in" and "rule-out" diagnostic decision rules.

The EAG management model was evaluated separately for children and adult subgroups. Within both the children and adult subgroup analyses, FeNO monitoring plus guidelines is expected to produce a small incremental QALY gain compared to guidelines alone. NIOX MINO and NIOX VERO are expected to be dominated as their marginal per-test cost is higher than that for NObreath. Within the adult subgroup, FeNO monitoring using NObreath versus guidelines alone is expected to cost approximately £2,100 per QALY gained. A similarly favourable result was produced within a further analysis based on a subgroup of women who are pregnant.[111] Importantly, these positive results for the adult subgroup are not held when alternative trials are used to inform the analysis.[107,109] Within the children subgroup, the incremental cost-effectiveness of guidelines plus FeNO monitoring using NObreath versus guidelines alone is expected to be approximately £45,200 per QALY gained. A more favourable ICER was produced when the analysis was based on the trial reported by Pijnenburg *et al*;[115] this may reflect differences in the characteristics of patients recruited to these trials, with the former being uncontrolled. The results in the children and adult subgroups are particularly sensitive to assumptions regarding changes in ICS use over time, the annual number of nurse visits for FeNO monitoring and the duration over which FeNO monitoring is assumed to impact upon exacerbation rates and ICS use.

8.2 Generalisability of results

8.2.1 Generalisability of evidence relating to FeNO in the diagnosis of asthma

- The clinical evidence was heterogeneous in terms of clinical characteristics and results, and studies were selected for modelling based on their similarity to UK practice, and similarity to the subgroups of interest as defined in the protocol (i.e. those difficult to diagnose, or the wider population of those presenting with symptoms of asthma). As such, no single study can be generalised to the whole population and this should be noted when interpreting the results of this assessment.
- Some of the subgroups of interest to the appraisal were not modelled. These groups were the elderly, pregnant women, and smokers/those exposed to environmental tobacco smoke. This was due to limitations in the identified evidence. Only inferences as to the generalisability of results from other studies to these populations can be made.
- The EAG model is "blunt" in that it assumes that all misdiagnoses are assumed to be later corrected by subsequent tests. The model is not specific about what these tests are.
- In addition, all but one of the studies used to inform the diagnostic accuracy parameters (Sivan *et al*[99]) was undertaken in adults. As a consequence, the EAG model does not fully capture differences in the likely diagnostic pathways between children and adult subgroups.

8.2.2 Generalisability of evidence relating to FeNO in the management of asthma

- In adults, the studies used in the model were Shaw 2007[108], Smith 2005[107] and Syk.[109] Each study has its own merits in terms of generalisability.
    - Shaw 2007[108] followed UK practice in terms of the comparator arm management strategy. It also recruited a population from primary care and included mild to severe asthmatics regardless of atopic status. Smokers were excluded, so it is not clear if the results can be generalised to the UK smoking population. It was also not clear which FeNO device was used.
    - Smith 2005[107] recruited what is likely to be a population with mild to moderately severe asthma and used a different step-up/step-down protocol in both the control and intervention arms. It is unclear to what extent this study could be generalisable to the UK population, but it nevertheless provides some insight into the impact that different but plausible efficacy inputs have on the cost-effectiveness estimates.
    - Syk[109] is most notable for having recruited only atopic patients, only non-smokers and only mild to moderate asthmatics. This study is unlikely to have wide generalisability. Again, it nevertheless provides some insight into the impact that different but plausible efficacy inputs have on the cost-effectiveness estimates.

- In children, the two studies that were modelled were Szefler 2008[113] and Pijnenburg 2005,[115] largely because these two studies reported the most complete sets of data, and recruited different populations. Again, each study has its own merits in terms of generalisability.
    - Szefler 2008[113] had the lowest risk of bias amongst the studies available. It also recruited patients who were difficult to treat, one of the subgroups identified in the scope as being of especial interest, and so generalisability may be limited to this group. The step-up/step down protocol within this trial did not however allow for ICS to be decreased on the basis of low FeNO alone, making it less likely that a decrease in ICS will be seen in the intervention arm in comparison to some other protocols. Therefore, the generalisability of this study largely depends on what type of step-up/step-down protocol is likely to be adopted in the UK.
    - Pijnenburg 2005[115] adopted inclusion criteria which were likely to result in a population of asthmatics who have more stable disease. The step-up/step-down protocol also does not allow for ICS to be decreased on the basis of low FeNO alone, requiring that symptoms are also low. As such, the generalisability of this study is also largely depends on what type of step-up/step-down protocol is likely to be adopted in the UK.

- One study was found which recruited pregnant women. The management strategy allowed step-down on the basis of FeNO alone. This study can be generalised within the population of pregnant women.

8.2.3 Equivalence of devices

- As the equivalence of devices is not assured, the generalisability of these results to all three devices is also not assured.
- It is thought that estimates of diagnostic accuracy and efficacy in managing asthma are probably achievable by all devices, as correlation between measurements is good. However, the actual value that should be used as a cut-off in diagnosis and management is much more difficult to generalise, and further research may be required to estimate the most appropriate values.

8.3 Strengths and limitations of the assessment

8.3.1 Strengths of the assessment

The assessment includes systematic reviews of equivalence of devices, diagnostic accuracy, management efficacy and test failures which have been undertaken according to robust and high quality methods.

The scope of the assessment was agreed by NICE and the SCM during an extensive scoping exercise.

The existing economic evidence base models have been formally critiqued using the Drummond checklist and assessed in terms of adherence of the individual studies to the NICE Reference Case.[132]

The two economic models have been developed to a high standard, based on the decision problem rather than being limited by the available empirical evidence. Both EAG models explicitly address the trade-off between expected additional health gains resulting from the more accurate diagnosis of asthma and the health loss associated with displacing existing services. Whilst many of the parameters included in these models are subject to considerable uncertainty, the use of a modelling framework helps elucidate which parameters are likely to be most important for decision-making.

The assessment report has been peer reviewed by NICE, other experienced HTA researchers and leading experts in the diagnosis and management of inflammatory airways diseases.

8.3.2 Limitations of the assessment

This assessment is subject to several limitations. It is important to note that these limitations are principally sourced in the evidence base, rather than the methods used to interrogate and evaluate it. Overall, the evidence base for this assessment was not of the highest quality. No end to end studies were found which estimated the clinical utility of FeNO in the diagnosis of asthma, and no studies were found which used NIOX VERO or NObreath. As such, clinical validity studies were included and a review of the equivalence of devices was conducted. This leads to the following limitations:

- The benefits and harms associated with diagnosis of asthma using FeNO have been estimated based on modelling of the consequences of being true-positive, true-negative, false-positive and false-negative. This includes a large number of assumptions and extrapolations many of which cannot be substantiated with empirical evidence.

- The equivalence of devices is assumed, and this may not hold true in practice. As such, FeNO cut-off values as reported in the primary research may not be applicable to measurements using other devices.
- NObreath will always dominate other devices as its efficacy has been assumed to be equivalent, but its unit cost is less.

No study provided estimates relating to the additional diagnostic value of FeNO to the whole UK diagnostic pathway. This limits the scope of the economic analysis.

No short-term diagnosis of asthma is 100% accurate, and as such all diagnostic studies included in the review had a flawed reference standard. However, in the absence of any alternative, these reference standards were considered to be 100% accurate. A better reference standard would have been long term follow-up of patients. Only one study used such a reference standard (Sivan 2009).[99]

None of the management studies in children included a step-up/step-down protocol that allowed ICS to be stepped down on the basis of FeNO alone. This will limit the degree to which ICS use can be reduced and means that one of the major putative benefits of FeNO management has not actually been assessed empirically: the identification of ICS non-responsive asthmatics who can be taken off ICS therapy with no loss of control.

The EAG diagnostic model is based on evidence identified through the systematic review of FeNO. The diagnostic accuracy of other non-FeNO comparators (spirometry, airways reversibility (MCT) and bronchodilator reversibility) was based on comparative studies identified through the review process. It is possible that other studies not identified within the review could be considered relevant to the model. The use of the Hunter *et al* case control study[151] does however mean that all non-FeNO diagnostic options are assessed consistently within the same study.

The EAG diagnostic model and the Price/Aerocrine diagnostic models draw a number of naïve indirect comparisons across studies; this is a limitation of the evidence base rather than the assessment. It does however limit the confidence that can and should be placed on the findings of these diagnostic models.

The EAG management model is based on short-term evidence of the comparative efficacy of FeNO versus guidelines. The extrapolation of these benefits to the longer-term is subject to

considerable uncertainty. Again, this limitation reflects the evidence base rather than the model itself.

Two previous systematic reviews of the effectiveness of FeNO monitoring to guide management were identified. Petsky et al.[175] performed a Cochrane review in 2008, which was updated with data from two new studies in 2009. A total of six studies were included in the update (two adult studies[87,108] and four children/ adolescents[113,115,176,177]), all of which compared adjustments of asthma therapy based on FeNO with conventional methods (typically clinical symptoms and spirometry). A meta-analysis was performed for seven outcomes: number of patients with > 1 exacerbation; exacerbation rates; FEV1% predicted at the final study visit; FeNO at the final visit; symptom score; ICS dose at final visit; and geometric mean change in FeNO from baseline. There was some suggestion of benefits associated with FeNO for several outcomes, in particular number of subjects with > 1 exacerbation, exacerbation rates, FEV1% predicted at final visit, and geometric change in FeNO from baseline; however, none of these results were statistically conclusive. There were also some results that suggested inconsistent effects between adult and child cohorts. FeNO appeared to have some beneficial effect on symptom score in adults (mean difference: -0.14, 95% CI: -0.42, 0.14), but not children (mean difference: 0.04, 95% CI: -0.11, 0.20), and FeNO management lowered ICS dose in adults (mean difference: -450.03 µg, 95% CI: -676.73, -223.34), but not children (mean difference: 140.18, 95% CI: 28.94, 251.43). Furthermore, there were some limitations to the meta-analysis, particularly with respect to the children's studies. There was substantial clinical heterogeneity among the study cohorts, with no two studies using exactly the same step up/down protocols. The study by de Jongste 2009,[177] which included a telemedical component, was not of relevance to our current assessment, making the results of this meta-analysis not directly applicable to this review.

It can be seen that there is a high degree of agreement between the Petsky et al.[175] review and our own review, especially with realtion to the lack of statistically significant effects, and some differences between adults and children. The strength of our review lies in the inclusion of subsequently published studies, the focus on exacerbation rates rather than number of people with an exacerbation, and the *a priori* separation of both children and pregnant women into different subgroups.

The second review was an AIC manufacturer's submission to NICE.[178] This review updated the meta-analyses of the number of patients with > 1 exacerbation and exacerbation rates from the aforementioned Cochrane review[175] with a study of FeNO-guided asthma management in pregnant women.[111] Inclusion of this study resulted in improvements on all

measures of exacerbations, especially asthma exacerbation rates in adults (mean difference: -0.27, 95% CI: -0.42, -0.12), and relative rate of asthma exacerbations in adults (relative rate: 0.57, 95% CI: 0.41, 0.80). However, since it is known that pregnancy can substantially affect the course of asthma,[179] it was arguably inappropriate to include the cohort of pregnant women in a meta-analysis of adults with asthma. Indeed, in the current review of FeNO-guided management, we have interpreted the results of the Powell study of pregnant women[111] separately from the main results for just this reason.

8.4 Research Recommendations

This appraisal has been limited by several key evidence gaps that would benefit from further research. It could be argued that this technology is currently under-researched and that any conclusions drawn at this stage may be unduly affected by this lack of evidence. However, some of the problems with the evidence base seem intractable in terms of practicalities, and it could also be argued that the available evidence does point towards some benefits to the technology, albeit benefits that are difficult to quantify with certainty.

Some key problems and suggested research priorities are listed here.

- The equivalence of devices is not assured. There are several ways this problem could be addressed, none of which offer a panacea:
    1. Additional extensive equivalence testing of all devices in relation to one another to ascertain what is driving the heterogeneity in study results. This may be expensive and time-consuming, and may still reveal high levels of disagreement between studies owing to the evidence of variability between devices of the same design.
    2. A network meta-analysis of the existing evidence. This was precluded in this project owing to time and resource constraints. There is likely to be a high degree of uncertainty in any such analysis, based on current evidence, and its results may not be useful.
    3. Derivation and validation studies conducted using the devices in question to develop unique cut-off points for each device for management and diagnosis. This may also be expensive and time-consuming.
    4. Explore the option of using intra-subject relative change to assess control when managing asthma. There is already evidence relating to this approach, but it appears to be in comparatively early stages of development. This is not likely to be a useful option in diagnosis.

- Cut-off values are highly variable and are largely based on derivation studies not validation studies. This problem is related to problems with the equivalence of devices. Possible research priorities relating to this include
    1. Large validation studies (possibly preceded by derivation studies) for cut-off values in all populations of interest, using a number of available devices. Whilst expensive and time consuming, these studies could be very valuable.
- The clinical utility of diagnosis of asthma using FeNO compared to current practice is not informed by direct evidence. Possible research priorities relating to this include
    1. A study which charts the clinical utility of diagnosis of asthma using FeNO versus diagnosis of asthma using current guidelines, against a reference standard of long term follow-up of diagnosis to correct for the mis-diagnoses of both diagnostic approaches.
- It is unclear which step-up/step-down protocol offers the best efficacy. Possible research priorities relating to this include
    1. RCT studies which compare different management protocols to one another. It may be that different protocols are necessary in different populations.
    2. Studies which aim to derive the best cut-off points for management protocols. This may be influenced by the specifics of the step-up/step-down protocols.
- It is unclear how treatment effects will progress over time. Long term studies following patients for a number of years could address this evidence gap.

8.5 Conclusions

There is considerable uncertainty associated with all analyses within this assessment. This is largely due to the limitations of the evidence base.

Studies using the devices that are the focus of this review were not available for all analyses, and in the absence of an alternative, equivalence has been assumed between devices. However, there is not a strong indication across the literature to support this assumption.

The clinical evidence relating to the use of FeNO for diagnosis of asthma is highly heterogeneous and difficult to interpret in the context of the insertion of FeNO into a diagnostic pathway. This is compounded by a lack of certainty as to the equivalence of devices used in the primary research studies to the devices that are the focus of this assessment.

The health economic analysis indicates that FeNO could have value in both the diagnostic and management settings. In particular, the diagnostic model indicates that FeNO plus bronchodilator reversibility dominates many other diagnostic tests and may render airway reversibility cost-ineffective. In the management setting, FeNO-guided management has the potential to appear cost-effective, although this is largely dependent on the expected duration over which it continues to impact upon medication decisions. The conclusions drawn from both models require strong technical value judgements with respect to several aspects of the decision problem in which little or no empirical evidence exists.

9. REFERENCES

1. de Roberto, M., Locatelli, F., Sunyer, J., Burney, P. "Differences in Incidence of Reported Asthma Related to Age in Men and Women". *American Journal of Respiratory and Critical Care Medicine* 2000; 162(1):68-74.

2. Green, R.H., Brightling, C.E., McKenna, S., Hargadon, B., Parker, D., Bradding, P. et al. Asthma exacerbations and sputum eosinophil counts: a randomised controlled trial. *Lancet* 2002; 360(9347):1715-1721.

3. Wardlaw, A., Brightling, C., Green, R., Woltmann, G., Pavord, I. Eosinophils in asthma and other allergic diseases. *Br Med Bull* 2000; 56(4):985-1003. Available from http://www.ncbi.nlm.nih.gov/pubmed/11359633

4. Berry, M., Morgan, A., Shaw, D., Parker, D., Green, R., Brightling, C. et al. Pathological features and inhaled corticosteroid response of eosinophilic and non-eosinophilic asthma. *Thorax* 2007; 62(12):1043-1049.

5. Schleich, F.N., Seidel, L., Sele, J., Manise, M., Quaedvlieg, V., Michils, A. et al. Exhaled nitric oxide thresholds associated with a sputum eosinophil count >=3% in a cohort of unselected patients with asthma. *Thorax* 2010; 65(12):1039-1044.

6. Pavord, I.D., Shaw, D.E., Gibson, P.G., Taylor, D.R. Inflammometry to assess airway diseases. *Lancet* 2008; 372(9643):1017-1019.

7. Sur, S., Crotty, T., Kephart, G., Hyma, B., Colby, T., Reed, C. et al. "Sudden-Onset Fatal Asthma: A Distinct Entity with Few Eosinophils and Relatively More Neutrophils in the Airway Submucosa?" . *American Review of Respiratory Disease* 1993; 148(3):713-719.

8. British Thoracic Society/Scottish Intercollegiate Guidelines Network. British Guideline on the Management of Asthma: A national clinical guideline. 2013.

9. Wenzel, S. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nature Medicine* 2012; 18(5):716-725.

10. Gotshall, R. Exercise-Induced Bronchoconstriction. *Drugs* 2002; 62(12):1725-1739.

11. Anderson, H., Gupta, R., Strachan, D., Limb, E. 50 years of asthma: UK trends from 1955 to 2004. *Thorax* 2007; 62(1):85-90. Available from http://www.ncbi.nlm.nih.gov/pubmed/17189533

12. Boodhna, G., Hall, J. Respiratory symptoms and asthma in children. *HSE* 2011; Available from https://catalogue.ic.nhs.uk/publications/public-health/surveys/heal-surv-resp-heal-eng-2010/heal-surv-eng-2010-resp-heal-ch4-symp-chil.pdf

13. Shepherd, J., Rogers, G., Anderson, R., Main, C., Thompson Coon, J., Hartwell, D. et al. ICS and LABAs for the treatment of chronic asthma in adults and children aged 12 years and over: Systematic review and economic analysis. Final report to NICE. 2006.

14. Burr, M., Davies, B., Hoare, A., Jones, A., Williamson, I., Holgate, S. et al. A confidential inquiry into asthma deaths in Wales. *Thorax* 1999; 54:985-989.

15. Bucknall, C., Slack, R., Godley, C., Mackay, T., Wright, S. Scottish Confidential Inquiry into Asthma Deaths (SCIAD), 1994-6. *Thorax* 1999; 54(978):984.

16. Sturdy, P., Victor, C., Anderson, H., Bland, J., Butland, B., Harrison, B. et al. Psychological, social and health behaviour risk factors for deaths certified as asthma: a national case-control study. *Thorax* 2002; 57:1034-1039.

17. Sturdy, P., Butland, B., Anderson, H., Ayres, J.G., Bland, J., Harrison, B. et al. Deaths certified as asthma and use of medical services: a national case-control study. *Thorax* 2005; 60:909-915.

18. Jones, K., Berrill, W., Bromly, C., Hendrick, D. A confidential enquiry into certified asthma deaths in the North of England, 1994-96: influence of co-morbidity and diagnostic inaccuracy. *Respir Med* 1999; 93:923-927.

19. Brown, P., Greville, H., Finucane, K. Asthma and irreversible airflow obstruction. *Thorax* 1984; 39(2):131-136.

20. The Childhood Asthma Management Program Research Group. Long-term effects of budesonide or nedocromil in children with asthma. *N Eng J Med* 2000; 343(15):1054-1063.

21. Agertoft, L., Pederson, S. Effect of long-term treatment with inhaled budesonide on adult height in children with asthma. *N Eng J Med* 2000; 343(15):1064-1069.

22. Dweik, R.A., Boggs, P.B., Erzurum, S.C., Irvin, C.G., Leigh, M.W., Lundberg, J.O. et al. An official ATS clinical practice guideline: interpretation of exhaled nitric oxide levels (FENO) for clinical applications. *American Journal of Respiratory & Critical Care Medicine* 184[5], 602-615.

23. Kostikas, K., Minas, M., Papaioannou, A.I., Papiris, S., Dweik, R.A., Kostikas, K. et al. Exhaled nitric oxide in asthma in adults: the end is the beginning?. [Review]. *Current Medicinal Chemistry* 2011; 18(10):1423-1431.

24. Aerocrine. 'Clinical guide to Interpretation of FeNO Values'. Sponsor's submission. 2013.

25. National Institute for Health and Care Excellence. Measurement of exhaled nitric oxide concentration in asthma; NIOX MINO and NObreath. Final scope. 2013.

26. Petsky, H.L., Cates, C.J., Li, A., Kynaston, J.A., Turner, C., Chang, A.B. et al. Tailored interventions based on exhaled nitric oxide versus clinical symptoms for asthma in children and adults. [Review] [34 refs][Update of Cochrane Database Syst Rev. 2008;(2):CD006340; PMID: 18425949]. *Cochrane Database of Systematic Reviews* 2009;(4):CD006340.

27. Jartti, T., Wendelin-Saarenhovi, M., Heinonen, I., Hartiala, J., Vanto, T., Jartti, T. et al. Childhood asthma management guided by repeated FeNO measurements: a meta-analysis. [Review]. *Paediatric Respiratory Reviews* 2012; 13(3):178-183.

28. Spahn, A.J.D., Malka, J., Mahr, T.A., Dorinsky, P.M. Meta analysis of asthma exacerbation rates in pediatric studies during asthma managed using fractional exhaled nitric oxide versus standard clinical parameters. *Journal of Allergy and Clinical Immunology* 2013; 131(2 SUPPL. 1):AB194.

29. Reference Manager Bibliographic Software [Thomson Reuters: Philadelphia, PA; 2013.

30. National Institute for Health and Clinical Excellence. Clinical guideline development methods. 2012.

31. Measurement, O. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. *American Journal of Respiratory and Critical Care Medicine* 2005; 171(8):912-930.

32. Anonymous. Cochrane Handbook for systematic reviews of interventions Version 5.1.0. 2011.

33. Centre for Reviews and Dissemination. Systematic reviews. CRD's guidance for undertaking reviews in healthcare. 2009.

34. Whiting, P., Rutjes, A., Westwood, M., Mallett, S., Deeks, J., Reitsma, J. et al. "QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies.". *Annals of Internal Medicine* 2011; 155(8):529-536.

35. RevMan 5. 2011; available from http://ims.cochrane.org/revman

36. Olaguibel, J.M., Parra, A., Alvarez, M.J., Quirce, S., Lopez, R., Olaguibel, J.M. et al. Measurements of fractional exhaled nitric oxide with 2 portable electrochemical sensors: a comparative study. *Journal of Investigational Allergology & Clinical Immunology* 2011; 21(4):322-323.

37. Bland, M., Altman, D. "Statistical methods for assessing agreement between two methods of clinical measurement.". *The Lancet* 1986; 327(8476):307-310.

38. Menzies, D., Nair, A., Lipworth, B.J., Menzies, D., Nair, A., Lipworth, B.J. Portable exhaled nitric oxide measurement: Comparison with the "gold standard" technique. *Chest* 2007; 131(2):410-414.

39. Grob, N.M., Laskowski, D., Dweik, R.A., Grob, N.M., Laskowski, D., Dweik, R.A. A technical report on exhaled nitric oxide measurement: asthma monitoring in athletes. *Journal of Breath Research* 2008; 2(3):37027.

40. Korn, S., Telke, I., Kornmann, O., Buhl, R., Korn, S., Telke, I. et al. Measurement of exhaled nitric oxide: comparison of different analysers. *Respirology* 2010; 15(8):1203-1208.

41. Chen, W., Purohit, A., Barnig, C., Casset, A., de, B.F., Chen, W. et al. Niox and Niox Mino: comparison of exhaled NO in grass pollen allergic adult volunteers. *Allergy* 2007; 62(5):571-572.

42. Hemmingsson, T., Linnarsson, D., Gambert, R., Hemmingsson, T., Linnarsson, D., Gambert, R. Novel hand-held device for exhaled nitric oxide-analysis in research and clinical applications.[Erratum appears in J Clin Monit Comput. 2005 Dec;19(6):463-4]. *Journal of Clinical Monitoring & Computing* 2004; 18(5-6):379-387.

43. Pizzimenti, S., Bugiani, M., Piccioni, P., Heffler, E., Carosso, A., Guida, G. et al. Exhaled nitric oxide measurements: correction equation to compare hand-held device to stationary analyzer. *Respiratory Medicine* 2008; 102(9):1272-1275.

44. Khalili, B., Boggs, P.B., Bahna, S.L., Khalili, B., Boggs, P.B., Bahna, S.L. Reliability of a new hand-held device for the measurement of exhaled nitric oxide. *Allergy* 2007; 62(10):1171-1174.

45. Alving, K., Janson, C., Nordvall, L. Performance of a new hand-held device for exhaled nitric oxide measurement in adults and children. *Respiratory research* 2006; 7 Available from http://respiratory-research.com/content/7/1/67

46. Vahlkvist, S., Sinding, M., Skamstrup, K., Bisgaard, H. Daily home measurements of exhaled nitric oxide in asthmatic children during natural birch pollen exposure. *Journal of Allergy and Clinical Immunology* 2006; 117(6):1272-1276.

47. McGill, C., Malik, G., Turner, S.W., McGill, C., Malik, G., Turner, S.W. Validation of a hand-held exhaled nitric oxide analyzer for use in children. *Pediatric Pulmonology* 2006; 41(11):1053-1057.

48. Kalliola, S., Malmberg, P., Rito, T., Pelkonen, A.S., Makela, M.J., Kalliola, S. et al. Can we use portable nitric oxide analyzer in young children? *Pediatric Pulmonology* 2011; 46(7):627-631.

49. Ozier, A., Girodet, P.O., Marthan, R., Berger, P. Reliability of the hand-held NIOX MINO device for monitoring asthma control. *Fundamental and Clinical Pharmacology* 2010; 24:105.

50. Ozier, A., Girodet, P.-O., Bara, I., Tunon De, L.J.M., Marthan, R., Berger, P. Control maintenance can be predicted by exhaled NO monitoring in asthmatic patients. *Respiratory medicine* 2011; 105(7):989-996.

51. Fortuna, A.M., Fexias, T., Gonzalez, M., Casan, P. Determination of exhaled nitric oxide in healthy volunteers and asthmatic patients - comparison of two devices [Abstract]. *European Respiratory Journal* 2006; 28:791s.

52. Fortuna, A.M., Feixas, T., Casan, P., Fortuna, A.M., Feixas, T., Casan, P. [Measurement of fraction of exhaled nitric oxide with the portable NIOX-MINO monitor in healthy adults]. [Spanish]. *Archivos De Bronconeumologia* 2007; 43(3):176-179.

53. Fukuhara, A., Saito, J., Sato, S., Sato, Y., Nikaido, T., Saito, K. et al. Validation study of asthma screening criteria based on subjective symptoms and fractional exhaled nitric oxide. *Annals of Allergy, Asthma and Immunology* 2011; 107(6):480-486.

54. Kim, S.H., Moon, J.Y., Kwak, H.J., Kim, S.I., Park, D.W., Kim, J.W. et al. Comparison of two exhaled nitric oxide analyzers: the NIOX MINO hand-held electrochemical analyzer and the NOA280i stationary chemiluminescence analyzer. *Respirology* 2012; 17(5):830-834.

55. Yoon, H., Kwak, H., Kim, T., Sohn, J., Shin, D., Park, S. et al. Measurement of exhaled nitric oxide: Comparison of two analyzers. *Allergy: European Journal of Allergy and Clinical Immunology* 2011; 66:535.

56. de Laurentiis, G., Maniscalco, M., Cianciulli, F., Stanziola, A., Marsico, S., Lundberg, J.O. et al. Exhaled nitric oxide monitoring in COPD using a portable analyzer. *Pulmonary Pharmacology & Therapeutics* 2008; 21(4):689-693.

57. Peche, R., Michils, A., Baldassarre, S., Mourid, Z., Muylem, A. Exhaled nitric oxide measured by using NIOX MINO and chemo-luminesence analyser: a comparative study [Abstract]. *European Respiratory Journal* 2007; 30:365s.

58. Michils, A., Peche, R., Baldassarre, S., Mourid, Z., Van, M.A. Comparisons between portable and chemoluminescence exhaled nitric oxide measurements. *European Respiratory Journal* 2008; 32(1):243-244. Available from http://erj.ersjournals.com/cgi/reprint/32/1/243

59. Logan Research Ltd. Benchmark testing report. comparison testing carried out by Logan Fesearch Ltd for Bedfont Scientific Ltd - Reference the NObreath Nitric Oxide monitoring product. Internal report supplied by Bedfont (manufacturers). 2013.

60. Park, Y.A., Lee, J.H., Lee, Y.J., Song, T.W., Kim, K.W., Sohn, M.H. et al. Comparison of exhaled nitric oxide analysers in childhood asthma. *Journal of Allergy and Clinical Immunology* 2011; 127(2 SUPPL. 1):AB62.

61. 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. *Pediatric Allergy and Immunology* 2009; 20(7):679-685.

62. Chladkova, J., Havlinova, Z., Chyba, T., Krcmova, I., Chladek, J. Analysis of single-breath profiles of exhaled nitric oxide in children with allergy and asthma: guideline-derived plateau concentrations compared to results of automatic evaluation by two analyzers. *Journal of Asthma* 2008; 45(9):820-826.

63. Boot, J., de Ridder, L., de Kam, M., Calderon, C., Mascelli, M., Diamant, Z. Comparison of exhaled Nitric Oxide measurements between Niox Mino electrochemical and Ecomedics chemiluminescence analyzer. *Respiratory Medicine* 2008; 102:1667-1671.

64. Lund, T.K., Lund, T.K. Asthma in elite athletes: how do we manage asthma-like symptoms and asthma in elite athletes? *The Clinical Respiratory Journal* 2009; 3(2):123.

65. Fukuhara, A., Sato, S., Saito, J., Sato, Y., Nikaido, T., Inokoshi, Y. et al. Conversion equations of fractional exhaled nitric oxide (FeNO) levels measured by two portable and a stationary analyzers. Unpublished abstract submitted by Bedfont (manufacturers). 2013.

66. Hedlund, A. A randomized, multi-center study to determine the agreement between the NIOX MINO® Nitric Oxide Monitoring system and the NIOX VERO@ device using the 1O-second exhalation mode. Sponsors submission received 18th July 2013. 2013.

67. Pisi, R., Aiello, M., Tzani, P., Marangio, E., Olivieri, D., Chetta, A. et al. Measurement of fractional exhaled nitric oxide by a new portable device: comparison with the standard technique. *Journal of Asthma* 2010; 47(7):805-809.

68. Antus, B., Horvath, I., Barta, I., Antus, B., Horvath, I., Barta, I. Assessment of exhaled nitric oxide by a new hand-held device. *Respiratory Medicine* 2010; 104(9):1377-1380.

69. Hiel, D., von, S., I, Sundblad, B.M., Larsson, K., Palmberg, L., Hiel, D. et al. Evaluation of respiratory effects related to high-pressure cleaning in a piggery with

and without robot pre-cleaning. *Scandinavian Journal of Work, Environment & Health* 2009; 35(5):376-383.

70. Kapande, K.M., McConaghy, L.A., Douglas, I., McKenna, S., Hughes, J.L., McCance, D.R. et al. Comparative repeatability of two handheld fractional exhaled nitric oxide monitors. *Pediatric Pulmonology* 2012; 47(6):546-550.

71. Kapande, K.M., Shields, M.D., McConaghy, L.A. Comparative repeatability of two handheld fractional exhaled nitric oxide monitors. *Irish Journal of Medical Science* 2011; 180:S460.

72. Hussain, S., Laumbach, R., Coleman, J., Youssef, H., Kelly-McNeil, K., Ohman-Strickland, P. et al. Controlled exposure to diesel exhaust causes increased nitrite in exhaled breath condensate among subjects with asthma. *Journal of Occupational & Environmental Medicine* 2012; 54(10):1186-1191.

73. Peden, D., Reed, C.E. Environmental and occupational allergies. *Journal of Allergy and Clinical Immunology* 2010; 125(2 SUPPL. 2):S150-S160.

74. Global strategy for asthma management and prevention. *Global Initiative for Asthma* 2007; Available from http://www.ginasthma.org

75. Schneider, A., Schwarzbach, J., Faderl, B., Welker, L., Karsch-Volk, M., Jorres, R.A. FENO measurement and sputum analysis for diagnosing asthma in clinical practice. *Respiratory medicine* 2013; 107(2):209-216.

76. Fortuna, A.M., Feixas, T., Gonzalez, M., Casan, P., Fortuna, A.M., Feixas, T. et al. Diagnostic utility of inflammatory biomarkers in asthma: exhaled nitric oxide and induced sputum eosinophil count.[Erratum appears in Respir Med. 2012 Apr;106(4):599]. *Respiratory Medicine* 2007; 101(11):2416-2421.

77. Schneider, A., Tilemann, L., Schermer, T., Gindner, L., Laux, G., Szecsenyi, J. et al. Diagnosing asthma in general practice with portable exhaled nitric oxide measurement - results of a prospective diagnostic study. *Respiratory research* 2009; 10 Available from http://respiratory-research.com/content/10/1/15

78. Schneider, A., Tilemann, L., Schermer, T., Gindner, L., Laux, G., Szecsenyi, J. et al. 'Diagnosing Asthma in General Practice with Portable Exhaled Nitric Oxide Measurement - Results of a Prospective Diagnostic Study: FENO <= 16 ppb better than FENO <= 12 ppb to rule out mild and moderate to severe asthma. *Respiratory research* 2009; 10 Available from http://respiratory-research.com/content/10/1/64

79. Hsu, J.Y., Wang, C.Y., Cheng, Y.W., Chou, M.C. Optimal value of fractional exhaled nitric oxide in inhaled corticosteroid treatment for patients with chronic cough of unknown cause. *Journal of the Chinese Medical Association* 2013; 76(1):15-19.

80. Hahn, P.Y., Morgenthaler, T.I., Lim, K.G. Use of exhaled nitric oxide in predicting response to inhaled corticosteroids for chronic cough. *Mayo Clinic Proceedings* 2007; 82(11):1350-1355.

81. Sato, S., Saito, J., Sato, Y., Ishii, T., Xintao, W., Tanino, Y. et al. Clinical usefulness of fractional exhaled nitric oxide for diagnosing prolonged cough. *Respiratory medicine* 2008; 102(10):1452-1459.

82. Prieto, L., Ferrer, A., Ponce, S., Palop, J., Marin, J. Exhaled nitric oxide measurement is not useful for predicting the response to inhaled corticosteroids in subjects with chronic cough. *Chest* 2009; 136(3):816-822. Available from http://www.chestjournal.org/content/136/3/816.full.pdf+html

83. Schleich, F.N., Asandei, R., Manise, M., Sele, J., Seidel, L., Louis, R. Is FENO50 useful diagnostic tool in suspected asthma? *International Journal of Clinical Practice* 2012; 66(2):158-165.

84. ElHalawani, S.M., Ly, N.T., Mahon, R.T., Amundson, D.E. Exhaled nitric oxide as a predictor of exercise-induced bronchoconstriction. *Chest* 2003; 124(2):639-643.

85. Arora, R., Thornblade, C.E., Dauby, P.A.L., Flanagan, J.W., Bush, A.C., Hagan, L.L. Exhaled nitric oxide levels in military recruits with new onset asthma. *Allergy and Asthma Proceedings* 2006; 27(6):493-498.

86. Heffler, E., Guida, G., Marsico, P., Bergia, R., Bommarito, L., Ferrero, N. et al. Exhaled nitric oxide as a diagnostic test for asthma in rhinitic patients with asthmatic symptoms. *Respiratory medicine* 2006; 100(11):1981-1987.

87. Smith, A.D., Cowan, J.O., Brassett, K.P., Filsell, S., McLachlan, C., Monti-Sheehan, G. et al. Exhaled nitric oxide: a predictor of steroid response. *American Journal of Respiratory & Critical Care Medicine* 2005; 172(4):453-459.

88. de la Barra, S.L., Smith, A.D., Cowan, J.O., Herbison, G.P., Taylor, D.R. Predicted versus absolute values in the application of exhaled nitric oxide measurements. *Respiratory Medicine* 2011; 105(11):1629-1634.

89. Pedrosa, M., Cancelliere, N., Barranco, P., Lopez-Carrasco, V., Quirce, S. Usefulness of Exhaled Nitric Oxide for Diagnosing Asthma. *Journal of Asthma* 2010; 47(7):817-821.

90. Smith, A.D., Cowan, J.O., Filsell, S., McLachlan, C., Monti-Sheehan, G., Jackson, P. et al. Diagnosing asthma: comparisons between exhaled nitric oxide measurements and conventional tests. *American Journal of Respiratory & Critical Care Medicine* 2004; 169(4):473-478.

91. Cordeiro, D., Rudolphus, A., Snoey, E., Braunstahl, G.-J. Utility of nitric oxide for the diagnosis of asthma in an allergy clinic population. *Allergy and Asthma Proceedings* 2011; 32(2):119-126. Available from http://docserver.ingentaconnect.com/deliver/connect/ocean/10885412/v32n2/s8.pdf?expires=1301406513&id=61992699&titleid=6151&accname=Elsevier+Science&checksum=BAE762D5A46CF56B3CC4A9B7C32379A4

92. Bobolea, I.D., Barranco, P., Lopez-Carrasco, V., Calderon, O., Guillen, D., Quirce, S. Is methacholine challenge sufficient to rule out bronchial hyperresponsiveness in patients with suspected asthma? *Journal of Allergy and Clinical Immunology* 2012; 129(2 SUPPL. 1):AB3.

93. Zhang, Y.M., Lin, J.T., Su, N., Chen, X., Liu, G.L., Yu, H.X. et al. [Values of fractional exhaled nitric oxide in the diagnosis of chronic cough]. [Chinese]. *Chung-Hua i Hsueh Tsa Chih [Chinese Medical Journal]* 2011; 91(18):1254-1258.

94. Pizzimenti, S., Heffler, E., Piccioni, P., Bugiani, M., Migliore, E., Guida, G. et al. Usefulness of exhaled nitric oxide (FeNO) measured by a portable analyzer to

diagnose Cough Variant asthma in a clinical setting of chronic cough. *Allergy: European Journal of Allergy and Clinical Immunology* 2009; 64:395.

95. Mathew, S., Cliff, I., Agarwal, S., Lim, A., Allen, M., Mustfa, N. Relationship between exhaled nitric oxide and methacholine challenge test in suspected asthma. *American journal of respiratory and critical care medicine* 2011; 183(1 MeetingAbstracts) Available from http://ajrccm.atsjournals.org/cgi/reprint/183/1_MeetingAbstracts/A5554?sid=0209d9f3-c74d-409d-8e6e-81ebe76d2328

96. Whiting, P., Rutjes, A., Reitsma, J., Bossuyt, P., Kleijnen, J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology* 2003; 3(1):25.

97. Linkosalo, L., Lehtimaki, L., Holm, K., Kaila, M., Moilanen, E., Linkosalo, L. et al. Relation of bronchial and alveolar nitric oxide to exercise-induced bronchoconstriction in atopic children and adolescents. *Pediatric Allergy & Immunology* 2012; 23(4):360-366.

98. Ramser, M., Hammer, J., Amacher, A., Trachsel, D. The value of exhaled nitric oxide in predicting bronchial hyperresponsiveness in children. *Journal of Asthma* 2008; 45(3):191-195.

99. Sivan, Y., Gadish, T., Fireman, E., Soferman, R. The Use of Exhaled Nitric Oxide in the Diagnosis of Asthma in School Children. *Journal of Pediatrics* 2009; 155(2):211-216.

100. Woo, S.-I., Lee, J.-H., Kim, H., Kang, J.-W., Sun, Y.-H., Hahn, Y.-S. Utility of fractional exhaled nitric oxide (FENO) measurements in diagnosing asthma. *Respiratory medicine* 2012; 106(8):1103-1109.

101. Malinovschi, A., Backer, V., Harving, H., Porsbjerg, C. The value of exhaled nitric oxide to identify asthma in smoking patients with asthma-like symptoms. *Respiratory Medicine* 2012; 106(6):794-801.

102. Mahut, B., Peyrard, S., Delclaux, C., Mahut, B., Peyrard, S., Delclaux, C. Exhaled nitric oxide and clinical phenotypes of childhood asthma. *Respiratory Research* 2011; 12:65.

103. Hanson, J., DeLurgio, S., Williams, D., Dinakar, C. Ambulatory fractional exhaled nitric oxide (FENO) measurement in children 4-7years ofage. *Annals of Allergy, Asthma and Immunology* 2012; 109:A22.

104. Riva-Velasco, E., Krishnan, S., Dozor, A.J. Relationship between exhaled nitric oxide and exposure to low-level environmental tobacco smoke in children with asthma on inhaled corticosteroids. *Journal of Asthma* 2012; 49(7):673-678.

105. Tamasi, L., Bohacs, A., Bikov, A., Andorka, C., Rigo, J., Jr., Losonczy, G. et al. Exhaled nitric oxide in pregnant healthy and asthmatic women. *Journal of Asthma* 2009; 46(8):786-791.

106. Simpson, J.L., McDonald, V.M., Gibson, P.G. Exhaled nitric oxide is not a marker of eosinophilic inflammation in older Australians. *Respirology* 2010; 15:A53.

107. Smith, A.D., Cowan, J.O., Brassett, K.P., Herbison, G.P., Taylor, D.R., Smith, A.D. et al. Use of exhaled nitric oxide measurements to guide treatment in chronic asthma. *New England Journal of Medicine* 2005; 352(21):2163-2173.

108. Shaw, D.E., Berry, M.A., Thomas, M., Green, R.H., Brightling, C.E., Wardlaw, A.J. et al. The use of exhaled nitric oxide to guide asthma management: a randomized controlled trial. *American Journal of Respiratory & Critical Care Medicine* 2007; 176(3):231-237.

109. Syk, J., Malinovschi, A., Johansson, G., Unden, A.-L., Andreasson, A., Lekander, M. et al. Anti-inflammatory treatment of atopic asthma guided by exhaled nitric oxide: a randomized controlled trial. 2013.

110. Calhoun, W.J., Ameredes, B.T., King, T.S., Icitovic, N., Bleecker, E.R., Castro, M. et al. Comparison of physician-, biomarker-, and symptom-based strategies for adjustment of inhaled corticosteroid therapy in adults with asthma: the BASALT randomized controlled trial. *JAMA* 2012; 308(10):987-997.

111. Powell, H., Murphy, V.E., Taylor, D.R., Hensley, M.J., McCaffery, K., Giles, W. et al. Management of asthma in pregnancy guided by measurement of fraction of exhaled nitric oxide: a double-blind, randomised controlled trial. *Lancet* 2011; 378(9795):983-990.

112. Fritsch, M., Uxa, S., Horak, F., Jr., Putschoegl, B., Dehlink, E., Szepfalusi, Z. et al. Exhaled nitric oxide in the management of childhood asthma: a prospective 6-months study. *Pediatric Pulmonology* 2006; 41(9):855-862.

113. Szefler, S.J., Mitchell, H., Sorkness, C.A., Gergen, P.J., O'Connor, G.T., Morgan, W.J. et al. Management of asthma based on exhaled nitric oxide in addition to guideline-based treatment for inner-city adolescents and young adults: a randomised controlled trial. *Lancet* 2008; 372(9643):1065-1072.

114. Verini, M., Consilvio, N.P., Di, P.S., Cingolani, A., Spagnuolo, C., Rapino, D. et al. FeNO as a Marker of Airways Inflammation: The Possible Implications in Childhood Asthma Management. *Journal Of Allergy* 2010; 2010, 2010.

115. Pijnenburg, M.W., Bakker, E.M., Hop, W.C., de Jongste, J.C., Pijnenburg, M.W., Bakker, E.M. et al. Titrating steroids on exhaled nitric oxide in children with asthma: a randomized controlled trial. *American Journal of Respiratory & Critical Care Medicine* 2005; 172(7):831-836.

116. Petsky, H., Li, A.M., Kynaston, J.A., Turner, C., Chang, A.B. Dual-center randomised trial on tailored asthma therapy based on exhaled nitric oxide (FENO) versus routine clinical care. *American journal of respiratory and critical care medicine* 2010; 181(1 MeetingAbstracts) Available from http://ajrccm.atsjournals.org/cgi/reprint/181/1_MeetingAbstracts/A3928?sid=272f60a2-7bd7-44a2-80d7-0a7591379e41

117. Scott, M., Raza, A., Karmaus, W., Mitchell, F., Grundy, J., Kurukulaaratchy, R.J. et al. Influence of atopy and asthma on exhaled nitric oxide in an unselected birth cohort study. *Thorax* 2010; 65(3):258-262.

118. Frischer, T., Eber, E., Eichler, I., Horak, E., Riedler, J., Gotz, M. et al. Consensus guidelines for drug therapy of bronchial asthma in children and adolescents. *Wien Klin Wochenschr* 1999; 111:900-902.

119. National Asthma Education and Prevention Program. *National Heart, Lung, and Blood Institute (NHLBI)* 2013; Available from http://www.nhlbi.nih.gov/about/naepp/

120. Pijnenburg, M.W., Lissenberg, E.T., Hofhuis, W., Ghiro, L., Ho, W.C., Holland, W.P. et al. Exhaled nitric oxide measurements with dynamic flow restriction in children aged 4-8 yrs. *European Respiratory Journal* 2002; 20(4):919-924.

121. Beck-Ripp, J., Griese, M., Arenz, S., Koring, C., Pasqualoni, B., Bufler, P. et al. Changes of exhaled nitric oxide during steroid treatment of childhood asthma. *European Respiratory Journal* 2002; 19(6):1015-1019.

122. Inoue, H., Niimi, A., Takeda, T., Matsumoto, H., Ito, I., Otsuka, K. et al. Pathophysiological characteristics of asthma in the elderly. *American journal of respiratory and critical care medicine* 2010; 181(1 MeetingAbstracts) Available from http://ajrccm.atsjournals.org/cgi/reprint/181/1_MeetingAbstracts/A5096?sid=b20c1fa6-2f0d-46eb-a9ed-8be599e0e2e0

123. Columbo, M., Wong, B., Panettieri, R.A., Rohr, A.S. Asthma in the elderly: The role of exhaled nitric oxide measurements. *Journal of Allergy and Clinical Immunology* 2012; 129(2 SUPPL. 1):AB8.

124. Roh, Y.H. Usefulness of asthma control test questionnaire, FEV1 and Exhaled Nitric Oxide Level (FENO) for the clinical assessment of elderly asthma. *World Allergy Organization Journal* 2012; 5:S160.

125. Smith, A.M., Villareal, M., Bernstein, D.I., Swikert, D.J., Smith, A.M., Villareal, M. et al. Asthma in the elderly: risk factors and impact on physical function. *Annals of Allergy, Asthma, & Immunology* 2012; 108(5):305-310.

126. Ross, J., Baptist, A. Factorsassociatedwithasthmaquality of lifeand controlamong olderadults. *Annals of Allergy, Asthma and Immunology* 2011; 107(5 SUPPL. 1):A5.

127. Michils, A., Louis, R., Peche, R., Baldassarre, S., Van, M.A. Exhaled nitric oxide as a marker of asthma control in smoking patients. *European Respiratory Journal* 2009; 33(6):1295-1301. Available from http://erj.ersjournals.com/cgi/reprint/33/6/1295

128. Kostikas, K., Papaioannou, A.I., Tanou, K., Giouleka, P., Koutsokera, A., Minas, M. et al. Exhaled NO and exhaled breath condensate pH in the evaluation of asthma control. *Respiratory medicine* 2011; 105(4):526-532.

129. Hromis, S., Milutinov, S., Zvezdin, B., Kopitovic, I., Maksimovic, O. Exhaled nitric oxide in assessing of asthma control in smoking patients pilot study. *Allergy* 2012; 67:461.

130. Lehtimaki, L., Turjanmaa, V., Kankaanranta, H., Saarelainen, S., Hahtola, P., Moilanen, E. Increased bronchial nitric oxide production in patients with asthma measured with a novel method of different exhalation flow rates. *Annals of Medicine* 2000; 32(6):417-423.

131. Drummond, M., Sculpher, M., Torrance, G., O'Brien, B., Stoddart, G. Methods for the Economic Evaluation of Health Care Programmes. 2005. Oxford, Oxford University Press.

132. NICE. Diagnostics Assessment Programme manual. Manchester; 2011.

133. Weinstein, M., O'Brien, B., Hornberger, J., Jackson, J., Johannesson, M., McCabe, C. et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. *Value in Health* 2003; 6(1):9-17.

134. Price, D., Berg, J., Lindgren, P., Price, D., Berg, J., Lindgren, P. An economic evaluation of NIOX MINO airway inflammation monitor in the United Kingdom. *Allergy* 2009; 64(3):431-438.

135. Berg, J., Lindgren, P. Economic evaluation of FE(NO) measurement in diagnosis and 1-year management of asthma in Germany. *Respir Med* 2008; 102(2):219-231.

136. Peters, J., Stevenson, M., Beverley, C., Lim, J.N.W., Smith, S. The clinical effectiveness and cost-effectiveness of inhaler devices used in the routine management of chronic asthma in older children: A systematic review and economic evaluation. *Health Technology Assessment* 2002; 6(5)

137. Briggs, A.H., Bousquet, J., Wallace, M.V., Busse, W.W., Clark, T.J., Pedersen, S.E. et al. Cost-effectiveness of asthma control: an economic appraisal of the GOAL study. *Allergy* 2006; 61(5):531-536.

138. Doull, I., Price, D., Thomas, M., Hawkins, N., Stamuli, E., Tabberer, M. et al. Cost-effectiveness of salmeterol xinafoate/fluticasone propionate combination inhaler in chronic asthma. *Current Medical Research and Opinion* 2007; 23(5):1147-1159.

139. Wilson, E.C., Price, D., Musgrave, S.D., Sims, E.J., Shepstone, L., Murdoch, J. et al. Cost effectiveness of leukotriene receptor antagonists versus long-acting beta-2 agonists as add-on therapy to inhaled corticosteroids for asthma: a pragmatic trial. *PharmacoEconomics* 2010; 28(7):597-608.

140. Paggiaro, P., Buseghin, G., Nicolini, G., Patel, S., Iannazzo, S., Zaniolo, O. et al. The cost-effectiveness of step down from high dose ics/laba combination therapy in asthma in the UK setting. *Value in Health* 2011; 14(7):A493-A494.

141. Booth, P.C., Capsey, L.J., Langdon, C.G., Wells, N.E.J. A comparison of the cost-effectiveness of alternative prophylactic therapies in the treatment of adult asthma. *British Journal of Medical Economics* 1995; 8(2):65-72.

142. Barnes, N.C., Thwaites, R.M., Price, M.J., Barnes, N.C., Thwaites, R.M., Price, M.J. The cost-effectiveness of inhaled fluticasone propionate and budesonide in the treatment of asthma in adults and children. *Respiratory Medicine* 1999; 93(6):402-407.

143. Payne, D.N., Qiu, Y., Zhu, J., Peachey, L., Scallan, M., Bush, A. et al. Airway inflammation in children with difficult asthma: relationships with airflow limitation and persistent symptoms. *Thorax* 2004; 59(10):862-869.

144. Everden, P., Lloyd, A., Hutchinson, J., Plumb, J., Everden, P., Lloyd, A. et al. Cost-effectiveness of eformoterol Turbohaler versus salmeterol Accuhaler in children with symptomatic asthma. *Respiratory Medicine* 2002; 96(4):250-258.

145. Price, M.J., Briggs, A.H. Development of an economic model to assess the cost effectiveness of asthma management strategies. *PharmacoEconomics* 2002; 20(3):183-194.

146. Buxton, M.J., Sullivan, S.D., Andersson, L.F., Lamm, C.J., Liljas, B., Busse, W.W. et al. Country-specific cost-effectiveness of early intervention with budesonide in mild asthma. *European Respiratory Journal* 2004; 24(4):568-574.

147. Price, D., Wiren, A., Kuna, P., Price, D., Wiren, A., Kuna, P. Cost-effectiveness of budesonide/formoterol for maintenance and reliever asthma therapy. *Allergy* 2007; 62(10):1189-1198.

148. Kemp, L., Haughney, J., Barnes, N., Sims, E., Von, Z.J., Hillyer, E.V. et al. Cost-effectiveness analysis of corticosteroid inhaler devices in primary care asthma management: A real world observational study. *Clinicoeconomics & Outcomes Research* 2010; 2:75-85.

149. Aerocrine Ltd. Submission to NICE - Assessing the Impact of FeNO in the Management and Diagnosis of Asthma. Slideset and Excel model. 2013. 2013.

150. Smith, A.D., Taylor, D.R. Is exhaled nitric oxide measurement a useful clinical test in asthma? *Curr Opin Allergy Clin Immunol* 2005; 5(1):49-56.

151. Hunter, C., Brightling, C., Woltmann, G., Wardlow, A., Pavord, I. A comparison of the validity of different diagnostic tests in adults with asthma. *Chest* 2002; 121:1051-1057.

152. British Thoracic Society and Scottish Intercollegiate Guidelines Network. British Guideline on the Management of Asthma, 2007 update. 2007.

153. Department of Health. NHS Reference Costs 2005. 2005. DH:London.

154. Curtis, L., Netten, A. Unit Costs of Health and Social Care 2005. 2005. Personal Social Services Research Unit at the University of Kent, Canterbury, UK.

155. British Medical Association. British National Formulary 51. 2006. London:Pharmaceutical Press.

156. Dupont, L.J., Demedts, M.G., Verleden, G.M., Dupont, L.J., Demedts, M.G., Verleden, G.M. Prospective evaluation of the validity of exhaled nitric oxide for the diagnosis of asthma. *Chest* 2003; 123(3):751-756.

157. Jayaram, L., Pizzichini, M.M., Cook, R.J., Boulet, L.P., LemiÇùre, C., Pizzichini, E. et al. Determining asthma treatment by monitoring sputum cell counts: effect on exacerbations. *The European Respiratory Journal : Official Journal of the European Society for Clinical Respiratory Physiology* 2006; 27(3):483-494.

158. Andersson, F., Stahl, E., Barnes, P.J., LÇôfdahl, C.G., O'Byrne, P.M., Pauwels, R.A. et al. Adding formoterol to budesonide in moderate asthma--health economic results from the FACET study. *Respiratory Medicine* 2001; 95(6):505-512.

159. Tattersfield, A.E., Postma, D.S., Barnes, P.J., Svensson, K., Bauer, C.A., O'Byrne, P.M. et al. Exacerbations of asthma: a descriptive study of 425 severe exacerbations. The FACET International Study Group. *American Journal of Respiratory and Critical Care Medicine* 1999; 160(2):594-599.

160. Szende, A., Svensson, K., Stahl, E., Meszaros, A., Berta, GY. Psychometric and utility-based measures of health status of asthmatic patients with different disease control level. *PharmacoEconomics* 2004; 22(8):537-547.

161. Akinbami, L., Moorman, J. Asthma Prevalence, Health Care Use, and Mortality: United States, 2005-2009. National Health Statistics Report. No.32. 2011.

162. Department of Health. NHS Reference Costs 2011-12. DH:London. 2012.

163. Kind, P., Dolan, P., Gudex, C., Williams, A. Variations in population health status: results from a United Kingdom national questionnaire survey. *BMJ (Clinical Research Ed)* 1998; 316(7133):736-741.

164. Piacentini, G.L., Peroni, D.G., Bonafiglia, E., Chinellato, I., Bodini, A., Boner, A.L. Childhood asthma control test (C-ACT) and nasal eosinophil inflammation in asthmatic children. *Allergy: European Journal of Allergy and Clinical Immunology* 2010; 65(6):796-797.

165. Andersson, F., Borg, S., Barnes, P.J., Lofdahl, C.G., O'Byrne, P., Postma, D. et al. A high dose of inhaled budesonide is more cost-effective than a low dose in moderate asthma. *Journal of Allergy and Clinical Immunology* 2000; 105(1):S102-S103.

166. Curtis, L. Unit costs of health and social care. 2012. Personal Social Services Research Unit. PSSRU:Kent.

167. Main, C., Shepherd, J., Anderson, R., Rogers, G., Thompson Coon, J., Liu, Z. et al. Inhaled corticosteroids and long-acting beta2-agonists for the treatment of chronic asthma in children under the age of 12 years: Systematic review and economic analysis. Final report to NICE. 2006.

168. McTaggart-Cowan, H., Marra, C., Yang, Y., Brazier, J., Kopec, J., FitzGerald, J. et al. The validity of generic and condition-specific preference-based instruments: the ability to discriminate asthma control status. *Quality of Life Research* 2008; 17(3):453-462.

169. Sullivan, P., Slejko, J., Sculpher, M.J., Ghushchyan, V. Catalogue of EQ-5D scores for the United Kingdom. *Medical Decision Making* 2011; 31(6):800-804.

170. Ara, R., Brazier, J.E. Populating an Economic Model with Health State Utility Values: Moving toward Better Practice. *Value in Health* 2010; 13(5):509-518.

171. Lloyd, A., Price, D., Brown, R. The impact of asthma exacerbations on health-related quality of life in moderate to severe asthma patients in the UK. *Prim Care Respir J* 2007; 16(1):22-27.

172. BMJ Publishing Group Ltd. RCPCH Publications Ltd and the Royal Pharmaceutical Society of Great Britain 2013. *British National Formulary* 2013; Available from http://www.bnf.org/

173. Davis, S. A review of the psychometric performance of the EQ-5D in patients with asthma. Report prepared for NICE by the Decision Support Unit. 2010.

174. Office for National Statistics. Interim life tables 2009-2011. 2011.

175. Petsky, H.L., Cates, C.J., Li, A., Kynaston, J.A., Turner, C., Chang, A.B. Tailored interventions based on exhaled nitric oxide versus clinical symptoms for asthma in children and adults. [Review] [34 refs][Update of Cochrane Database Syst Rev. 2008;(2):CD006340; PMID: 18425949]. *Cochrane Database of Systematic Reviews.(4):CD006340, 2009.* 2009; [4], CD006340.

176. Fritsch, M., Uxa, S., Horak, F., Jr., Putschoegl, B., Dehlink, E., Szepfalusi, Z. et al. Exhaled nitric oxide in the management of childhood asthma: a prospective 6-months study. *Pediatr Pulmonol* 2006; 41(9):855-862.

177. de Jongste, J.C., Carraro, S., Hop, W.C., CHARISM Study Group, Baraldi, E., De Jongste, J.C. et al. Daily telemonitoring of exhaled nitric oxide and symptoms in the treatment of childhood asthma. *American Journal of Respiratory & Critical Care Medicine* 2009; 179(2):93-97.

178. Aerocrine. "Meta Analysis of Asthma Exacerbation Rates with FeNO Guided Asthma Management" Aerocrine submission to NICE. 2013.

179. Tan, K.S., Thomson, N.C. Asthma in pregnancy. *American Journal of Medicine* 2000; 109(9):727-733.

180. Moher, D., Liberati, A., Tetzlaff, J., Altman, D. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 2009; 6(6):e1000097.

10. APPENDICES

Appendix 1: Study Protocol

Measurement of exhaled nitric oxide concentration in asthma: NIOX MINO and NObreath

FINAL PROTOCOL

HTA Reference No. 1260

5 March 2013

Project Lead: Jon Minton

Health Economics and Decision Science

School of Health and Related Research (ScHARR)

The University of Sheffield

Regent Court

30 Regent Street

Sheffield S1 4DA

Tel: (+44) (0)114 222 0836

Fax: (+44) (0)114 272 4095

E-mail: j.minton@sheffield.ac.uk

Address for correspondence

Major documentation should be sent to:

1 the project lead (Jon Minton, j.minton@sheffield.ac.uk);

2 the project administrator (Andrea Shippam, a.shippam@sheffield.ac.uk);

3 and the managing director of ScHARR-TAG (Eva Kaltenthaler, e.kaltenthaler@sheffield.ac.uk).

Abbreviations

BTS	British Thoracic Society
CEAC	Cost effectiveness acceptability curve
CEAF	Cost effectiveness acceptability frontier
CRD	Centre for Reviews and Dissemination
DAR	Diagnostic Assessment Group
EAG	Evidence Assessment Group
FeNO	Exhaled Nitric Oxide (Assumed synonymous with FeNO)
FeNO	Fractional Exhaled Nitric Oxide (Assumed synonymous with FeNO)
FEV_1	Forced expiry volume in first second
FN	False Negative
FP	False Positive
HRQoL	Health related quality of life
IAD	Inflammatory Airway Disease
ICER	Incremental cost effectiveness ratio
ICS	Inhaled corticosteroids
LRTS	Lower respiratory tract symptoms
LTRA	leukotriene receptor antagonist
mL/s	millilitres per second
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
OCS	Oral corticosteroids
PEF	Peak expiratory flow
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
QALY	Quality adjusted life year
SCM	Specialist committee members
SIGN	Scottish Intercollegiate Guidelines Network
TN	True Negative
TP	True Positive
UK	United Kingdom

Contents

1. Plain English Summary

Nitric oxide monitors measure fraction of exhaled (FeNO) nitric oxide levels. Two hand-held nitric oxide monitors have been identified: NIOX MINO, developed by Aerocrine Inc; and NObreath, developed by Bedfont Scientific Ltd. High FeNO levels in people with symptoms suggestive of asthma, such as coughing and wheezing, may suggest that the patient has a type of asthma that could be treated with inhaled corticosteroids (ICS). In someone already diagnosed with asthma, changes in FeNO levels can indicate how well a patient is responding to ICS-based medication, whether medication is being adhered to, and whether the dosage of medication should be increased or decreased (titrated). [1]

The main aim of this assessment is to consider the role of these FeNO monitors in (1) the diagnosis of asthma in patients with suspected asthma, and (2) the management and monitoring of patients diagnosed with asthma. The decision problem will be considered separately for adults and children, as the diagnosis and treatment for these patient groups are slightly different. As diagnoses of asthma are routinely made in primary care without extensive testing, the emphasis within the diagnostic part of the decision problem will be on patients whose symptoms are difficult to diagnose.

Systematic reviews will be conducted to find evidence for the diagnostic and management applications of this these devices. The evidence produced by these reviews will be combined with other sources of evidence to construct an economic model. This model will be used to examine the expected costs and health outcomes associated with the use of these devices in diagnostic and management settings.

2. Decision problem

2.1 Purpose of the decision to be made

The objective of the evaluation is to assess the clinical and cost-effectiveness of FeNO measurement in people with asthma. This can be separated into two distinct questions:

a) What is the clinical and cost-effectiveness of the nitric oxide monitors included in this evaluation for use in the diagnosis of asthma in adults and children?
b) What is the clinical and cost-effectiveness of the nitric oxide monitors included in this evaluation for use in the management and monitoring of asthma in adults and children?

As the cut-off values used in diagnostic technologies affect their sensitivity (true positive rate) and specificity (true negative rate) it is also important to determine if an ‘optimal’ cut-off value can be identified for the use of NIOX MINO/NObreath for either diagnosis or

management purposes. Any exploration of test-specific cut-offs will be consistent with the CE-mark instructions of the interventions included in the assessment.

Three key questions emerged from the scoping workshops held in February 2013. One question relates to the diagnostic application of the devices; the other two relate to the management applications:

- **Diagnosis question**: Does FeNO concentration help to identify individuals who are most likely to respond to corticosteroid therapy?
- **Management question 1**: Does FeNO concentration help to optimise (i.e. lead to appropriate increases or decreases in) corticosteroid therapy doses during patient management? In particular, can exhaled nitric oxide concentration be used to safely reduce the dose of corticosteroid therapy when appropriate?
- **Management question 2**: Does FeNO concentration help to identify individuals who are not complying with corticosteroid therapy and can compliance be improved?

2.2 Clear definition of the intervention

2.2.1 NIOX MINO

NIOX MINO determines FeNO concentration in a breath sample. The device is small, hand-held and portable, and it can be used by both adults and children. It requires a 10 second exhalation of breath by the patient, at an exhalation pressure of 10 - 20 cm H_2O to maintain a fixed flow rate of 50±5 mL/s. The last 3 seconds of the 10 second exhalation is analysed by a calibrated electrochemical sensor, to give a definitive result in parts per billion. Clinical cut-off values can be applied to the FeNO values to categorise readings as low, intermediate or high according to the reference ranges for age less than 12 years and 12 years or more (Aerocrine. 'Guide to Interpretation of eNO Values').

NIOX MINO is pre-calibrated and designed to ensure a service- and calibration-free system. It can be used as a stand-alone device or connected to a PC for monitoring with the NIOX MINO Data Management Program and for use with Electronic Medical Record systems.

NIOX MINO is CE-marked and was launched in the UK in November 2004. It is currently available in 8 GP surgeries and used in more than 90 hospitals across the UK.

The manufacturer claims that NIOX MINO is indicated for use as follows:

- To diagnose the specific type of airway inflammation to guide treatment
- To predict the onset of asthma symptoms or loss of asthma controls due to eosinophilic airway inflammation
- To monitor compliance to corticosteroid therapy and effectiveness of treatment (frequency of exacerbations).

2.2.2 NObreath

NObreath (Bedfont Scientific Ltd.) is a diagnostic monitoring device that measures FeNO produced by airway inflammation. The reading is presented in parts per billion and is claimed to be directly related to the severity of inflammatory disease (for example, asthma). NObreath requires 12 seconds of exhalation of breath in adults and 10 seconds in children.
NObreath weighs approximately 400g (including batteries). It has a battery life that lasts up to 120 tests. The device is CE marked.

2.3 Populations and relevant subgroups

2.3.1 Diagnosis

The population of interest is people with clinical characteristics suggestive of asthma.
Subgroups:

- Scoping workshop attendees considered that exhaled nitric oxide measurement had the greatest potential to benefit people who are difficult to diagnose.
- Certain groups of patients may experience different outcomes from the use of FeNO when compared to the main population under assessment (for example, FeNO levels tend to be lower in smokers than non-smokers). Such groups should be assessed separately if evidence allows.

2.3.2 Management

The population of interest is patients diagnosed with asthma. There are two subgroups of particular interest:

- Those with good asthma control who are being considered for a dose reduction
- Those with uncontrolled asthma who are experiencing exacerbations or worsening of symptoms, and are being considered for a dose increase of ICS or are being checked for compliance to treatment.

Certain groups of patients may experience different outcomes from the use of FeNO when compared to the main population under assessment (for example, FeNO levels tend to be lower in smokers than non-smokers). Such groups should be assessed separately if evidence allows.

2.4 Place of the intervention in the diagnostic/treatment pathway(s)

Scoping workshop attendees considered that the interventions should be assessed when added to current practice.

2.4.1 Diagnosis

Asthma is diagnosed on the basis of symptoms and objective tests of lung function (such as peak expiratory flow [PEF] rate and forced expiratory volume in the first second [FEV_1]) and percentage predicted FEV_1 (calculated as a percentage of the predicted FEV_1 for a person of the same height, sex and age without diagnosed asthma). Variability of PEF and FEV_1, either spontaneously or in response to therapy, is a characteristic feature of asthma. The severity of asthma is judged according to symptoms and the amount of medication required to control the symptoms, and is based on British Guidelines for the Management of Asthma from the British Thoracic Society (BTS) and Scottish Intercollegiate Guidelines Network (SIGN).[2]

The diagnosis of asthma is a clinical one and there is no standardised definition of the condition. Central to all definitions is the presence of symptoms (wheezing, breathlessness, chest tightness, and cough) and of variable airflow obstruction. More recently descriptions of asthma have included airway hyper responsiveness and airway inflammation. It is unclear how these features relate to each other, how they are best measured and how they contribute to the clinical manifestations of asthma.

2.4.1.1 Diagnosis of asthma in children

2.4.1.1.1 Current pathway

A flow chart of the diagnostic pathway for asthma in children is given in Figure 1 with locations for the use of FeNO measurements. Diagnosis in children is clinically-based on recognising a characteristic pattern of episodic symptoms in the absence of an alternative explanation. Clinical features that increase the probability of asthma include:

- More than one of the following symptoms - wheeze, cough, difficulty breathing, chest tightness - particularly if these are frequent and recurrent; are worse at night and in the early morning; occur in response to, or are worse after, exercise or other triggers, such as exposure to pets; cold or damp air, or with emotions or laughter; or occur apart from colds
- Personal history of atopic disorder
- Family history of atopic disorder and/or asthma
- Widespread wheeze heard on auscultation
- History of improvement in symptoms or lung function in response to adequate therapy.

If asthma is suspected, an initial clinical assessment should be carried out to estimate the probability of asthma. According to the British Guidelines on the Management of Asthma,[2] based on initial clinical assessment, an individual child can be classed into one of three groups:

- High probability – diagnosis of asthma likely
- Low probability – diagnosis other than asthma likely
- Intermediate probability – diagnosis uncertain

For children identified as having a low probability of asthma, a more detailed investigation and specialist referral should be considered. For children with a high probability of asthma, a trial of treatment should be started immediately. The response to treatment should be reassessed every 6 months. Those with a poor response to treatment should undergo more detailed investigations.

According to the British Guidelines on the Management of Asthma, there is insufficient evidence at first consultation to make a firm diagnosis of asthma in some children, particularly those below the age of 4 to 5 years.[2] For these children who can perform spirometry and for whom airway obstruction is evident, change in forced expiratory flow volume or peak expiratory flow monitoring should be assessed in response to an inhaled bronchodilator and/or the response to a trial of treatment for a specified period.

In children with an intermediate probability of asthma who can perform spirometry and have no evidence of airway obstruction tests for atopic status, assessment of bronchodilator reversibility and if possible, bronchial hyper-responsiveness using methacholine, exercise or mannitol should be considered. These tests are performed in secondary care. In such cases, specialist referral should always be considered.

Other investigations to diagnose asthma in children include tests of eosinophilic airway inflammation using induced sputum or exhaled nitric oxide concentrations, tests of atophy by skin test or blood eosinophilia or by chest x-ray.

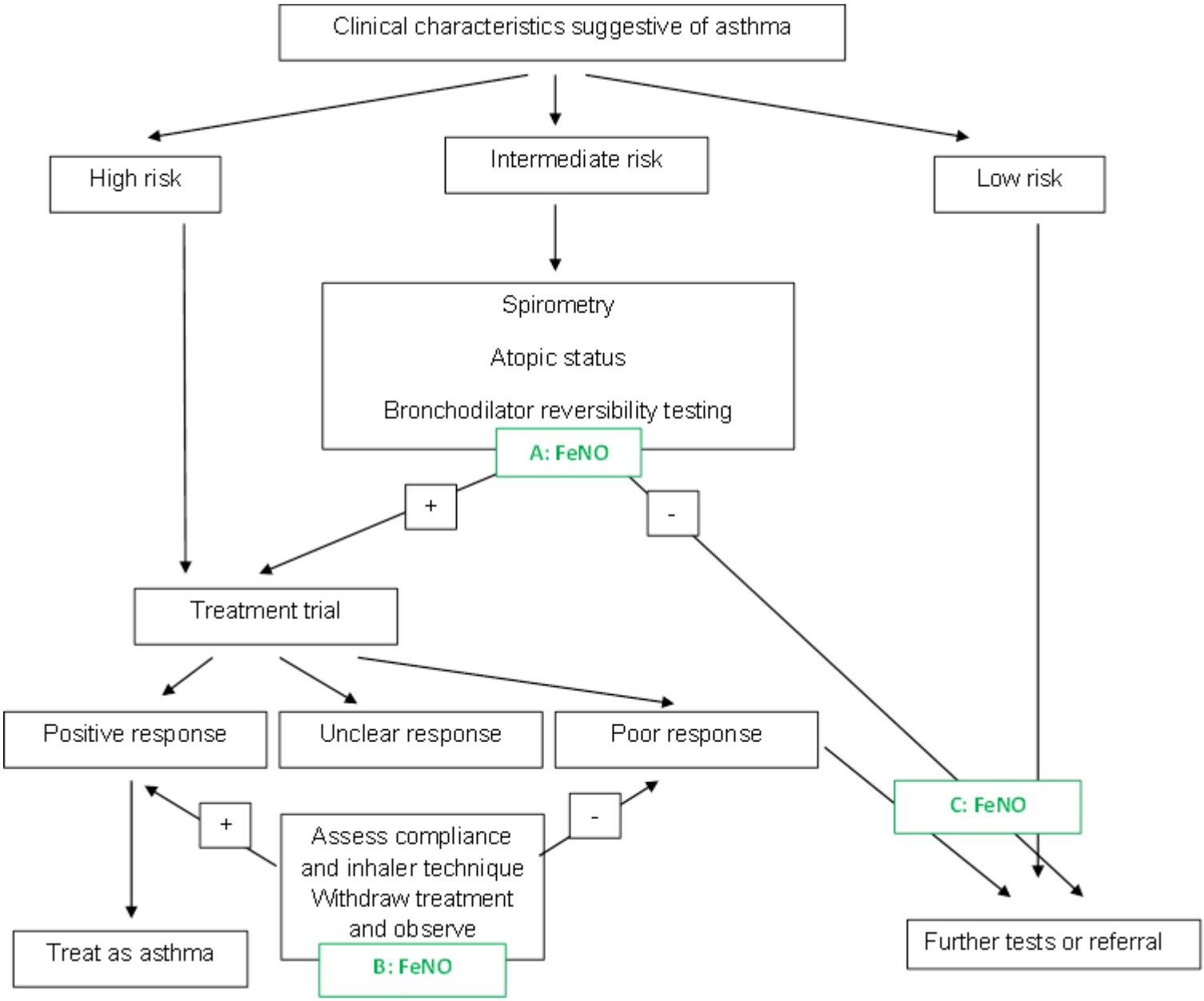
2.4.1.1.2 Position of FeNO in the pathway

FeNO is thought to be of most use in Positions A, B and C as shown in Figure 1. This equates to patients who are difficult to diagnose. In primary care, FeNO measurement may help to reduce the number of individuals who are subjected to a trial of treatment with inhaled corticosteroids and may reduce the number of referrals to secondary care. In secondary care, FeNO measurement may help to reduce the use of more expensive diagnostics (for example, tests of airway hyperresponsiveness) and reduce the number of individuals who are subjected to a trial of treatment with inhaled corticosteroids. During the scoping workshop, clinical specialists suggested that individuals who have had their FeNO measured in primary care, but have been referred to secondary care, will have their FeNO level measured again.

Position B can also be considered as a management strategy (see Section 2.4.2), but as patients undergoing a trial of treatment may not yet have been diagnosed as asthmatic, it may

be necessary to assess this use as both a diagnostic strategy and a management strategy. The availability of evidence may dictate how this position is assessed.

Figure 1. Flow chart for the diagnosis of asthma in children as described in the British Guidelines on the Management of Asthma,[2] with possible positions for the addition of FeNO testing.



2.4.1.2 Diagnosis of asthma in adults

2.4.1.2.1 Current pathway

A flow chart of the diagnostic pathway for asthma in adults is presented in Figure 2. Diagnosis in adults is also based on the clinical history and includes the recognition of a

characteristic pattern of symptoms and signs and the absence of an alternative explanation for them. Unlike in children, spirometry is tested initially to assess the presence and severity of airflow obstruction.

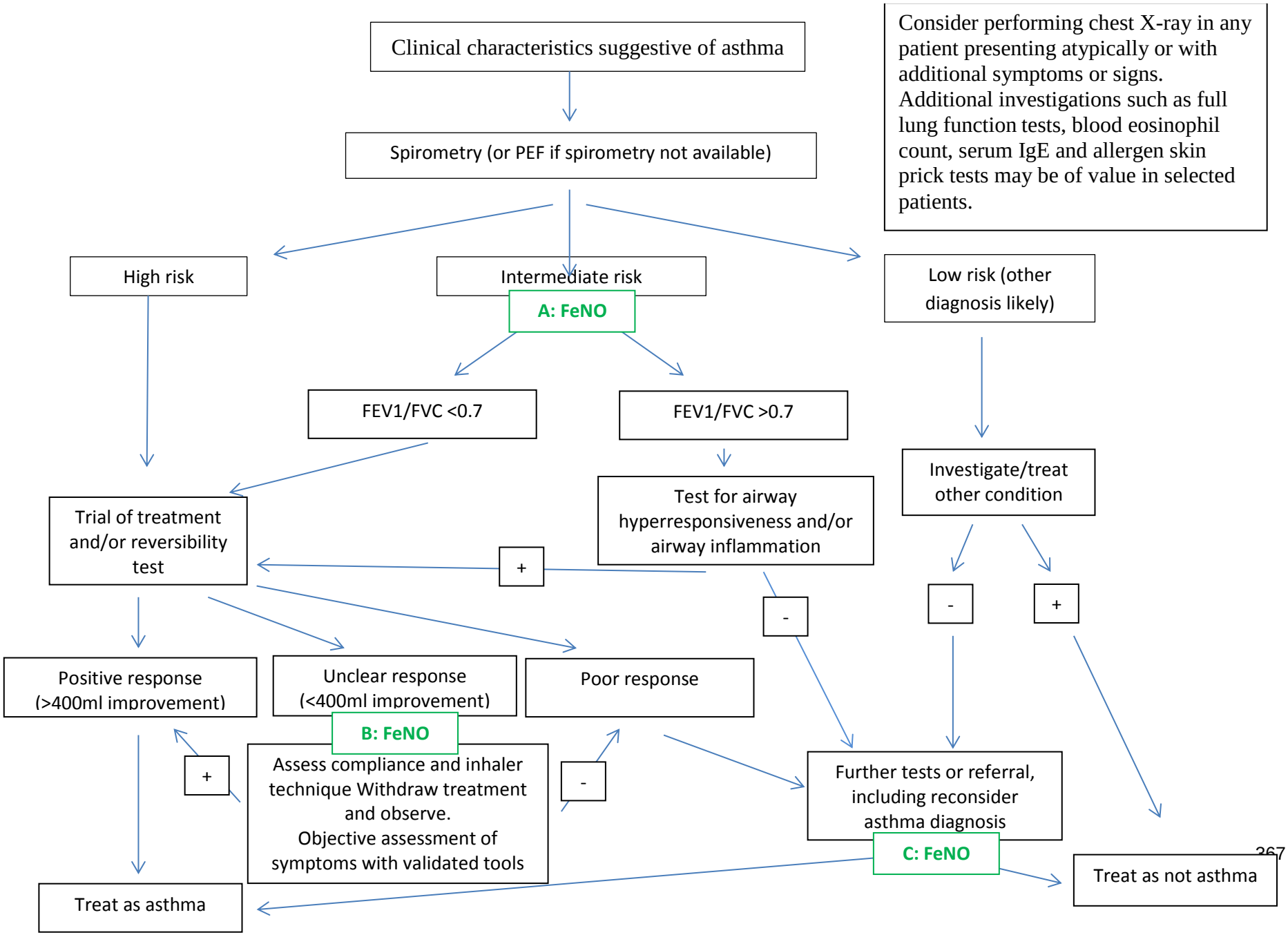
As in the diagnosis of children, adults are also classified as having a high, low or intermediate probability of asthma. Chest x-ray and specialist referral may be considered in any patient presenting atypically or with additional symptoms or signs.

2.4.1.2.2 Position of FeNO in the pathway

FeNO is thought to be of most use in Positions A, B and C as shown in Figure 2. This equates to patients who are difficult to diagnose. In primary care, FeNO measurement may help to reduce the number of individuals who are subjected to a trial of treatment with inhaled corticosteroids and may reduce the number of referrals to secondary care. In secondary care, FeNO measurement may help to reduce the use of more expensive diagnostics (for example, tests of airway hyperresponsiveness) and reduce the number of individuals who are subjected to a trial of treatment with inhaled corticosteroids. During the scoping workshop, clinical specialists suggested that individuals who have had their FeNO measured in primary care, but have been referred to secondary care, will have their FeNO level measured again.

Position B can also be considered as a management strategy (see Section 2.4.2), but as patients undergoing a trial of treatment may not yet be diagnosed as asthmatic, it may be necessary to assess this use as both a diagnostic strategy and a management strategy. The availability of evidence may dictate how this position is assessed.

Figure 2, Flow chart for the diagnosis of asthma in adults as described in the British Guidelines on the Management of Asthma,[1] with positions for the addition of ENO testing.



2.4.2 Monitoring and management

For both children and adults, asthma is monitored and managed in primary care by routine clinical reviews on at least an annual basis. These reviews include (but are not limited to) assessment of patient's symptom score (using a validated questionnaire), exacerbations, oral corticosteroid use, time off school or work, growth, inhaler technique and in adults, lung function assessed by spirometry of peak expiratory flow. Patients are managed in a stepwise manner, with escalation of medication until control is reached. This approach to pharmacological management for children and adults is represented in Tables 1 and 2 respectively (taken from the British guidelines).[2] Patients are started on the step that most closely matches the severity of their symptoms.

2.4.2.1 Monitoring asthma in children

The British Guideline on the Management of Asthma[2] states that asthma in children is best monitored in primary care by routine clinical review on at least an annual basis. The factors that should be monitored and recorded include:

- Symptom score, for instance Children's Asthma Control test or Asthma Control Questionnaire
- Exacerbations, oral corticosteroid use and time off school/nursery due to asthma since last assessment
- Inhaler technique
- Adherence to treatment, which can be assessed by reviewing prescription refill frequency
- Possession of and use of self management plan/personalised asthma action plan
- Exposure to tobacco smoke
- Growth (height and weight centile)

The guideline is indistinct about the use of biomarkers in monitoring asthma. It states: "a better understanding of the natural variability of biomarkers independent of asthma is required and studies are needed to establish whether subgroups of patients can be identified in which biomarker guided management is effective".

2.4.2.2 Monitoring asthma in adults

According to the guideline,[2] symptom-based monitoring is adequate in the majority of adults with asthma. Those with poor lung function and with a history of exacerbations in the previous year may be at a greater risk of future exacerbations for a given level of symptoms.

Asthma in adults is best monitored in primary care by routine clinical review on at least an annual basis. The factors that should be monitored and recorded include:

- Symptomatic asthma control: best assessed using directive questions such as the Asthma Control Questionnaire or Asthma Control Test
- Lung function, assessed by spirometry or by PEF
- Exacerbations, oral corticosteroid use and time off work or school since last assessment
- Inhaler technique
- Adherence to treatment, which can be assessed by reviewing prescription refill frequency
- Bronchodilator reliance, which can be assessed by prescription refill frequency
- Possession of and use of self management plan/personal action plan

2.4.2.3 Management in adults and children

Asthma management aims to control symptoms (including nocturnal symptoms and exercise-induced asthma), prevent exacerbations and achieve the best possible lung function, with minimal side effects of treatment. The British Guideline on the Management of Asthma recommends a stepwise approach to treatment in both adults and children. Treatment is started at the step most appropriate to the initial severity of the asthma, with the aim of achieving early control of symptoms and optimising respiratory function. Control is maintained by stepping up treatment as necessary and stepping down when control is good (see Tables 1 and 2).

2.4.2.4 Position of FeNO in the management pathway

Experts suggested that FeNO measurement may be helpful in individuals diagnosed with asthma to facilitate titration of corticosteroid therapy, to check for compliance with medication, and ultimately to lead to better asthma control. It is likely that management decisions would be based on a combination of the monitoring information collected at review and FeNO measurements. Possible contingencies based on this information will be elicited from clinicians during the course of the assessment.

Step 1: Mild intermittent asthma	**Step 2: Regular preventer therapy**	**Step 3: Initial add-on therapy**	**Step 4: Persistent poor control**	**Step 5: Continuous or frequent use of oral steroids**
Inhaled short-acting B_2-agonist Prescribe inhalers only after the patient has received training in the use of the device and has demonstrated satisfactory technique **0-5 years pMDI and spacer are preferred delivery system.**	Add inhaled corticosteroid (ICS) 200-400mcg/day (BDP or equivalent) Start dose of inhaled corticosteroid appropriate to severity of disease. 200mcg/day is an appropriate dose for most children **Special instructions for under 5 years** Use a leukotriene receptor antagonist (LTRA) if inhaled corticosteroid cannot be used	**Special instructions for under 5 years** In the under 5 years and those already taking inhaled corticosteroids consider adding LTRA. In those already taking LTRA consider adding ICS 200-400mcg/day (BDP or equivalent).	**Special instructions for under 5 years** Refer to paediatrician	
		Special instructions for 5-12 years Add inhaled long-acting B2-agonist (LABA) and assess response. **If response good** - continue. Consider combination inhalers in those for whom LABA are effective at controlling symptoms. **If response poor** discontinue and increase ICS to 400mcg/day (BDP or equivalent). **If response still poor**, add other therapies.	**Special instructions for 5-12 years** Increase inhaled corticosteroid up to 800mcg/day (BDP or equivalent) Consider referral to paediatrician	**Special instructions for 5-12 years** Use daily steroid tablet in lowest dose to provide adequate control Maintain high-dose ICS at 800mcg (BDP or equivalent) per day Refer to paediatrician

1. Table 1: Asthma in children. - Summary of stepwise management (from British Thoracic Society/SIGN Guidelines for the Management of Asthma).[2]

Step 1: Mild intermittent asthma	**Step 2: Regular preventer therapy**	**Step 3: Initial add-on therapy**	**Step 4: Persistent poor control**	**Step 5: Continuous or frequent use of oral steroids**
Inhaled short-acting B_2-antagonist Prescribe inhalers only after the patient has received training in the use of the device and has demonstrated satisfactory technique	Add inhaled corticosteroid (ICS) 200-800mcg/day (BDP or equivalent) Start dose of inhaled corticosteroid appropriate to severity of disease. 400mcg/day (BDP or equivalent) is an appropriate dose for most patients	1. Add inhaled long-acting B_2-agonist (LABA) and assess control of asthma: Good response to LABA Continue LABA Combination inhalers should be considered in those for whom LABA are effective at controlling symptoms. Benefit from LABA but control still inadequate Continue LABA and increase inhaled steroid dose to 800 mcg/day BDP or equivalent (if not already on this dose) No response to LABA Stop LABA and increase inhaled steroid to 800mcg/ day. BDP or equivalent 2. If control still inadequate, Institute trial of other therapies, leukotriene antagonist or SR theophylline receptor	Consider trials of: Increased dose of inhaled corticosteroid up to 2000mcg/day (BDP or equivalent) Consider adding a fourth drug eg leukotriene receptor antagonist, SR theophylline or B_2-agonist tablet	Use daily steroid tablet in lowest dose to provide adequate control Maintain high dose inhaled corticosteroids at 2000mcg/day (BDP or equivalent) Consider other treatments to minimise the use of oral steroids Refer patient for specialist care
Regular review of patients as treatment is stepped down is important. Patients should be maintained at the lowest possible dose of inhaled corticosteroid. Any reduction in inhaled steroids should be undertaken slowly, every three months, as patients deteriorate at different rates. Inhaled corticosteroid reduction in severe asthma should be reduced by 25% only, 50% for more stable patients				
In selected patients at Step 3 who are poorly controlled, or in selected patients at step 2 who are poorly controlled, the use of budesonide/formoterol in a single inhaler as rescue medication and maintenance therapy can be an effective treatment option.				

2. Table 2: Asthma in adults. - Summary of stepwise management (from British Thoracic Society/SIGN Guidelines for the Management of Asthma).[2]

2.5 Relevant comparators

The relevant comparators are diagnosis or management according to the current UK guideline,[2] as described in Sections 2.4.1.2.1 and 2.4.1.1.1.

2.5.1 Diagnosis

This comprises following the established diagnostic pathway without the use of FeNO measurements (See Figures 1 & 2 for children and adults respectively).

2.5.2 Management

This comprises following the established diagnostic pathway without the use of FeNO measurements (See Tables 1 & 2 for children and adults respectively).

2.5.3 Healthcare setting

Primary care and secondary care.

2.5.4 Outcomes

2.5.4.1 Clinical considerations

The intermediate measures for consideration include:

- Diagnostic test accuracy
- Test failure rate

The clinical outcomes for consideration include:

- Asthma control which includes asthma symptoms
- Exacerbation rate. Including frequency of exacerbations requiring unscheduled contact with healthcare professionals, visit to accident and emergency departments or hospitalisations.
- Clinical complications associated with acute exacerbations
- Levels of inhaled corticosteroids
- Use of oral corticosteroids
- Adverse effects of treatments (including bronchodilators and steroids)
- Health-related quality of life
- Mortality

2.5.4.1 Cost considerations

- Cost of equipment, reagents and consumables
- Maintenance and renewal of equipment
- Cost associated with acute exacerbations
- Cost of further investigations avoided

3 Methods for assessing the outcomes arising from the use of the interventions

A systematic review will be conducted to identify evidence relevant to the decision problem. There will be two main reviews, each reporting results for adults and children separately. The two main reviews are:

1. A review of diagnostic accuracy
2. A review of the efficacy of monitoring and management strategies

The inclusion and exclusion criteria are detailed for each review separately here.

3.1 Diagnostic accuracy review

3.1.1 Population

The primary population is patients presenting with clinical characteristics suggestive of asthma. The main relevant subgroups within this population are:

- Those presenting with clinical characteristics suggestive of asthma and who are difficult to diagnose. This patient group roughly equates to the "intermediate" group in the patient pathway (Position A in Figures 1 and 2), and those in Positions D and E
- Women during pregnancy
- Older people
- Smokers (whose FeNO levels may be affected by smoking)

Studies will be included if they recruited a wider population but report *a priori* subgroup analyses for the populations of relevance to this review.

3.1.2 Interventions

Studies will be included if they report results relating to the clinical validity or clinical utility of NIOX MINO or NObreath.

If data are not available for the clinical validity and clinical utility of the interventions, studies will be included if they report clinical validity or clinical utility of FeNO measured by

chemiluminescence (the gold standard in FeNO measurement) and supporting evidence of the analytic validity of NIOX MINO and NObreath compared to chemiluminescence will be presented. A full systematic review of analytic validity may not be necessary if a recent (2010 onwards), good quality systematic review of this evidence is found. Otherwise, a full systematic review of analytic validity of NIOX MINO and NObreath will be conducted.

Studies will be included if they utilise the standard cut-off rates defined in the ATS guidelines,[3] and as recommended by the manufacturers of NIOX MINO[4] (Bedfont do not provide an interpretation guide). These are as follows:

- FeNO less than 25 ppb (< 20 ppb in children) indicates that eosinophilic inflammation and responsiveness to corticosteroids are less likely
- FeNO greater than 50 ppb (>35 ppb in children) indicates that eosinophilic inflammation and, in symptomatic patients, responsiveness to corticosteroids are likely
- FeNO values between 25 ppb and 50 ppb (20–35 ppb in children) should be interpreted cautiously and with reference to the clinical context

Where evidence is available, and subject to the approval of NICE, studies using other cut-off points will also be included to allow modelling of the optimum values for cost effectiveness.

Expiratory flow rate and exhalation time are important factors that can affect measurements. The standard methodology defined in the ATS/BTS guidelines recommends

- Expiratory flow rate of 50mL per second, though other flow rates can be used in certain situations if desired or to measure flow-independent parameters.
- An exhalation time of 10 seconds may be necessary to establish a stable plateau in FeNO for evaluation over a 3 second window.[5]

These guidelines will be interpreted, based on clinical advice given at the scoping workshop and the design specifications of the two interventions, as:

- Expiratory flow rate of 50mL per second (0.05L/sec)
- An exhalation time of at least 10 seconds.

Studies of FeNO measurement by NIOX MINO, NObreath or chemiluminescence will only be included if they comply with these parameters.

3.1.3 Comparators

Studies of clinical validity and utility will be included if the comparator for the diagnosis of asthma comprises any combination of the tests and clinical characteristics described in the UK guideline. [2]

Studies will be excluded if the comparator uses tests to diagnose asthma that are not used as standard in the UK or if the comparator includes the use of FeNO measurement. If studies are identified which compare different locations of FeNO testing head-to-head, within a diagnostic pathway comparable to clinical practice in the UK, and these do not include a comparator arm without FeNO, these could be presented separately as additional evidence for consideration by the assessment committee.

Studies of analytic validity will be included if they compare the intervention devices to FeNO values measured by chemiluminescense at the flow rate and exhalation time stated for the interventions. If no studies at this flow rate and exhalation time are found, any flow rate or exhalation time will be included.

3.1.4 Outcomes

Studies of clinical utility will be included if they report any of the following outcomes at any time point:

- Incidence of acute exacerbations, including those requiring unscheduled contact with healthcare professionals, visits to accident and emergency departments or hospitalisations. As patients can experience more than one exacerbation within the timeframe of follow-up, the rate of exacerbations is the preferred outcome measurement. Other measures (time-to-event data; numbers of patients experiencing an exacerbation) will only be considered if insufficient data are available for the rate of exacerbations. Any definition of exacerbation will be acceptable.
- Asthma control which includes asthma symptoms, either reported individually or by use of a standardised patient outcome measure or symptom score.
- Clinical complications associated with acute exacerbations
- Levels of inhaled corticosteroids
- Use of oral corticosteroids
- Adverse effects of treatments (including bronchodilators and steroids)
- Health-related quality of life
- Mortality
- Test failure rate

Studies of clinical validity will be included if they report data that allow the extraction of the numbers of patients who are true positive, true negative, false positive and false negative against the reference standard. Studies which report test failure rates will also be included.

Studies of analytic validity will be included if they report the ability of the test to measure FeNO accurately, as compared to chemiluminescence in humans.

3.1.5 Study design

There are three types of evidence that may inform reviews of diagnostic accuracy: clinical utility, clinical validity and analytic validity.[6] This review will include the highest level of evidence, namely clinical utility studies, which follow patients from diagnostic test to clinical outcomes (also known as end-to-end studies and which demonstrate the ability of the test to improve patient outcomes). If no evidence is found at this level, clinical validity studies (which compare the diagnosis of patients by the intervention with a reference standard diagnostic; this is influenced by both calibration of the test and its ability to differentiate between patients with and without disease) will be included. If there is no evidence at this level, studies of analytic validity linked to studies of the clinical utility (or if no utility studies are available, clinical validity) of FeNO will be included. The inclusion criteria for each level are given below.

For the review of clinical utility, RCTs will be included where available. If sufficient high quality evidence is not available from RCTs, the next best level of evidence will be included, according to the well-established hierarchy of evidence.[7] It may be preferable, however, to draw conclusions from well-designed studies of clinical validity rather than poor quality studies of clinical utility.

For the review of clinical validity, studies are likely to be prospective cohort studies, cross sectional studies or retrospective cohort studies. If studies of these designs are not located, other study designs will be considered (e.g. case control studies). Both studies deriving cut-off values for diagnosis and studies validating existing cut-off values for diagnosis will be included.

For the review of analytic validity, the most recent or comprehensive good quality systematic review will be included where available. If this has been conducted recently (2010 onwards), no further evidence will be included. However, if no good quality recent reviews are identified, studies of analytic validity will be included if they have been conducted in humans. Studies performed *in vitro* on gas samples will not be included unless no test evidence is found in humans. Studies of inter-rater reliability or inter-subject repeatability will be excluded.

In all three reviews, studies with the following characteristics will be excluded:

- studies not meeting the inclusion criteria

- animal models
- preclinical and biological studies
- editorials and opinion pieces
- studies only published in languages other than English
- reports published as meeting abstracts will only be included where comparable data do not exist in full published studies and where sufficient methodological details are reported to allow critical appraisal of study quality.

3.2 Management review

An existing Cochrane review [8] will be updated with searches from 2009. Additional data for exacerbation rates will be included as reported in an update and reanalysis of the same review provided by Aerorcrine.[9]

3.2.1 Population

The population of interest is patients diagnosed with asthma. There are two subgroups of particular interest:

- Those with good asthma control who are being considered for a dose reduction
- Those with uncontrolled asthma who are experiencing exacerbations or worsening of symptoms, and are being considered for a dose increase of ICS or are being checked for compliance to treatment.

And three further subgroups within each of these categories:

- Women during pregnancy
- Older people
- Smokers (whose FeNO levels may be affected by smoking)

Studies will be included if they recruit whole asthma populations or if they recruit patients exclusively from any of the subgroups.

3.2.2 Interventions

Studies using NIOX MINO or NObreath will be included if they comply with the flow rate specifications listed in Section 3.1.2.

If studies using NIOX MINO or NObreath are not located, studies using chemiluminescence to measure FeNO at the flow rate and specification listed in Section 3.1.2 will be included.

Only studies using FeNO measurements in:

- routine annual monitoring
- dose titration indicated during routine monitoring
- assessment of compliance

will be included in the review. Studies where FeNO is measured on a regular basis (ie not during routine annual review), with the intention of predicting exacerbations or loss of control, will be excluded.

Any protocols and cut-off values for management decisions or compliance monitoring will be included.

3.2.3 Comparators

Studies comparing the interventions to any other management strategy that does not utilise FeNO measurements will be included. Studies using management strategies that closely match all or part of UK practice as described in the UK guidelines[2] will be included. If no studies which closely match UK practice are found, studies using other management strategies will be included.

3.2.4 Study design

Randomised controlled trials will be included. If insufficient RCT evidence is identified, other study designs will be included according to the hierarchy of evidence for efficacy trials.[7]

3.3 Search strategy

The search strategy will comprise the following main elements:

- Searching of electronic databases, clinical trial registers and websites
- Reference tracking of retrieved papers
- Citation searching
- Contact with experts in the field

The electronic databases and websites to be searched will include the following:

- MEDLINE and Medline in Process
- EMBASE
- The Cochrane Library (including Cochrane Database of Systematic Reviews, and the Cochrane Controlled Trials Register)
- Web of Knowledge Science Citation Index Expanded
- Conference Proceedings Citation Index – Science
- Clinicaltrials.gov
- metaRegister of Controlled Trials

- Manufacturer and User Facility Device (MAUDE)
- EuroScan International Network

A comprehensive Medline search strategy for the diagnostic review is provided in Appendix A. During the course of the review, two iterative approaches will be considered upon further discussion with the DAR review team:

- Application of study design filters (RCTs, SRs, diagnostic) to the strategy (Appendix A) to identify an initial collection of relevant papers which will then be used to inform smaller searches for similar papers.
- Key papers provided by NICE and the manufacturers and those identified by the review team will be used to inform and design iterative searches for similar papers that were not retrieved in the search above.

No language restrictions will be applied.

3.4 Data extraction strategy

Data will be extracted by one reviewer using a standardised data extraction form and checked by another. Any discrepancies will be resolved by discussion, with involvement of a third reviewer when necessary. Data will be extracted from the existing Cochrane review[8] and manufacturer's meta-analysis [9] where possible. Any discrepancies between sources will be checked against the original journal articles. If time allows, attempts will be made to contact authors for any missing data. Data from multiple publications of the same study will be extracted and quality assessed as a single study.

3.5 Quality assessment strategy

This review is likely to draw on evidence provided by several different study designs. Each study design will be assessed according to the principles outlined in the CRD handbook and the Cochrane handbook.[10,11] RCT studies will be assessed according to the Cochrane risk of bias tool, with additional questions taken from the CRD guidelines if relevant. Diagnostic accuracy studies will be assessed using QUADAS II.[12] Other study designs will be assessed using tools or adapted tools specific to the study design.

Quality assessment will be conducted by one reviewer and checked by a second. A third reviewer will be consulted in cases of disagreement.

3.6 Methods of analysis/synthesis

Studies will be tabulated and discussed in a narrative review in the following groups and population subgroups:

Diagnostic review

- Adults with clinical characteristics of asthma
    - Patients who are difficult to diagnose
    - Pregnant women
    - Older age adults
    - Smokers
- Children with clinical signs of asthma
    - Patients who are difficult to diagnose

Management review

- Adults with clinical signs of asthma
    - Pregnant women
    - Older age adults
    - Smokers
- Children with clinical signs of asthma

If sufficient clinically and statistically homogenous data are available, data will be pooled using appropriate meta-analytic techniques to estimate a summary measure of effect on relevant outcomes. The following subgroup analyses will be considered:

- Studies with similar comparators; studies conducted in the UK or with a comparator with a high degree of similarity to UK guidelines in terms of diagnostic pathway and management pathway
- Studies using different cut of values for interpretation
- Location of care (primary or secondary care)

Clinical, methodological and statistical heterogeneity will be investigated in sensitivity analyses. Sources of heterogeneity that may be investigated include (list not exhaustive)

- Duration of study (treatment and measurement of outcomes)
- Definition of asthma
- Definition of exacerbation
- Study quality

4. Methods for synthesising evidence of cost effectiveness

4.1 Identifying and systematically reviewing published cost-effectiveness studies

A systematic review of existing models of NIOX MINO/NObreath in either diagnosis or management of asthma will be conducted. Because the estimated clinical and cost-effectiveness of FeNO monitors depends on patient management pathways, it is important that existing models of patient management for patients with asthma are reviewed. Due to finite resources, a full systematic review of this broader asthma will not be conducted, although a thorough search of existing asthma management models will be conducted.

The primary outcome from the model will be an estimate of the incremental cost per additional quality-adjusted life year (QALY) gained associated with the use of the devices as compared against current UK guidelines. Secondary outcomes (the health benefits listed above) will also be presented. The economic analysis will adopt an NHS and Personal Social Services (PSS) perspective. Costs and health outcomes will be discounted at an annual rate of 3.5%. Modelling assumptions will be drawn from expert clinical expert opinion where required. Health-related quality of life (HRQoL) data will be reviewed and used to generate the quality adjustment weights required to estimate QALYs. Costs will be derived from national sources (e.g. NHS reference costs, national unit costs, British National Formulary)[13,14] and data provided by the manufacturers.

4.2 Evaluation of costs and cost effectiveness

The economic analysis will follow the NICE Reference Case. [6]

The costs of the treatment will include both the costs of performing the diagnostic tests, and downstream costs which result from using the information made available by the test. These downstream costs could be much larger than the test costs. Correctly estimating these downstream costs requires that the long-term consequences of a diagnostic test result are known. Sufficiently realistic modelling of the patient management pathways of those patients categorised as ICS responsive or ICS non-responsive will thus be required.

4.3 Development of a health economic model

If a review of existing economic models, does not identify a suitable model, a *de novo* model will be developed. This model will be based around the UK asthma diagnostic and care management pathway. The model will involve two components: a diagnostic component, and a management component. Both components could involve the use of NIOX MINO/NObreath. This means it is

expected there will be at least three interventions to compare against the reference case of routine diagnosis and management:

- **Reference case**: NIOX MINO/NObreath used for neither diagnosis nor management;
- **Comparator 1**: NIOX MINO/NObreath used for diagnosis but not for management;
- **Comparator 2**: NIOX MINO/NObreath used for management but not diagnosis;
- **Comparator 3**: NIOX MINO/NObreath used for both management and diagnosis.

A model has already been identified, but if it is not suitable the development of a *de novo* model is likely to be an iterative process. [15] A conceptual model will be developed in conjunction with clinical experts to capture the current pathway of care for the diagnosis and management asthma and how the new tests would change the pathway if routinely available in the NHS. The conceptual model will indicate the data requirements which will be sought both from the published literature and within commercial-in-confidence data held by the manufacturers. The model is likely to evolve following discussions with project stakeholders and the specialist committee members (SCMs), and according to the availability of evidence. It is anticipated that there may be limited evidence for some of the parameters that will be included in the economic model. Therefore, the uncertainty around the parameter estimates will be modelled to take this into account. A range of scenarios will be presented varying main model assumptions to identify parameters that impact the most the ICER and to represent the uncertainty in parameter estimates. Furthermore, probabilistic sensitivity analysis (PSA) will also be undertaken using standard Monte Carlo simulation methods. The uncertainty in each parameter will be characterised using a probability distribution. The decision uncertainty will be presented as the probability that each intervention is the most cost-effective for a given cost-effectiveness threshold. Decision uncertainty will be represented using cost-effectiveness acceptability curves (CEACs) and cost-effectiveness acceptability frontiers (CEAFs).

The methods and approaches described in the NICE Decision Support Unit Technical Support Document 13 will be used in structuring and parameterising the model. [16]

The next two subsections of this report will provide illustrative conceptual models of the short-term and long-term components of the economic models. However, the final model or models used may differ substantially from these.

4.3.1 Illustrative short-term diagnostic model

An illustrative decision tree describing the diagnostic component of the economic model is shown in Figure 1 below.

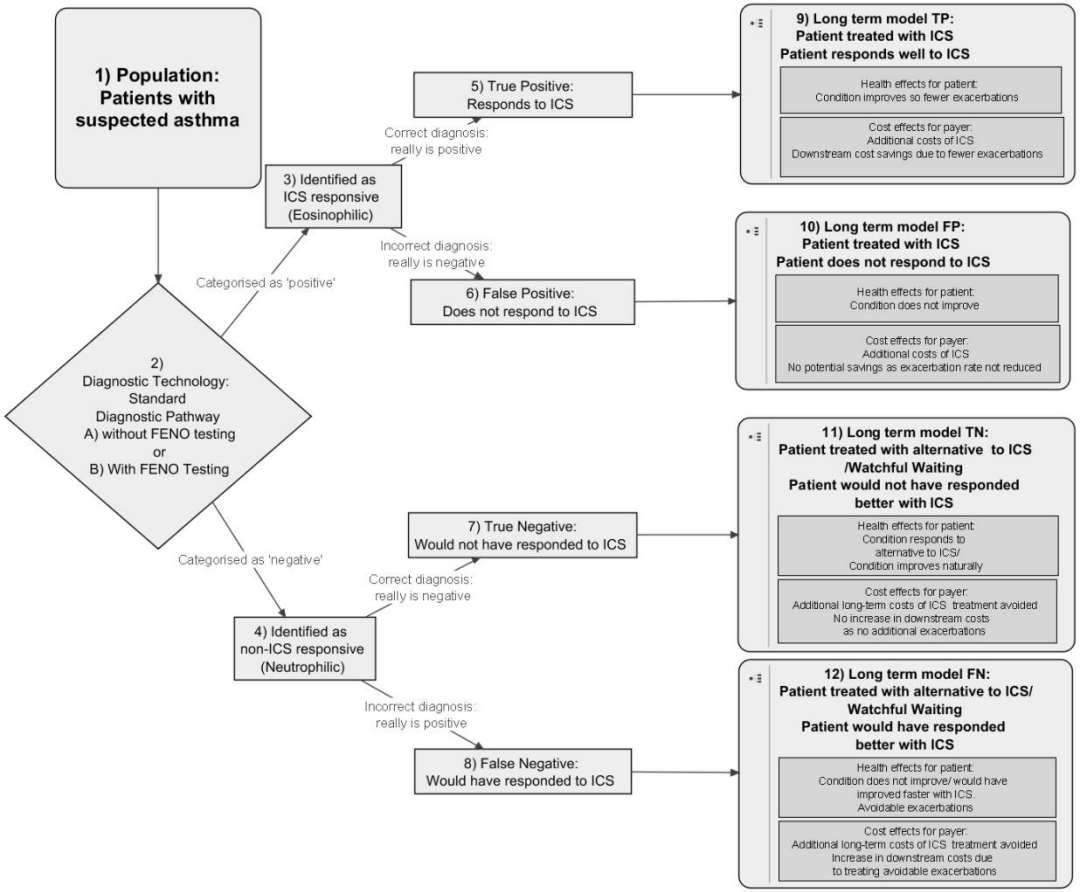


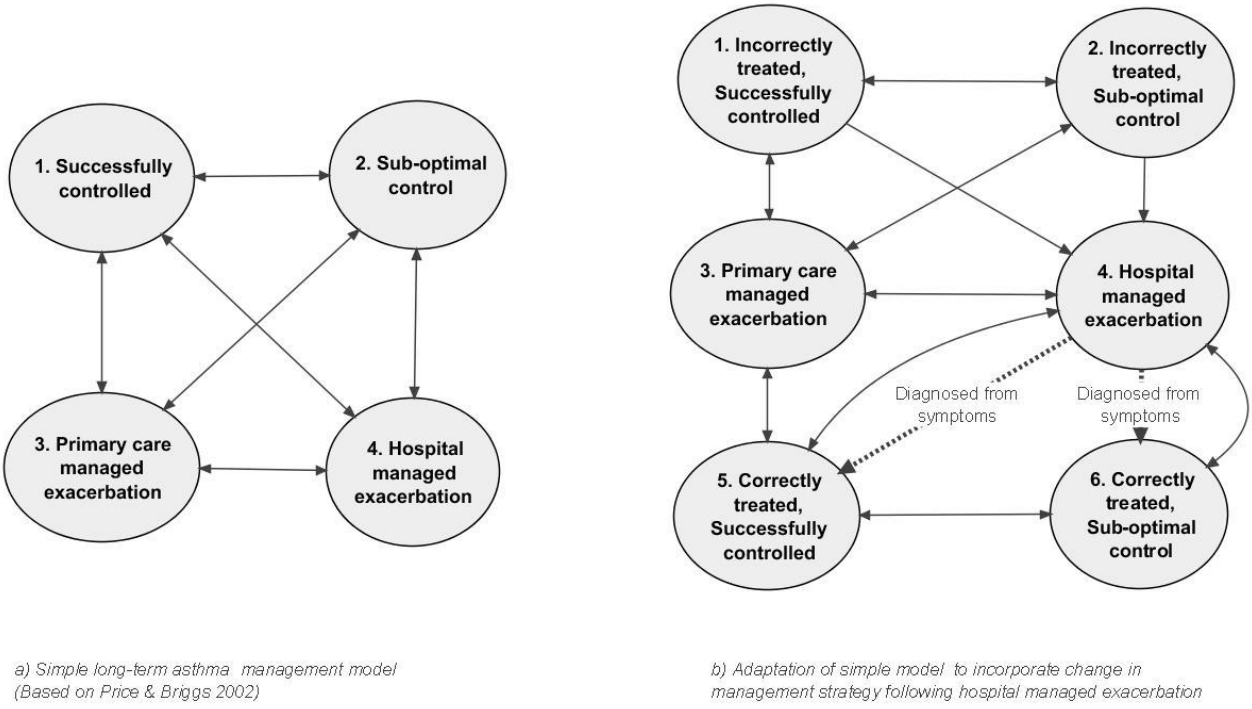
Figure 1 Illustrative short term diagnostic model

The different end states of the diagnostic part of the model are four separate nodes. These nodes correspond to the expected patient care pathways and clinical outcomes for true positives (TPs), false positives (FPs), true negatives (TNs) and false negatives (FNs). Both TPs and FPs will receive the same patient management, although in the model only the true positives will see a significant improvement in their symptoms as a result.

Similarly, both TNs and FNs will receive the same patient management, but the health consequences will be better for TNs than FNs. The best clinical outcomes are achieved if all patients are correctly diagnosed - i.e. all patients are either true positives or true negatives. However misclassification errors mean that some proportion of patients will either be FPs or FNs. The marginal clinical utility of a better diagnostic test results from how the test results lead to increases in the proportion of patients receiving appropriate outcomes, and a reduction in the proportion receiving inappropriate outcomes. However, as the cost and health consequences of TPs, FPs, TNs and FNs are not necessarily equal, the most clinically or cost effective outcomes may not result from a strategy with the lowest rate of misclassification.

4.3.1 Illustrative long-term management model

An illustrative conceptual model describing the management component of the economic model is shown in Figure 2 below**.**



a) Simple long-term asthma management model (Based on Price & Briggs 2002)

b) Adaptation of simple model to incorporate change in management strategy following hospital managed exacerbation

Figure 2 Illustrative long-term management model

The conceptual model shown in Figure 2a is based on an existing asthma management model. [17] It is flexible enough that the four long-term nodes in the diagnostic model could all, in principle, be represented by different parameterisations of the same model structure. It could also be adapted structurally to incorporate a range of alternative scenarios, such as that shown in Figure 2b. In this figure it is assumed that, if a patient who has been receiving suboptimal treatment due to an incorrect diagnosis experiences a severe exacerbation which leads to hospitalisation, then as a result more comprehensive tests are conducted in secondary care, leading to a correct diagnosis, and so improved condition management afterwards. Similar structural adaptations could be produced following expert clinical guidance. Models featuring more states will require more model parameters to be populated, and where good quality data are not available to do this may mean stronger assumptions will be required in their construction.

A wide range of scenarios and sensitivity analyses will be conducted to evaluate the influence of a range of types of uncertainty on estimated outcomes and decision uncertainty.

5. Handling information from the companies

All relevant data submitted by the manufacturers/sponsors will be considered if received by the TAR team no later than 1 May 2013. Data arriving after this date are unlikely to be considered, except for data specifically requested by the EAG. If the data meet the inclusion criteria for the review they will be extracted and quality assessed in accordance with the procedures outlined in this protocol.

Any economic evaluations included in company submissions, provided that they comply with NICE's advice on presentation, will be assessed for clinical relevance, reasonableness of assumptions and appropriateness of the data used in the economic model. If the EAG judge that the existing economic evidence is not robust, then further work will be undertaken, either by adapting what already exists or developing a *de novo* model.

Any 'commercial in confidence' data taken from a company submission, and specified as confidential in the check list, will be highlighted in blue and underlined in the assessment report (followed by an indication of the relevant company name e.g. in brackets).Any 'academic in confidence' data provided by manufacturers, and specified as such, will be highlighted in yellow and underlined in the assessment report. Any confidential data used in the cost-effectiveness models will also be highlighted.

References

1. Payne DN, Adcock IM, Wilson NM, Oates T, Scallan M, Bush A (2001) Relationship between exhaled nitric oxide and mucosal eosinophilic inflammation in children with difficult asthma, after treatment with prednisalone. Am J Respir Crit Care Med 164: 1376-1381.

2. British Thoracic Society and Scottish Intercollegiate Guidelines Network. (2012) British Guideline on the Management of Asthma: a national clinical guideline.

3. Dweik RA, Boggs PB, Erzurum SC, Irvin CG, Leigh MW, Lundberg JO, Olin AC, Plummer AL, Taylor DR (2011) An official ATS clinical practice guideline: interpretation of exhaled nitric oxide levels (FeNO) for clinical applications. Am J Respir Crit Care Med 184: 602-615. 184/5/602 [pii];10.1164/rccm.9120-11ST [doi].

4. Aerocrine. (2012) Guide to Interpretation of eNO Values.

5. 2005) ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. Am J Respir Crit Care Med 171: 912-930. 171/8/912 [pii];10.1164/rccm.200406-710ST [doi].

6. NICE (2008) Guide to the methods of technology appraisal.

7. National Institute for Health and Clinical Excellence (2012) Clinical guideline development methods.

8. Petsky HL, Cates CJ, Li A, Kynaston JA, Turner C, Chang AB (2009) Tailored interventions based on exhaled nitric oxide versus clinical symptoms for asthma in children and adults (Review).

9. Aerocrine (2013) Meta Analysis of Asthma Exacerbation Rates with FeNO Guided Asthma Management.

10. [Anonymous] (2011) Cochrane Handbook for systematic reviews of interventions Version 5.1.0.

11. Centre for Reviews and Dissemination (2009) Systematic reviews. CRD's guidance for undertaking reviews in healthcare.

12. Whiting PF RAWMMSDJRJLMSJaBP (2011) QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Ann Intern Med 155: 529-536.

13. Curtis L (2010) Unit Costs of Health and Social Care 2010.

14. DoH (2011) NHS Reference Costs 2009-2010. 2012.

15. Price D, Berg J, Lindgren P (2009) An economic evaluation of NIOX MINO airway inflammation monitor in the United Kingdom. Allergy 64: 431-438.

16. Kaltenthaler E, Tappenden P, Paisley S, Squires H (2011) NICE DSU Technical Support Document 13: Identify and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models.

17. Price MJ, Briggs AH (2002) Development of an economic model to assess the cost effectivneess of asthma management strategies. Pharmacoeconomics 20: 183-194.

Competing interests of authors

None

Timetable/milestones

Milestone	Date to be completed
TAR Centre meets with the assessment subgroup to discuss the draft scope and the draft protocol	19-Feb-13
Final protocol prepared by TAR Centre and sent to NICE cc NETSCC, HTA.	05-Mar-13
Progress report to NETSCC, HTA (stripped version forwarded to NICE for info)	05-Jun-13
Draft version of DAR and executable economic model from TAR Centre to NICE	31-Jul-13
TAR Centre deliver final DAR simultaneously to NICE and NETSCC, HTA including executable (or redacted if applicable) economic model	29-Aug-13
NICE circulate stripped DAR and executable or redacted economic model to registered stakeholders	05-Sep-13
Registered stakeholder comments on DAR and economic model sent to TAR Centre	23-Sep-13
TAR Centre responds verbally or submits a written response to relevant stakeholder comments for inclusion in DAC meeting papers	01-Oct-13
4 week public consultation on the DCD	27-Nov-13

Consultation comments on DCD forwarded from NICE to AG for optional response	02-Dec-13
NICE receives AG's responses (written or verbal) to DCD consultation comments	05-Dec-13
Second Diagnostics Advisory Committee meeting (to consider final recommendations)	11-Dec-13

Appendices

Appendix A Search Strategy

1. NIOX MINO.mp.
2. aerocrine.mp.
3. (niox adj5 (monitor$ or chemiluminescence or analyser$ or sensor or device$ or desktop)).mp.
4. NObreath.mp.
5. bedfont.mp.
6. or/1-5
7. exp cough/
8. cough$.mp.
9. phlegm.mp.
10. sputum.mp.
11. mucus.mp.
12. wheez$.mp.
13. chest pain/
14. chest pain$.mp.
15. (chest adj5 tight$).tw.
16. ((lower respiratory or lrt) adj5 symptom$).tw.
17. (lower airway adj5 symptom$).tw.
18. ((trache$ or wind pipe or lung$ or bronch$) adj3 symptom$).tw.
19. exp lung/ or trachea/
20. symptom$.tw.
21. 19 and 20
22. or/7-18,21
23. exp asthma/
24. asthma$.mp.
25. exp respiratory hypersensitivity/
26. exp bronchial hyperreactivity/
27. bronchial spasm/
28. bronchospas$.mp.
29. exp Bronchoconstriction/
30. bronchoconstric$.mp.

31. (bronch$ adj3 constrict$).mp.
32. (bronch$ adj5 spas$).mp.
33. (airway$ adj5 (obstruct$ or inflammation$)).mp.
34. ((bronchial$ or respiratory or airway$ or lung$) adj3 (hypersensitiv$ or hyperreactiv$ or allerg$ or insufficiency)).mp.
35. or/23-34
36. Nitric Oxide/
37. nitric oxide.mp.
38. 36 or 37
39. (exhal$ or expir$ or alveolar).mp.
40. 38 and 39
41. exhaled NO.mp.
42. eno.mp.
43. fe?no$.mp.
44. (fractional adj2 NO).mp.
45. or/40-44
46. 22 and 45

Where applicable, the following filters could be applied:

RCT filter

1. Randomized controlled trials as Topic/
2. Randomized controlled trial/
3. Random allocation/
4. randomized controlled trial.pt.
5. Double blind method/
6. Single blind method/
7. Clinical trial/
8. exp Clinical Trials as Topic/
9. controlled clinical trial.pt.
10. or/1-9
11. (clinic$ adj25 trial$).ti,ab.
12. ((singl$ or doubl$ or treb$ or tripl$) adj (blind$ or mask$)).tw.
13. Placebos/
14. Placebo$.tw.

15. (allocated adj2 random).tw.
16. or/11-15
17. 10 or 16
18. Case report.tw.
19. Letter/
20. Historical article/
21. 18 or 19 or 20
22. exp Animals/
23. Humans/
24. 22 not (22 and 23)
25. 21 or 24
26. 17 not 25

Systematic review filter

1. meta-analysis as topic/
2. (meta analy$ or metaanaly$).tw.
3. Meta-Analysis/
4. (systematic adj (review$1 or overview$1)).tw.
5. "Review Literature as Topic"/
6. or/1-5
7. (cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or cinhal or science citation index or bids or cancerlit).ab.
8. ((reference adj list$) or bibliograph$ or hand-search$ or (relevant adj journals) or (manual adj search$)).ab.
9. ((selection adj criteria) or (data adj extraction)).ab.
10. "review"/
11. 9 and 10
12. comment/ or editorial/ or letter/
13. Animals/
14. Humans/
15. 13 not (13 and 14)
16. 12 or 15
17. 6 or 7 or 8 or 11
18. 17 not 16

Diagnostic filter

1. exp "sensitivity and specificity"/
2. sensitivity.tw.
3. specificity.tw.
4. ((pre-test or pretest) adj probability).tw.
5. post-test probability.tw.
6. predictive value$.tw.
7. likelihood ratio$.tw.
8. *Diagnostic Accuracy/
9. or/1-9

Appendix B: Additional information that is needed by NETSCC, HTA and NICE.

Appendix B.1 Details of External Assessment Group

Jon Minton
Research Associate
Health Economics and Decision Science
School of Health and Related Research
The University of Sheffield
Regent Court, 30 Regent Street
Sheffield, S1 4DA
Tel: +44 (0) 114 222 0836
E-mail: j.minton@sheffield.ac.uk

Susan Harnan
Research Fellow
Health Economics and Decision Science
School of Health and Related Research
The University of Sheffield
Regent Court, 30 Regent Street
Sheffield, S1 4DA
Tel: (+44) (0)114 222 0869
E-mail: s.harnan@sheffield.ac.uk

Munira Essat
Research Associate
Health Economics and Decision Science
School of Health and Related Research
The University of Sheffield
Regent Court, 30 Regent Street
Sheffield, S1 4DA
Tel: (+44) (0)114 222 0860
E-mail: m.essat@sheffield.ac.uk

Ruth Wong
Information Specialist

Health Economics and Decision Science
School of Health and Related Research
The University of Sheffield
Regent Court, 30 Regent Street
Sheffield, S1 4DA
Tel: (+44) (0)114 222 0797
E-mail: ruth.wong@sheffield.ac.uk

Paul Tappenden
Reader
Health Economics and Decision Science
School of Health and Related Research
The University of Sheffield
Regent Court, 30 Regent Street
Sheffield, S1 4DA
Tel: (+44) (0)114 222 0855
E-mail: p.tappenden@sheffield.ac.uk

Appendix 2: Medline strategies for searches for the Clinical Review

1. Management review

1	NIOX MINO.mp.
2	aerocrine.mp.
3	(niox adj5 (monitor$ or chemiluminescence or analyser$ or sensor or device$ or desktop)).mp.
4	NObreath.mp.
5	bedfont.mp.
6	or/1-5
7	exp cough/
8	cough$.mp.
9	phlegm.mp.
10	sputum.mp.
11	mucus.mp.
12	wheez$.mp.
13	chest pain/
14	chest pain$.mp.
15	(chest adj5 tight$).tw.
16	((lower respiratory or lrt) adj5 symptom$).tw.
17	(lower airway adj5 symptom$).tw.
18	((trache$ or wind pipe or lung$ or bronch$) adj3 symptom$).tw.
19	exp lung/ or trachea/
20	symptom$.tw.
21	19 and 20
22	or/7-18,21
23	exp asthma/
24	asthma$.mp.
25	exp respiratory hypersensitivity/
26	exp bronchial hyperreactivity/
27	bronchial spasm/
28	bronchospas$.mp.
29	exp Bronchoconstriction/
30	bronchoconstric$.mp.
31	(bronch$ adj3 constrict$).mp.
32	(bronch$ adj5 spas$).mp.
33	(airway$ adj5 (obstruct$ or inflammation$)).mp.
34	((bronchial$ or respiratory or airway$ or lung$) adj3 (hypersensitiv$ or hyperreactiv$ or allerg$ or insufficiency)).mp.
35	or/23-34
36	Nitric Oxide/
37	nitric oxide.mp.
38	36 or 37
39	(exhal$ or expir$ or alveolar or fractional).mp.
40	38 and 39 (5228)
41	exhaled NO.mp.
42	eno.mp.
43	fe?no$.mp.
44	(fractional adj2 NO).mp.
45	or/40-44

46	22 and 45
47	35 and 45
48	6 or 46 or 47
49	limit 48 to yr="2009 -Current"

2. Systematic reviews search

1	NIOX MINO.mp.
2	aerocrine.mp.
3	(niox adj5 (monitor$ or chemiluminescence or analyser$ or sensor or device$ or desktop)).mp.
4	NObreath.mp.
5	bedfont.mp.
6	or/1-5
7	exp cough/
8	cough$.mp.
9	phlegm.mp.
10	sputum.mp.
11	mucus.mp.
12	wheez$.mp.
13	chest pain/
14	chest pain$.mp.
15	(chest adj5 tight$).tw.
16	((lower respiratory or lrt) adj5 symptom$).tw.
17	(lower airway adj5 symptom$).tw.
18	((trache$ or wind pipe or lung$ or bronch$) adj3 symptom$).tw.
19	exp lung/ or trachea/
20	symptom$.tw.
21	19 and 20
22	or/7-18,21
23	exp asthma/
24	asthma$.mp.
25	exp respiratory hypersensitivity/
26	exp bronchial hyperreactivity/
27	bronchial spasm/
28	bronchospas$.mp.
29	exp Bronchoconstriction/
30	bronchoconstric$.mp.
31	(bronch$ adj3 constrict$).mp.
32	(bronch$ adj5 spas$).mp.
33	(airway$ adj5 (obstruct$ or inflammation$)).mp.
34	((bronchial$ or respiratory or airway$ or lung$) adj3 (hypersensitiv$ or hyperreactiv$ or allerg$ or insufficiency)).mp.
35	or/23-34
36	Nitric Oxide/
37	nitric oxide.mp.
38	36 or 37
39	(exhal$ or expir$ or alveolar or fractional).mp.
40	38 and 39 (5228)

41	exhaled NO.mp.
42	eno.mp.
43	fe?no$.mp.
44	(fractional adj2 NO).mp.
45	or/40-44
46	22 and 45
47	35 and 45
48	6 or 46 or 47
49	meta-analysis as topic/
50	(meta analy$ or metaanaly$).tw.
51	Meta-Analysis/
52	(systematic adj (review$1 or overview$1)).tw.
53	"Review Literature as Topic"/
54	or/49-53 (96944)
55	(cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or cinhal or science citation index or bids or cancerlit).ab.
56	((reference adj list$) or bibliograph$ or hand-search$ or (relevant adj journals) or (manual adj search$)).ab.
57	((selection adj criteria) or (data adj extraction)).ab.
58	"review"/
59	57 and 58
60	comment/ or editorial/ or letter/
61	Animals/
62	Humans/
63	61 not (61 and 62)
64	60 or 63
65	54 or 55 or 56 or 59
66	65 not 64
67	48 and 66

3. RCT studies search

1	NIOX MINO.mp.
2	aerocrine.mp.
3	(niox adj5 (monitor$ or chemiluminescence or analyser$ or sensor or device$ or desktop)).mp.
4	NObreath.mp.
5	bedfont.mp.
6	or/1-5
7	exp cough/
8	cough$.mp.
9	phlegm.mp.
10	sputum.mp.
11	mucus.mp.
12	wheez$.mp.
13	chest pain/
14	chest pain$.mp.
15	(chest adj5 tight$).tw.
16	((lower respiratory or lrt) adj5 symptom$).tw.
17	(lower airway adj5 symptom$).tw.

18	((trache$ or wind pipe or lung$ or bronch$) adj3 symptom$).tw.
19	exp lung/ or trachea/
20	symptom$.tw.
21	19 and 20
22	or/7-18,21
23	exp asthma/
24	asthma$.mp.
25	exp respiratory hypersensitivity/
26	exp bronchial hyperreactivity/
27	bronchial spasm/
28	bronchospas$.mp.
29	exp Bronchoconstriction/
30	bronchoconstric$.mp.
31	(bronch$ adj3 constrict$).mp.
32	(bronch$ adj5 spas$).mp.
33	(airway$ adj5 (obstruct$ or inflammation$)).mp.
34	((bronchial$ or respiratory or airway$ or lung$) adj3 (hypersensitiv$ or hyperreactiv$ or allerg$ or insufficiency)).mp.
35	or/23-34
36	Nitric Oxide/
37	nitric oxide.mp.
38	36 or 37
39	(exhal$ or expir$ or alveolar or fractional).mp.
40	38 and 39 (5228)
41	exhaled NO.mp.
42	eno.mp.
43	fe?no$.mp.
44	(fractional adj2 NO).mp.
45	or/40-44
46	22 and 45
47	35 and 45
48	6 or 46 or 47
1	NIOX MINO.mp.
2	aerocrine.mp.
3	(niox adj5 (monitor$ or chemiluminescence or analyser$ or sensor or device$ or desktop)).mp.
4	NObreath.mp.
5	bedfont.mp.
6	or/1-5

49	Randomized controlled trials as Topic/
50	Randomized controlled trial/
51	Random allocation/
52	randomized controlled trial.pt.
53	Double blind method/
54	Single blind method/
55	Clinical trial/
56	exp Clinical Trials as Topic/
57	controlled clinical trial.pt.
58	or/49-57
59	(clinic$ adj25 trial$).ti,ab.
60	((singl$ or doubl$ or treb$ or tripl$) adj (blind$ or mask$)).tw.
61	Placebos/
62	Placebo$.tw.
63	(allocated adj2 random).tw.
64	or/59-63
65	58 or 64
66	Case report.tw.
67	Letter/
68	Historical article/
69	66 or 67 or 68
70	exp Animals/
71	Humans/
72	70 not (70 and 71)
73	69 or 72
74	65 not 73
75	48 and 74

4. Diagnostic studies search

1	NIOX MINO.mp.
2	aerocrine.mp.
3	(niox adj5 (monitor$ or chemiluminescence or analyser$ or sensor or device$ or desktop)).mp.
4	NObreath.mp.
5	bedfont.mp.
6	or/1-5
7	exp cough/
8	cough$.mp.
9	phlegm.mp.
10	sputum.mp.
11	mucus.mp.
12	wheez$.mp.
13	chest pain/
14	chest pain$.mp.
15	(chest adj5 tight$).tw.
16	((lower respiratory or lrt) adj5 symptom$).tw.
17	(lower airway adj5 symptom$).tw.
18	((trache$ or wind pipe or lung$ or bronch$) adj3 symptom$).tw.

19	exp lung/ or trachea/
20	symptom$.tw.
21	19 and 20
22	or/7-18,21
23	exp asthma/
24	asthma$.mp.
25	exp respiratory hypersensitivity/
26	exp bronchial hyperreactivity/
27	bronchial spasm/
28	bronchospas$.mp.
29	exp Bronchoconstriction/
30	bronchoconstric$.mp.
31	(bronch$ adj3 constrict$).mp.
32	(bronch$ adj5 spas$).mp.
33	(airway$ adj5 (obstruct$ or inflammation$)).mp.
34	((bronchial$ or respiratory or airway$ or lung$) adj3 (hypersensitiv$ or hyperreactiv$ or allerg$ or insufficiency)).mp.
35	or/23-34
36	Nitric Oxide/
37	nitric oxide.mp.
38	36 or 37
39	(exhal$ or expir$ or alveolar or fractional).mp.
40	38 and 39 (5228)
41	exhaled NO.mp.
42	eno.mp.
43	fe?no$.mp.
44	(fractional adj2 NO).mp.
45	or/40-44
46	22 and 45
47	35 and 45
48	6 or 46 or 47
49	**exp "Sensitivity and Specificity"/**
50	**sensitivity.tw.**
51	**specificity.tw.**
52	**((pre-test or pretest) adj probability).tw.**
53	**post-test probability.tw.**
54	**predictive value$.tw.**
55	**likelihood ratio$.tw.**
56	**or/49-55**
57	**48 and 56**

5. Analytic validity studies search

1	NIOX MINO.mp.
2	aerocrine.mp.
3	(niox adj5 (monitor$ or chemiluminescence or analyser$ or sensor or device$ or desktop)).mp.
4	NObreath.mp.
5	bedfont.mp.
6	or/1-5

6. Trial registers and website search

Clinicaltrials.gov (http://www.clinicaltrials.gov/)
21st March 2013
16 studies found for: niox
10 studies found for: mino | asthma
12 studies found for: aerocrine
no studies found for: NObreath
no studies found for: bedfont
31 studies found for: fractional exhaled nitric oxide | asthma

metaRegister of Controlled Trials (http://www.controlled-trials.com/mrct/)

3 studies found for: niox
3 studies found for: mino
2 studies found for: aerocrine
no studies found for: NObreath
1 study found for: bedfont
2 studies found for: fractional exhaled nitric oxide

Manufacturer and User Facility Device (MAUDE)
(http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm)
No records were found with
Brand Name: *NIOX MINO* Report Date From: *01/01/2000* Report Date To: *02/28/2013*
No records were found with
Manufacturer: *aerocrine* Report Date From: *01/01/2010* Report Date To: *02/28/2013*
No records were found with
Brand Name: *NObreath* Report Date From: *01/01/2010* Report Date To: *02/28/2013*
No records were found with
Brand Name: *bedfont* Report Date From: *01/01/2010* Report Date To: *02/28/2013*

EuroScan International Network (http://euroscan.org.uk/)
'2' results for NIOX MINO
No records were found with aerocrine
No records were found with bedfont
No records were found with NObreath
13 results for fractional exhaled nitric oxide asthma

Appendix 3: Clarification of scope: communication with SCM clinicians

Queries on scope and study selection	Study ID's	Our thoughts	Clifford Godley	John White	Hasan Arshad	Our decision
Population						
1. Unselected populations (not selected on basis of asthmatic symptoms, but patients with symptoms may be selected from the wider cohort, to establish diagnostic cut-off) – eg birth cohorts, a school year, a college year, a group in a particular profession	483, 525, 654, 661, 1109, 3011, 6204	We would like to exclude these as they do not reflect how the device would work in UK clinical practice and the population selected would likely impact on the estimates of sensitivity and specificity.	Agreed	Agree	Agree	Exclude unselected populations
2. Very young children – what is minimum age? a. Patients with infant wheeze – if we are including very young children, are these patients equivalent to	ID 6047, 5136	We would like to exclude studies where patients are >10% under 4 years of age, in line with Aerocrine's CE documentation. Bedfont do not state a minimum age.	Agreed	It would be good if there were an alternate measure but practical difficulties mean cut-off at	NO can be measured reliably in younger children. Potentially, it could be a very useful test (if it works) precisely	Exclude studies with >10% patients under 5 years of age.

patients with suspected asthma?		Alternatively, we could consider a cut-off of 5 years, as the UK guidelines draw this distinction in their protocols, as children under this age cannot reliably perform spirometry, a key stage in the diagnostic process.		5 reasonable as you suggest	because other diagnostic methods (spirometry, reversibility and BHR) are not practical. However, in the absence of "Gold standard" it is difficult to conclude if NO is a useful diagnostic test. So agree to keep the minimum age 5 years.	
3. All patients already have asthma, and the reference standard is one of a. responsiveness to ICS, b. exercise induced bronchoconstriction, c. severe asthma, d.	ID 350, 613, 6014, 6164, 7500	a. Responsiveness to ICS studies could be seen to fall in-between diagnosis and management – see attached "Diagnostic pathway for asthma in children" – is this useful? Is	Responsiveness to ICS is effectively a diagnostic trial, a positive response to which lends support to the postulated diagnosis. I would therefore like to keep this	It is both - initially diagnostic	"Response to steroid" is primarily a management issue (although it can be argued that the test can be used to "diagnose" steroid responsive asthma). I think this question is better considered	No clear consensus – but keep these studies in, present separately Latterly think only include if they are being used diagnostically for asthma.

methacholine/mannitol responsiveness, e. high levels of sputum eosinophilia		this diagnostic or management? Do any of the other reference standards have any potential use in patients already diagnosed with asthma? We would probably like to exclude if not.	group, particularly if symptoms relapse off steroids	Yes as they might suggest alck of "control" like FeNO	with other management issues.	
4. Patients with chronic cough – are these patients equivalent to patients with asthma symptoms? These studies usually diagnose ICS responsiveness (but not whether they have asthma or not) or Cough Variant Asthma	328, 5878, 6383, 6385	Unsure what to do with these.	Raised N>O> will be a useful diagnostic sign post in the investigation of chronic cough. Normal N>O> will not preclude cough variant asthma but will make it less likely.	Include as area of clinical relevance	Keep this as it might be useful to know if NO helps to diagnose "cough variant asthma".	Keep in IF diagnose CVA, present separately
5. Study only includes children with positive skin prick test (no	6345	We would like to exclude – only includes atopic children, not on	Agreed	Others will have clearer view on this. I	Agree	Exclude studies which select patients on the basis

asthma symptoms necessarily)		basis of asthma symptoms, so does not match the population we are interested in in UK practice.		would include – sorry!		of positive skin prick test only.
6. Study only includes patients with sever refractory asthma or moderate asthma. Reference standard is one of: a. ICS responsiveness b. Eosinophilic phenotype	6321, 6257	a. Responsiveness to ICS studies could be seen to fall in-between diagnosis and management – see attached "Diagnostic pathway for asthma in children". – is this useful? Is this diagnostic or management? b. Is eosinophilic phenotype a useful outcome? Management or diagnostic?	I think it is used diagnostically in the first instance If they have asthma I would include	As above. Yes as relates to accurate diagnosis and management strategy	a. As suggested above, responsiveness to steroid is more of a management issue b. NO is closely associated with eosinophilic asthma but in itself it is not a useful outcome in terms of managem	Include studies which select patients with asthma and diagnose ICS responsiveness and/or eosinophilic phenotype. Present separately.

					ent.	
7. Study only included patients with suspected occupational asthma, reference standard was positive specific inhalation challenge	387	Is this population too specific, or is it useful?	Should be included - useful	Emphasis on this area eg NICE QS. It could be useful	This population is distinct and it could be useful to see if NO can help to diagnose occupational asthma.	Include studies in patients with occupational asthma
8. Patients with suspected western red cedar asthma, or suspicion of any other very specific type of asthma (eg occupational asthma)		Is this population too specific, or is it useful?	useful	Think I might let you off here	Western red cedar asthma is too specific and targets a small population. However, occupational asthma in general is a significant problem.	Include with occupational asthma studies
9. Only patients diagnosed with occupational asthma - reference standard/diagnostic target of ICS responsiveness	617	Is this population too specific, or is it useful?	useful	See 7	See above	As 7.
10. Study included	106	Is this population too	useful	See 7 include	Asthma and	Include studies in

patients with rhinitis AND asthma symptoms – useful group to include?		specific, or is it useful?			rhinitis coexist in significant proportion of patients. This is therefore a large group and useful to include.	patients with rhinitis AND asthma.
11. Population a little odd…. See data in far right column	6184	Inclusion criteria states "children 6 to 16 years of age referred to our pulmonary outpatient clinic for further diagnostic work-up of possible reactive airway disease", but the results section then states "the major complaints leading to referral were: exercise-induced shortness of breath, physician diagnosed asthma, chronic cough or miscellaneous leading symptoms" ie some patients were already diagnosed with asthma. Is this a relevant cohort? Is a	My understanding is that more than 75% of patients with unexplained chronic cough referred to clinics are ultimately considered to have airways disease. Physicians diagnosed asthma patients should be included .	It's mainly a clinical diagnosis – supported by response to Rx etc Recommendation is further tests if this doesn't help. This is a further test!	I think this is an appropriate population of childhood asthma. I did not find the inclusion criteria unusual.	Include this study.

		physician diagnosis of asthma sufficiently unreliable to class these patients as "patients with symptoms suggestive of asthma"?? Some other studies excluded such patients.	This heterogeneous group of patients compromise the real world population . Raised N.O. is a most invaluable result.			
12. Are patients with chronic cough and FEV_1>80% predicted equivalent to the hard to diagnose group?	6169	If so, we can include this study.	Yes	Yes	Yes	Include this study
Intervention						
13. Offline measurements – can we exclude studies that use this?	6204, 7038, 525	We would have to conduct an review of comparability between methods in order to include these studies and they would just add another source of heterogeneity to the results. We have enough data without including	Agreed	I don't know what offline means. Maybe that's what I am hence lack of understanding	Agree.	Exclude offline measurement.

		these studies. We would like to exclude.				
14. Tidal breathing measurements – can we exclude these (mostly in very young children)		If we are only including those aged 4 and up, this is irrelevant. We would like to exclude	Agreed	OK	Agree	Exclude tidal breathing methods
15. Studies which use a different flow rate, but convert to FeNO 50 – can we exclude (this specifically relates to an RCT management study that was included in the Cochrane review – attached to email)	ID 7704, smith 2005	We would like to exclude – not convinced that the "equivalence" can be assured.	Agreed	OK	As far as I can see the study is sound; so no reason to exclude the study.	Include, but maybe do a subgroup analysis where the results are excluded.
16. Weird flow rate - please see attached study.	ID 617	Unsure what to do.		Prob exclude	Could not find the study in the attachment	Send again to clinicians.
17. Alveolar and bronchial NO measurements – can we exclude?		We would like to exclude – not the use intended by the manufacturer in this application.	Agreed	OK Not relevant or likely to be to routine practice	Agree	Exclude alveolar and bronchial NO measurements.
18. Exhaled breath condensate – we have		We would like to exclude.	Agreed	OK – as above	Agree	Exclude exhaled breath condensate

excluded these.						studies.
19. Laser spectroscopy – this is not in scope, correct?		We would like to exclude.	Agreed	OK	Agree	Exclude laser spectroscopy
20. Nasal NO measurements – can we exclude?		We would like to exclude	Agreed	Might tell you something different. Used in CF I think OK to exclude	Agree	Exclude nasal NO measurement.
21. Portable FeNO devices other than NIOX MINO and NObreath: NO Vario and Medisoft devices – should we include these in the same way that we are including chemiluminescence evidence, as equivalent? We may not have any analytic validity study data to support equivalence of devices.		We would like to exclude, unless they provide data on a subgroup that we have no other data for. This is because we would have to review the equivalence evidence for these devices as well as NIOX MINO and NObreath, which widens the scope of that review.	Agreed	Keen to see this as generic FeNO but would need equivalence data to allow that. If can't be done then excluded by default	I am not sure of the validity or equivalency of NO Vario and Medisoft devices. I have not used them and have no experience. Unless the companies can provide these data, it might be better to exclude these studies.	Include if we have equivalency studies that allow the comparison of NOVario and Medisoft devices.
Reference Standard/ outcome/diagnostic target						

22. Is exercise induced asthma/bronchoconstriction a useful diagnostic target? Are they the same thing, even?? Are they the same as the "Airway hyperresponsiveness testing" listed in UK guidelines?	ID 5171, 6014 (population all asthmatics), 6047 (population wheezy children)	Unsure what to do. If same as UK airway hyperresponsiveness testing, we can include.	This is a form of hyperresponsivness testing along with methacholine and histamine challenge. It is diagnostic for asthma. Include	May not be the same. Suggest include	Exercise induced bronchial challenge provides a valid outcome. It estimates bronchial hyperresponsiveness just like methacholine or histamine bronchial challenge. For childhood asthma in general and exercise induced asthma in particular, it might be superior to other types of bronchial challenges.	Include studies which use EIB as the reference standard.
23. Studies which diagnose asthma severity, usually in an already diagnosed asthma population		Unsure if useful to diagnose asthma severity – where population is all asthmatic, this may occupy an intermediate position between the	<u>Useful if it facilitated the</u> diagnosis of asthma or helps identify uncontrolled asthma	Not sure how this helps really	If NO can provide an objective indication of asthma severity in those already diagnosed with the condition, then	Include diagnosis of asthma severity in already diagnosed population

		diagnostic strategy and the management strategy that has not yet been captured in the review – as discussed above in "population" item 3a.			this might be a valid use of the test.	Exclude diagnosis of asthma severity in an undiagnosed population, unless this facilitates diagnosis of asthma or uncontrolled asthma
24. Studies which diagnose ICS responsiveness, in a variety of populations eg a. chronic cough b. diagnosed asthmatics c. suspected asthma d. severe refractory asthma e. moderate asthma	eg ref 6169, 613, 617, 5878, 6164, 6321(sputum cell count), chronic cough papers 328, 6383, 6385	For chronic cough, we would have to present this data separately, as the diagnosis may not always be one of asthma, just one of ICS responsiveness. However, it may be better to exclude for consistency. Would this data be useful to the committee? For asthmatic populations, this may occupy an intermediate position between the diagnostic strategy and the management	Data would be very useful I feel I am comfortable to consider that patients with chronic cough who demonstrate ICS responsiveness should be considered to have cough variant asthma.	Data ARE pleural so "these data". Sorry again! Anyway ICS responsiveness will effectively be asthma though not all asthma is ICS responsive. Include if possible	Yes, it would be useful to know if NO can provide a reliable indication of steroid responsiveness in those with chronic cough. For other asthmatic populations, it would be useful to assess the value of NO in estimating steroid responsiveness, although, the answer is probably yes in most asthma	Include studies which diagnose ICS responsiveness in those with or without asthma.

		strategy that has not yet been captured in the review – as discussed above in "population" item 3a.			phenotypes.	
25. Asthma (as opposed to ICS responsiveness, ie diagnoses both eosinophilic and non-eosinophilic asthma)		This will influence the estimates of sensitivity and specificity we will get out of the review – if we are using ICS responsiveness as a reference standard, feno is likely to give higher diagnostic accuracy than if we are using a broader definition of asthma as the reference standard, if we believe that it is better at identifying eosinophilic asthma. It may be best to present both sets of data? Another thing to consider is what will be useful to the model –	Certainly it will be more important to highlight the differences between eosinophilic and neutrophilic asthma with regard to N.O. levels and likely steroid responsiveness.	Agree probably	Agree, it would be better to present both sets of data. Studies show that NO is better in diagnosing atopic than non-atopic asthma. Indeed,	Include studies which use asthma as the diagnostic reference standard.

		some of the RCT studies recruit atopic patients – so if we are matching data from the diagnostic review to the RCT data for the modelling, would we need to use the diagnostic studies which identify ICS responsive asthmatics to lead into the modelling of the management strategy?			some studies show that it is a marker for atopy, more than it is for asthma. Hence, it would be useful to assess the usefulness of NO (as a diagnostic test) separately in atopic and non-atopic asthma. I suspect the result will be very different.	
26. Studies which use Metacholine/Mannitol or Bronchial hyperresponsiveness challenge as the reference standard	1109, 7500, 6184		include	Include	That is valid if the result of methacholine or mannitol bronchial challenge (bronchial hyperresponsivene ss) is combined with symptoms and not used alone as a diagnostic marker of asthma.	Include studies which use methacholine/mann itol or bronchial hyperresponsivenes s challenge as the reference standard where this is combined with symptoms.

27. Studies which use exposure to trigger (occupational asthma) as a specific inhalation challenge as the reference standard	387	Is the reference standard equivalent to a diagnosis of asthma?	Usually diagnosis of occupational asthma is more robust than the clinical diagnosis of non occupational asthma.	Not always!	Fine, but need to consider this as a distinct asthma phenotype (occupational asthma)	Include studies in occupational asthma which use specific inhalation challenge as the reference standard, but present separately.
28. EOS and phenotype	6257, 6164	Similar to diagnosing ICS responsiveness and asthma – does this sit between diagnosis and management? Is it useful?	Sorry – couldn't identify these papers		No, it is not useful.	Send to clinicians again.
29. Sputum eosinophilia	350	IS this a relevant reference standard?	yes	Ian Pavord would say so	For eosinophilic asthma, yes. But that is not a useful phenotype to diagnose in terms of management in day to day practice.	Inlcude studies which use eosinophilia as the reference standard. Present separately.
30. Variety of possibly useless reference standards	6345 (also only recruited children	See table 1 in attached document ID 6345 – are any of these a useful	Sorry – couldn't identify these papers	Can't see it on any list	Could not find Table 1.	Sent to clinicians.

	with positive skin prick test).	reference standard?			As suggested before, studies recruiting those with positive skin test only i.e. atopy if combined with asthma should be considered separately as this asthma phenotype (atopic asthma) is likely to be sensitive to NO.	

Appendix 4: Data extraction forms

Review type	**Data extracted for:**			
	Study background	**Methods and devices**	**Study sample characteristics**	**Results**
Analytic validity	Author / year Source of funding/ conflicts of interest Study design Setting Inclusion/ exclusion criteria Code for population (e.g. asthmatics; asthmatics and healthy) Code for age (adults, children, adolescents)	Devices used: • Niox Mino • NOBreath • Others (1-4) Comparisons made: • Niox Mino vs chemiluminescence • NOBreath vs chemiluminescence • Niox Mino vs NOBreath • Niox Mino vs other Method of FeNO recording (i.e. guidelines used) Authors' conclusions Statistical methods	Total N patients recruited Number withdrawn (with reasons) Number analysed Mean age (±SD) Gender n/N male (%) Diagnosis FEV1%predicated Medication usage Smokers (current/ ex smokers) Atopic status	Test failure rates Mean FeNO Niox Mino ppb Mean FeNO NOBreath ppb Mean FeNO Chemiluminescence pbb Comparison data Correlation coefficient Regression Bland-Altman statistic Comparability of AUC Comparability of cut-off values Correction equation
Diagnostic	Author/ year Source of funding Study design (prospective / retrospective Validation or derivation? Dates Description of study timeline Setting Inclusion/ exclusion criteria Inclusion age range Code for age (children/ adolescents/ adults) Code for population (e.g. symptoms of asthma; difficult to diagnose) Medication permitted during	Name of device Code for device(Niox mino; Nobreath; Chemiluminescence) Code for FENO measurement method (e.g. ATS/ERS; Bespoke) Measurement method details if bespoke FeNO cut-off points Description of reference standard Code for reference standard (e.g. ATS/ERS 2005) Definition of asthma	Total N patients recruited Number not followed up (with reasons) Number of patients analysed Mean age (±SD) Gender n/N male (%) FEV1%predicated Mean FeNO50 ppb (±SD)] Symptoms Smokers Atopic/ allergic status Other sympotms Statistical methods	Prevalence of asthma Prevalence of positive result by reference standard Cut-off value True positive False positive False negative True negative Total Prevalence of ICS responsiveness (95% CI) True positive False positive False negative True positive Sensitivity asthma (95% CI)

	study			Specificity asthma (95% CI) Sensitivity ICS responsiveness (95% CI) Specificity ICS responsiveness (95% CI)
Management	Author/ year Source of funding Study design Timeline of study Setting Inclusion/ exclusion criteria Asthma diagnosis method Inclusion age range Code for population Code for subgroup	Description of intervention Concomitant treatments Code for intervention (e.g. 'control plus FeNO) FeNO measurement method Code for device and method FeNO cut-off Description of control Code for control (e.g. symptoms, spirometry) Step up/ down protocol Definition of exacerbation	Total N patients recruited Number of patients not followed up* Number of patients analysed* Mean age ±SD* Gender n/N male (%)* FEV1% predicated* Symptom/ severity score* FeNO* Medication usage* Smokers (current and ex)* Symptoms* Atopic/ allergic status* Statistical methods Time of outcome measurement	Number of exacerbations** HRQoL** Asthma control** Clinical complications associated with exacerbations** Use of oral corticosteroids** Levels and use of inhaled corticosteroids** Levels and use of other medications** Adverse events** Mortality rate** Compliance** Test failure rate**

Abbreviations: AUC, area under the curve; CI, confidence interval; FEV1%, forced expiratory volume; HRQoL, health-related quality of life

*: Where available, data were extracted separately for the whole cohort, intervention, and control groups

**: Where available, data were extracted separately for intervention and control groups, and for any reporting of between-group comparisons

Appendix 5: Quality assessment scoring criteria

1. Diagnostic studies

Risk of bias in diagnostic studies (for both children and adult populations) was assessed and described using the Bristol University QUADAS-2 tool. QUADAS-2 is structured around four domains of potential sources of bias in primary diagnostic studies: patient selection; index test; reference standard; and flow and timing. There are signalling questions within each of these domains which allow researchers to overview the potential sources of bias therein, and a summary domain score can be generated to provide an indication of overall potential for bias in each aspect of a study's methodology. These signalling questions, and our approach to scoring them, are detailed below. It should be noted that, in our risk of bias tables, we report only the domain summary scores, although we also narratively summarise our responses to signalling questions to support these judgments in the review text.

Domain 1: Patient selection

Signalling question 1: Was a consecutive or random sample of patients enrolled?

- Score 'yes' if the report states enrolment was consecutive or random
- Score 'no' if the report states another method of sampling
- Score 'unclear' if insufficient information was provided to make a judgment

Signalling question 2: Was a case-control design avoided?

- Score 'yes' if the study was not case-control
- Score 'no' if the study was a case-control
- Score 'unclear' if insufficient information was provided to make a judgment

Signalling question 3: did the study avoid inappropriate exclusions?

With respect to the current review, the population of interest was patients presenting with clinical characteristics of asthma, or those who are "difficult to diagnose", that is, patients who have already undergone some of the tests for asthma in the UK pathway and have not yet been confirmed to have asthma. The review scope also sought data on subgroups, in particular, women during pregnancy, older people, and smokers/passive smokers. Hence, this question was answered with respect to these groups. Where there were ambiguities, two reviewers would discuss whether the population was appropriate.

- Score 'yes' if the study had appropriately selected patients conforming to the groups outline above
- Score 'no' if the study made inappropriate exclusions from the group it set out to select
- Score 'unclear' if insufficient information was provided to make a judgment

Summary domain score: Could the selection of patients have introduced bias?

- Score as 'low risk' if the study scored 'yes' on all signalling questions above
- Score as 'high risk' if the study scored 'no' or 'unclear' on $\geq$2 of the individual items, or on either of case-control design and inappropriate exclusions
- Score as 'unclear risk' if the study scored 'no' or 'unclear' on signalling question 1.

Domain 2: Index tests

Signalling question 1: Were the index test results interpreted without knowledge of the results of the reference standard?

- Score 'yes' if index test was interpreted blind to the reference standard or the index test was clearly interpreted before the reference standard was known.
- Score 'no' if results of reference standard (UK guidelines pathway) were already known, or if

parts of the reference standard downstream of the position of the test in the UK pathway were already known. This will need to be scored with reference to the patient population.
- Score 'unclear' if unclear

Signalling question 2: If a threshold was used, was it pre-specified?
- Score 'yes' if pre-specified cut of values were used (validation study)
- Score 'no' if cut-off values were fitted to the data (derivation study)
- Score 'unclear' if unclear

Summary domain score: Could the conduct or interpretation of the index tests have introduced bias?
- Score as 'low risk' if the study scored 'yes' on both signalling questions
- Score as 'high risk' if cut-off values were fitted to the data (since derivation studies are likely to over-estimate the true diagnostic accuracy of a technology relative to clinical practice), or if the study scored 'no' or 'unclear' on both signalling questions
- Score as 'unclear' if the study scored 'no' or 'unclear' on signalling question 1 only

Domain 3: Reference standard

Signalling question 1: Is the reference standard likely to correctly classify the target condition?

No reference standard for Asthma is 100% sensitive or specific. Possibly the only way this could be achieved is with long term follow-up, but even this might be confounded by the fact that asthma can remiss and develop (eg as a comorbidity) over time. Hence, this item was scored with respect to UK guidelines:

- Score 'yes' if the reference standard conforms with all or part of the UK guidelines
- Score 'no' if the reference standard does not conform with UK guidelines i.e. uses a test that is not within the UK guidelines (we should be excluding these anyway)
- Score 'unclear' if unclear

Signalling question 2: Were the reference standard results interpreted without knowledge of the results of the index test?

- Score yes if the reference standard was interpreted blind to the index test or the reference standard was clearly interpreted before the index test was known.
- Score no if the results of the index test were known.
- Score unclear if unclear

Summary domain score: Could the conduct or interpretation of the index tests have introduced bias?
- Score as 'low risk' if the study scored 'yes' on both signalling questions
- Score as 'high risk' if the study scored 'no' or 'unclear' on both signalling questions
- Score as 'unclear' if the study scored 'no' or 'unclear' on either of the two signalling questions

Domain 4: Flow and timing

Signalling question 1: Was there an appropriate interval between index test(s) and reference standard?

- Score 'yes' if tests were conducted consecutively
- Score 'no' if index and reference tests were conducted >1 week apart
- Score unclear if unclear

Signalling question 2: Did all patients receive a reference standard?

- Score yes if yes
- Score no if no
- Score unclear if unclear

Signalling question 3: Did patients receive the same reference standard?

- Score yes if yes
- Score no if no
- Score unclear if unclear

Signalling question 4: Were all patients included in the analysis?

- Score yes if yes
- Score no if no
- Score unclear if unclear

Summary domain score: Could the patient flow have introduced bias?

- Score 'low risk' if the study scored 'yes' on all signalling questions
- Score 'high risk' if the study scored 'no' or 'unclear' on $\geq$3 items
- Score 'unclear risk' if the study score 'no' or 'unclear' on up to two items.

2. Management studies

The quality of the FeNO-guided management studies for adults and children was assessed using the Cochrane Collaboration's tool ('the tool') for assessing risk of bias in randomised controlled trials (RCTs). The tool is designed to address seven domains of bias: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting and 'other issues'. The tool provides a two-part assessment for risk of bias; the first part describes what was reported in each study for each domain, and the second part comprises the review authors' categorisation of the study as 'low', 'high', and 'uncertain' risk of bias (Table 1).

The criteria for risk of bias judgments as outlined in the Cochrane Handbook (Table 2) were used to assign study ratings. As recommended by Cochrane, we did not assign an overall numerical score for risk of bias in each study, but discussed how potential sources of bias among the research literature may be likely to affect the outcomes of the study.

Table 84: The Cochrane Collaboration's tool for assessing risk of bias[32]

Domain	Description	Review authors' judgement
Sequence generation.	Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.	Was the allocation sequence adequately generated?
Allocation concealment.	Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.	Was allocation adequately concealed?
Blinding of participants, personnel and outcome assessors *Assessments should be made for each main outcome (or class of outcomes).*	Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Was knowledge of the allocated intervention adequately prevented during the study?
Incomplete outcome data *Assessments should be made for each main outcome (or class of outcomes).*	Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.	Were incomplete outcome data adequately addressed?
Selective outcome reporting.	State how the possibility of selective outcome reporting was examined by the review authors, and what was found.	Are reports of the study free of suggestion of selective outcome reporting?
Other sources of bias.	State any important concerns about bias not addressed in the other domains in the tool. If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry.	Was the study apparently free of other problems that could put it at a high risk of bias?

Table 85: Criteria for judging risk of bias in the 'Risk of bias' assessment tool (from the Cochrane Handbook)

SEQUENCE GENERATION **Was the allocation sequence adequately generated? [Short form: *Adequate sequence generation*?]**	
Criteria for a judgement of 'YES' (i.e. low risk of bias).	The investigators describe a random component in the sequence generation process such as: • Referring to a random number table; • Using a computer random number generator; • Coin tossing; • Shuffling cards or envelopes; • Throwing dice; • Drawing of lots; • Minimization*. *Minimization may be implemented without a random element, and this is considered to be equivalent to being random.
Criteria for the judgement of 'NO' (i.e. high risk of bias).	The investigators describe a non-random component in the sequence generation process. Usually, the description would involve some systematic, non-random approach, for example: • Sequence generated by odd or even date of birth; • Sequence generated by some rule based on date (or day) of admission; • Sequence generated by some rule based on hospital or clinic record number. Other non-random approaches happen much less frequently than the systematic approaches mentioned above and tend to be obvious. They usually involve judgement or some method of non-random categorization of participants, for example: • Allocation by judgement of the clinician; • Allocation by preference of the participant; • Allocation based on the results of a laboratory test or a series of tests; • Allocation by availability of the intervention.
Criteria for the judgement of 'UNCLEAR' (uncertain risk of bias).	Insufficient information about the sequence generation process to permit judgement of 'Yes' or 'No'.
ALLOCATION CONCEALMENT **Was allocation adequately concealed? [Short form: *Allocation concealment*?]**	

Criteria for a judgement of 'YES' (i.e. low risk of bias).	Participants and investigators enrolling participants could not foresee assignment because one of the following, or an equivalent method, was used to conceal allocation: • Central allocation (including telephone, web-based, and pharmacy-controlled, randomization); • Sequentially numbered drug containers of identical appearance; • Sequentially numbered, opaque, sealed envelopes.
Criteria for the judgement of 'NO' (i.e. high risk of bias).	Participants or investigators enrolling participants could possibly foresee assignments and thus introduce selection bias, such as allocation based on: • Using an open random allocation schedule (e.g. a list of random numbers); • Assignment envelopes were used without appropriate safeguards (e.g. if envelopes were unsealed or nonopaque or not sequentially numbered); • Alternation or rotation; • Date of birth; • Case record number; • Any other explicitly unconcealed procedure.
Criteria for the judgement of 'UNCLEAR' (uncertain risk of bias).	Insufficient information to permit judgement of 'Yes' or 'No'. This is usually the case if the method of concealment is not described or not described in sufficient detail to allow a definite judgement – for example if the use of assignment envelopes is described, but it remains unclear whether envelopes were sequentially numbered, opaque and sealed.
BLINDING OF PARTICIPANTS, PERSONNEL AND OUTCOME ASSESSORS **Was knowledge of the allocated interventions adequately prevented during the study? [Short form: *Blinding*?]**	
Criteria for a judgement of 'YES' (i.e. low risk of bias).	Any one of the following: • No blinding, but the review authors judge that the outcome and the outcome measurement are not likely to be influenced by lack of blinding; • Blinding of participants and key study personnel ensured, and unlikely that the blinding could have been broken; • Either participants or some key study personnel were not blinded, but outcome assessment was blinded and the non-blinding of others unlikely to introduce bias.
Criteria for the judgement of 'NO' (i.e. high risk of bias).	Any one of the following: • No blinding or incomplete blinding, and the outcome or outcome measurement is likely to be influenced by lack of blinding; • Blinding of key study participants and personnel attempted, but likely that the blinding could have been broken; • Either participants or some key study personnel were not blinded, and the non-blinding of others likely to introduce bias.
Criteria for the judgement of 'UNCLEAR' (uncertain risk of bias).	Any one of the following: • Insufficient information to permit judgement of 'Yes' or 'No'; • The study did not address this outcome.

INCOMPLETE OUTCOME DATA **Were incomplete outcome data adequately addressed? [Short form: *Incomplete outcome data addressed*?]**	
Criteria for a judgement of 'YES' (i.e. low risk of bias).	Any one of the following: • No missing outcome data; • Reasons for missing outcome data unlikely to be related to true outcome (for survival data, censoring unlikely to be introducing bias); • Missing outcome data balanced in numbers across intervention groups, with similar reasons for missing data across groups; • For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk not enough to have a clinically relevant impact on the intervention effect estimate; • For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes not enough to have a clinically relevant impact on observed effect size; • Missing data have been imputed using appropriate methods.
Criteria for the judgement of 'NO' (i.e. high risk of bias).	Any one of the following: • Reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups; • For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce clinically relevant bias in intervention effect estimate; • For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes enough to induce clinically relevant bias in observed effect size; • 'As-treated' analysis done with substantial departure of the intervention received from that assigned at randomization; • Potentially inappropriate application of simple imputation.
Criteria for the judgement of 'UNCLEAR' (uncertain risk of bias).	Any one of the following: • Insufficient reporting of attrition/exclusions to permit judgement of 'Yes' or 'No' (e.g. number randomized not stated, no reasons for missing data provided); • The study did not address this outcome.
SELECTIVE OUTCOME REPORTING **Are reports of the study free of suggestion of selective outcome reporting? [Short form: *Free of selective reporting*?]**	
Criteria for a judgement of 'YES' (i.e. low risk of bias).	Any of the following: • The study protocol is available and all of the study's pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way; • The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified (convincing text of this nature may be uncommon).

Criteria for the judgement of 'NO' (i.e. high risk of bias).	Any one of the following: • Not all of the study's pre-specified primary outcomes have been reported; • One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not pre-specified; • One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected adverse effect); • One or more outcomes of interest in the review are reported incompletely so that they cannot be entered in a meta-analysis; • The study report fails to include results for a key outcome that would be expected to have been reported for such a study.
Criteria for the judgement of 'UNCLEAR' (uncertain risk of bias).	Insufficient information to permit judgement of 'Yes' or 'No'. It is likely that the majority of studies will fall into this category.
OTHER POTENTIAL THREATS TO VALIDITY **Was the study apparently free of other problems that could put it at a risk of bias? [Short form: *Free of other bias*?]**	
Criteria for a judgement of 'YES' (i.e. low risk of bias).	The study appears to be free of other sources of bias.
Criteria for the judgement of 'NO' (i.e. high risk of bias).	There is at least one important risk of bias. For example, the study: • Had a potential source of bias related to the specific study design used; or • Stopped early due to some data-dependent process (including a formal-stopping rule); or • Had extreme baseline imbalance; or • Has been claimed to have been fraudulent; or • Had some other problem.
Criteria for the judgement of 'UNCLEAR' (uncertain risk of bias).	There may be a risk of bias, but there is either: • Insufficient information to assess whether an important risk of bias exists; or • Insufficient rationale or evidence that an identified problem will introduce bias.

Appendix 6: PRISMA 2009[180] Flow Diagram (adapted) for the three reviews of clinical evidence, including the update search for Niox Vero

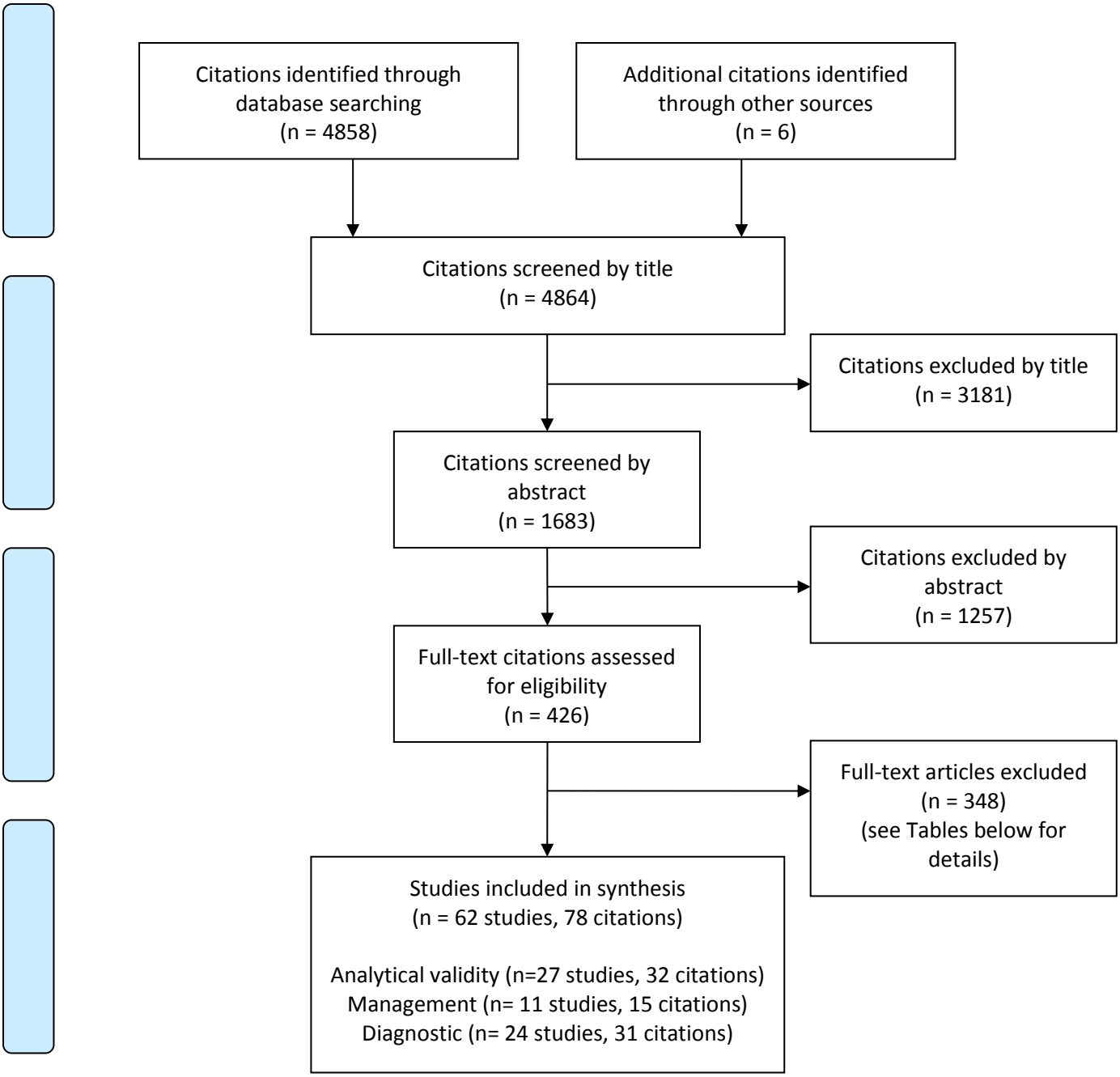


Table 86: Analytical Validity review – table of excluded studies with rationale

	Ref ID	Study	Reason for exclusion
1.	2649	Awabdy 2010	Wrong device
2.	619	Baraldi 1999	Tidal breathing
3.	6459	Baraldi 2000	NO analytical data
4.	5679	Becher 2010	NO analytical data
5.	7988	Becher 2010	NO analytical data
6.	2033	Boon 2010	Online versus offline
7.	6542	Boot 2007	Nasal NO
8.	from bedfont	Borrill 2006	Wrong device
9.	4893	Brooks 2011	Wrong device
10.	6588	Buchvald 2005	Measurement in healthy subjects
11.	6608	Byrnes 1997	No analytical data
12.	620	Canady 1999	Online versus offline
13.	2477	Castano 2011	No analytical data
14.	3083	Chai 2011	No analytical data
15.	7991	Cristescu 2013	Review
16.	7993	Fortuna 2006	Foreign language
17.	5001	Fuchs 2012	Single breath versus multiple breaths
18.	6887	Gill 2006	Inter reliability
19.	2428	Hemmingsson 2009	Not in humans
20.	5080	Ito Y 2010	NIOX MINO 6s versus 10s
21.	7039	Jobsis 2001	Offline
22.	7038	Jobsis 2001	Offline
23.	5931	Jung 2012	Nasal NO
24.	2938	Jung 2011	No analytical data
25.	5972	Koopman 2009	6s versus 10s in Niox
26.	537	Kumor 2004	Foreign language
27.	2839	Lee 2011	Repeatability
28.	6006	Lehtimaki 2000	Device not in scope
29.	6019	Linn WS 2009	Wrong device
30.	5193	Magori 2011	Device not in scope
31.	7261	Malik 2005	No comparison between devices
32.	494	Malik 2007	Wrong device
33.	347	Maniscalco 2010	Review
34.	485	McCurdy	Laser spectroscopy
35.	7299	McCurdy 2006	Laser spectroscopy
36.	7997	Montella 2011	Nasal NO
37.	530	Muller 2005	Wrong device
38.	5257	Munnik 2010	Not comparing relevant devices
39.	5281	Olaguibel 2011	Healthy volunteers
40.	574	Pijnenburg 2002	Online versus offline
41.	626	Rutgers 1998	Single breath v tidal breath

42.	738	Saito 2010	Foreign language
43.	7999	Sardon 2007	Foreign language
44.	7622	Sardon 2008	Foreign language
45.	8000	Sardon 2008	6s versus 10s
46.	379	Selby 2010	Repeatability
47.	5472	Taylor 2011	Reproducibility and long term performance in NIOX MINO only.
48.	2843	Thijs 2010	No comparison between devices
49.	3032	Tsuburai 2010	Offline
50.	5491	Tsuburai 2010	Foreign language
51.	7832	Turner 2006	Unable to obtain

Table 87: Management review – table of excluded studies with rationale

	Ref ID	Study	Reason for exclusion
1.	4800	Abba 2009	Unable to obtain
2.	2915	Acembekiroglu 2011	Not RCT study
3.	5595	Agache 2012	Not RCT study
4.	4821	Anderson 2012	Not RCT study
5.	5660	Ayars 2012	Not RCT study
6.	1963	Badzakova 2011	Not RCT study
7.	5663	Baek 2010	Not RCT study
8.	6455	Baptist 2008	Emergency care
9.	6457	Baraldi 1997	Not RCT study
10.	4850	Barreto 2009	Not RCT study
11.	4856	Bautista 2011	Not asthma
12.	4859	Beigelman 2009	Off line
13.	6492	Belda 2006	Not RCT study
14.	6521	Bisgaard 1999	Wrong flow rate
15.	6530	Bodini 2006	Not RCT study
16.	6531	Bodini 2007	Not RCT study
17.	4874	Bora 2011	Not RCT study
18.	5699	Bosque-Garcia 2011	Not RCT study
19.	5701	Bossley 2009	Not RCT study
20.	6561	Bratton 1999	Wrong flow rate
21.	6567	Brightling 2005	Review
22.	6570	Brindicci 2007	FeNO did not guide step up/down therapy
23.	6578	Bruce 2006	FeNO did not guide step up/down therapy
24.	5708	Bruce 2010	FeNO did not guide step up/down therapy
25.	6587	Buchvald 2003	FeNO did not guide step up/down therapy
26.	424	Bukstein 2011	Not RCT study
27.	2244	Bukstein 2012	Not RCT study
28.	6602	Bush 2005	Trial protocol
29.	403	Cabral 2009	Not RCT study
30.	6626	Carra S 2001	Wrong flow rate
31.	5723	Carter 2010	Not RCT study
32.	2327	Carvalho-Pinto 2010	Not RCT study
33.	5724	Castell 2011	Not RCT study
34.	4912	Chen 2010	Not RCT study
35.	4916	Chinellato 2012	Not RCT study
36.	4927	Clearie 2011	Not RCT study
37.	6399	clinical trial 2005	Trial protocol
38.	6684	Cohen 2008	Alveolar NO
39.	600	Colon-Semidey 2000	Not RCT study
40.	6697	Covar 2003	Not RCT study
41.	695	Cowan 2010	FeNO did not guide step up/down therapy
42.	4940	Cowan 2010	Not randomised to FeNO
43.	6700	Craig 2008	FeNO did not guide step up/down therapy

44.	6705	Crane 2002	Review
45.	5767	Crater 1999	Wrong flow rate
46.	264	Currie 2003	FeNO did not guide step up/down therapy
47.	563	Currie 2003	Not RCT study
48.	263	Currie 2003	Wrong flow rate
49.	4945	Dahlen 2009	FeNO did not guide step up/down therapy
50.	262	Dal Negro 2003	Wrong flow rate
51.	6732	de Gouw 1998	FeNO did not guide step up/down therapy
52.	6733	de Gouw 1998	Not RCT study
53.	6735	de Gouw 2001	Wrong flow rate
54.	4949	de Groot 2012	Not RCT study
55.	5778	de Jongste 2009	Daily monitoring
56.	6744	de KJ 2002	FeNO did not guide step up/down therapy
57.	4951	Debley 2012	Not RCT study
58.	6754	Delclaux 2008	Off line
59.	5790	Delgado-Corcoran 2004	Not RCT study
60.	6775	Diamant 2005	FeNO did not guide step up/down therapy
61.	2596	Domingo 2011	Not RCT study
62.	5807	Dressel 2007	Not RCT study
63.	4964	Dressel 2009	Not RCT study
64.	690	Dweik 2010	Not RCT study
65.	6948	Haldar 2008	Not RCT study
66.	318	Honkoop 2011	Ongoing study
67.	2128	Imaoka 2011	Not RCT study
68.	7085	Kelso 2006	Commentary
69.	7099	Kharitonov 2002	FeNO did not guide step up/down therapy
70.	702	Krcmova 2009	Not RCT study
71.	7215	Lonnkvist 2004	FeNO did not guide step up/down therapy
72.	345	McKinlay 2011	FeNO did not guide step up/down therapy
73.	444	Menzies 2008	FeNO did not guide step up/down therapy
74.	3014	Monforte 2010	Not RCT study
75.	376	Perez-de-Llano 2010	FeNO did not guide step up/down therapy
76.	7546	Rees 2006	Commentary
77.	3105	Ryan D 2013	Not RCT study
78.	456	Sordillo 2011	FeNO did not guide step up/down therapy
79.	6285	Syed 2011	No relevant outcomes
80.	551	Taylor 2004	FeNO did not guide step up/down therapy
81.	2389	Wanich 2009	Commentary

Table 88: Diagnostic review – table of excluded studies with rationale

	Ref ID	Study	Reason for exclusion
1.	5602	Alvarez-Gutierrez 2010	Foreign language
2.	5614	Andregnette 2011	No useable diagnostic data
3.	308	Arnold 2012	Emergency care
4.	726	Arochena 2012	No useable diagnostic data
5.	6436	Artlich 1996	Wrong flow rate
6.	5659	Avital 2001	Offline
7.	5668	Balinotti 2011	Infants
8.	6454	Baptist 2008	Emergency care
9.	6456	Baptist 2008	Emergency care
10.	4847	Barben 2013	Not FeNO for diagnosis
11.	5674	Barreto 2011	No useable diagnostic data
12.	2099	Bar-Yishay 2010	No useable diagnostic data
13.	493	Bastain 2011	Offline
14.	5677	Bayo 2008	Population asthma and non-asthma
15.	5681	Bell 2013	No useable diagnostic data
16.	1328	Berkman 2005	Wrong flow rate
17.	601	Berlyne 2000	Wrong flow rate
18.	4865	Bernstein 2009	No useable diagnostic data
19.	519	Berry 2005	Wrong flow rate
20.	2002	Bivins 2013	No useable diagnostic data
21.	5692	Blain 2009	Case study
22.	589	Bloemen 2010	No useable diagnostic data
23.	502	Bohadana 2011	Unselected population
24.	5694	Bommarito 2008	Offline
25.	5697	Boon 2012	Case control
26.	2488	Bozek 2010	No useable diagnostic data
27.	2487	Bozek 2010	Foreign language
28.	4882	Bozek 2011	Nasal NO
29.	5709	Brusselle 2010	No useable diagnostic data
30.	5712	Buchvald 2005	All asthmatic not diagnostic
31.	3099	Burnett 2011	No useable diagnostic data
32.	5721	Cardinale 2005	Wrong flow rate
33.	116	Carlstedt 2011	No useable diagnostic data
34.	4905	Castro-Rodriguez 2013	Age under 5
35.	5725	Caudri 2010	Offline
36.	5732	Chatkin 1999	Wrong flow rate
37.	655	Chawes 2010	Wrong flow rate
38.	114	Cherot-Kornobis 2011	No useable diagnostic data
39.	81	Chladkova 2012	Alveolar NO
40.	2814	Choi 2009	No useable diagnostic data
41.	352	Choi 2011	Not FeNO for diagnosis
42.	5741	Chow 2009	No useable diagnostic data
43.	365	Cibella 2011	No useable diagnostic data

44.	351	Ciprandi 2010	Data for both asthma and rhinitis
45.	5749	Ciprandi 2013	Data for both asthma and rhinitis
46.	5750	Cirillo 2013	Not diagnosis of asthma
47.	380	Clearie 2010	Not FeNO for diagnosis
48.	5754	Clearie 2010	No useable diagnostic data
49.	5755	Cleveland 2013	No useable diagnostic data
50.	5759	Columbo 2012	No useable diagnostic data
51.	4931	Consilvio 2010	Population obese
52.	3090	Consilvio 2010	No useable diagnostic data
53.	5763	Corradi 2002	Editorial
54.	6692	Corradi 2007	Review
55.	295	Crane 2012	No useable diagnostic data
56.	3412	Crothall 2012	No useable diagnostic data
57.	5771	Dallinga 2010	No useable diagnostic data
58.	5773	David 2007	Unselected population
59.	4947	de Bot 2013	No useable diagnostic data
60.	5781	de Meer 2005	Offline
61.	4698	de Winter-de Groot 2009	Nasal NO
62.	374	Debley 2010	Infants under 2 years
63.	5783	Debley 2011	Infants
64.	312	Decimo 2011	No useable diagnostic data
65.	5788	del Giudice 2004	Unselected population
66.	584	Delclaux 2002	Case control
67.	394	Demange 2009	Unselected population
68.	5791	Demange 2010	Unselected population
69.	4956	Dente 2010	No useable diagnostic data
70.	579	Deykin 2002	Wrong flow rate
71.	531	Deykin 2005	Offline
72.	5798	Diaconu 2010	No useable diagnostic data
73.	2010	Dichiaro 2009	No useable diagnostic data
74.	2059	Dichiaro 2010	No useable diagnostic data
75.	2189	Divjan 2009	No useable diagnostic data
76.	2040	Divjan 2011	No useable diagnostic data
77.	4961	Donohue 2013	Not feno for diagnosis
78.	570	Dupont 2003	Wrong flow rate
79.	364	Ekerljung 2011	No useable diagnostic data
80.	4471	Fernandez-Nieto 2009	No data on feno for diagnosis
81.	4987	Fireman 2010	No useable diagnostic data
82.	557	Franklin 2003	Wrong flow rate
83.	5838	Fujimura 2008	Case control
84.	5867	Grzelewski 2012	Diagnosing EIB in asthmatics - i.e. not diagnostic of asthma
85.	5043	Hafkamp-de Groen 2010	Study design only
86.	3087	Hardaker 2010	No useable diagnostic data
87.	435	Hardaker 2011	No useable diagnostic data

88.	332	Hogman 2011	No useable diagnostic data
89.	5901	Huang 2011	Unselected population
90.	2398	Hur 2013	No useable diagnostic data
91.	2128	Imaoka 2011	No useable diagnostic data
92.	2615	Imaoka 2011	No useable diagnostic data
93.	5915	Inoue 2010	No useable diagnostic data
94.	5917	Ishizuka 2011	No useable diagnostic data
95.	5081	Jackson 2009	No useable diagnostic data
96.	7038	Jobsis 2001	Offline
97.	7039	Jobsis 2001	No useable diagnostic data
98.	5955	Khurana 2011	No useable diagnostic data
99.	2409	Kim 2009	No useable diagnostic data
100.	2423	Kim 2009	No useable diagnostic data
101.	5962	Kim 2013	No useable diagnostic data
102.	5122	Klaassen 2012	No useable diagnostic data
103.	2537	Konstantinou 2009	FeNO in exacerbations
104.	5129	Konstantinou 2013	No useable diagnostic data
105.	5136	Kotaniemi-Syrjanen 2013	infants <3yrs
106.	5153	Larj 2010	Alveolar NO
107.	5997	Larson 2011	Alveolar NO
108.	5998	Larson 2011	No useable diagnostic data
109.	7161	Latzin 2006	No useable diagnostic data
110.	387	Lemiere 2010	Offline
111.	2746	Lemiere 2011	Not FeNO
112.	6014	Lex 2007	Diagnosis of EIB not asthma
113.	698	Li 2010	Foreign language
114.	2229	Linkosalo 2009	No useable diagnostic data
115.	402	Linn 2009	No useable diagnostic data
116.	6018	Linn 2009	No useable diagnostic data
117.	613	Little 2000	Wrong flow rate
118.	5188	Lund 2009	Case control study
119.	330	Mahut 2011	No useable diagnostic data
120.	5202	Malby Schoos 2012	No useable diagnostic data
121.	6044	Malinovschi 2009	No useable diagnostic data
122.	6045	Malinovschi 2012	Incorrect reference standard
123.	3237	Malka-Rais 2010	No useable diagnostic data
124.	5206	Malka-Rais 2010	No useable diagnostic data
125.	6047	Malmberg 2009	Age 3-7
126.	5208	Malmberg 2010	No useable diagnostic data
127.	7283	Martin 2007	No useable diagnostic data
128.	667	Martin 2009	No useable diagnostic data
129.	659	Martin 2010	No useable diagnostic data
130.	5220	Martin 2010	No useable diagnostic data
131.	5221	Martin 2012	No useable diagnostic data

132.	7285	Martins 2008	population not self-presenting for assessment
133.	6065	Matsunaga 2011	Case control
134.	7320	Meyts 2003	No useable diagnostic data
135.	6084	Mgaloblishvili 2009	No useable diagnostic data
136.	6085	Mgaloblishvili 2010	No useable diagnostic data
137.	6086	Mi 2011	No useable diagnostic data
138.	4639	Monforte 2009	No useable diagnostic data
139.	3014	Monforte 2010	No useable diagnostic data
140.	409	Motomura 2009	No useable diagnostic data
141.	6104	Motomura 2010	No useable diagnostic data
142.	6105	Motomura 2012	No useable diagnostic data
143.	5257	Munnik 2010	No useable diagnostic data
144.	7375	Murata 2007	No useable diagnostic data
145.	2640	Musk 2010	No useable diagnostic data
146.	5260	Musk 2011	No useable diagnostic data
147.	6108	Nagase 2011	Unclear how patients recruited
148.	5264	Nakajima 2011	No useable diagnostic data
149.	6110	Narang 2002	Case control study
150.	633	Nelson 1997	Case control study
151.	2479	Nikasinovic 2011	No useable diagnostic data
152.	482	Nishio 2007	No useable diagnostic data
153.	617	Obata 1999	Wrong flow rate
154.	2318	Oros 2010	No useable diagnostic data
155.	6126	Oshikata 2008	Foreign language
156.	2036	Perzanowski 2009	No useable diagnostic data
157.	2642	Perzanowski 2010	No useable diagnostic data
158.	7500	Porsbjerg 2008	Diagnosis of airway hyper-responsiveness
159.	6164	Porsbjerg 2009	Population all asthmatics
160.	3011	Porsbjerg 2010	Unselected population
161.	7503	Prasad 2006	No useable diagnostic data
162.	2492	Profita 2009	No useable diagnostic data
163.	5335	Puckett 2009	No useable diagnostic data
164.	5350	Raulf-Heimsoth 2010	No useable diagnostic data
165.	6190	Raulf-Heimsoth 2013	No useable diagnostic data
166.	2821	Reyes 2011	Wrong flow rate
167.	478	Robroeks 2007	Case control study
168.	2081	Rosa 2009	No useable diagnostic data
169.	6204	Rosa 2011	Offline measurement
170.	626	Rutgers 1998	case control study
171.	3105	Ryan 2013	Not diagnostic study
172.	6213	Sachs-Olsen 2010	Unselected population
173.	602	Sakai 2010	No useable diagnostic data
174.	6219	Sanchez-Vidaurre 2012	No data on FeNO
175.	350	Schleich 2010	Population all asthmatic

176.	612	Scollo 2000	Case control study
177.	6257	Silkoff 2005	Diagnosis of EOS+ phenotype
178.	2000	Simpson 2010	No useable diagnostic data
179.	5435	Smith 2012	No useable diagnostic data
180.	2272	Sobrevia 2010	No useable diagnostic data
181.	6267	Sordillo 2009	No useable diagnostic data
182.	654	Stahl 2009	Unselected population
183.	661	Sverrild 2010	Unselected population
184.	1109	Sverrild 2013	Unselected population
185.	6294	Tanaka 2011	No useable diagnostic data
186.	2102	Tatyana 2011	Age (under 5)
187.	2713	Taylor 2011	No useable diagnostic data
188.	7804	Terada 1999	Foreign language
189.	525	Thomas 2005	Unselected population
190.	5485	Tossa 2009	Study design only
191.	483	Travers 2007	Unselected population
192.	6321	Tseliou 2010	Population - severe refractory asthma
193.	495	Tworek 2006	Foreign language
194.	6332	van Amsterdam 2003	Wrong flow rate
195.	728	van de Kant 2011	Age (under 5)
196.	5500	van der Valk 2012	No useable diagnostic data
197.	6339	van Wonderen 2009	Study design only
198.	6344	Vieira 2009	No useable diagnostic data
199.	6345	Vieira 2011	population - children with positive skin prick test
200.	5512	Vitruba 2009	Foreign language
201.	2101	Votruba 2011	Not valid reference standard (biopsy results)
202.	6356	Wang 2009	No useable diagnostic data
203.	582	Warke 2002	Population is mix of asthmatics and healthy
204.	5523	Wedes 2009	Case control study
205.	7920	Wildhaber 1999	No useable diagnostic data
206.	7921	Wildhaber 2000	No useable diagnostic data
207.	6371	Yang 2011	Not FeNO
208.	6372	Yang 2011	Case control
209.	339	Yao 2011	Unselected population
210.	6377	Yawn 2013	Non-specific respiratory symptoms
211.	2827	Yoo 2012	No useable diagnostic data
212.	2830	Yoo 2013	No useable diagnostic data
213.	5553	Zhang 2011	Foreign language
214.	6384	Zhang 2011	Foreign language
215.	565	Zietkowski 2010	Case control study
216.	2776	Zietkowski 2010	No useable diagnostic data

Appendix 7: Sub Group PRISMA Flow Diagram

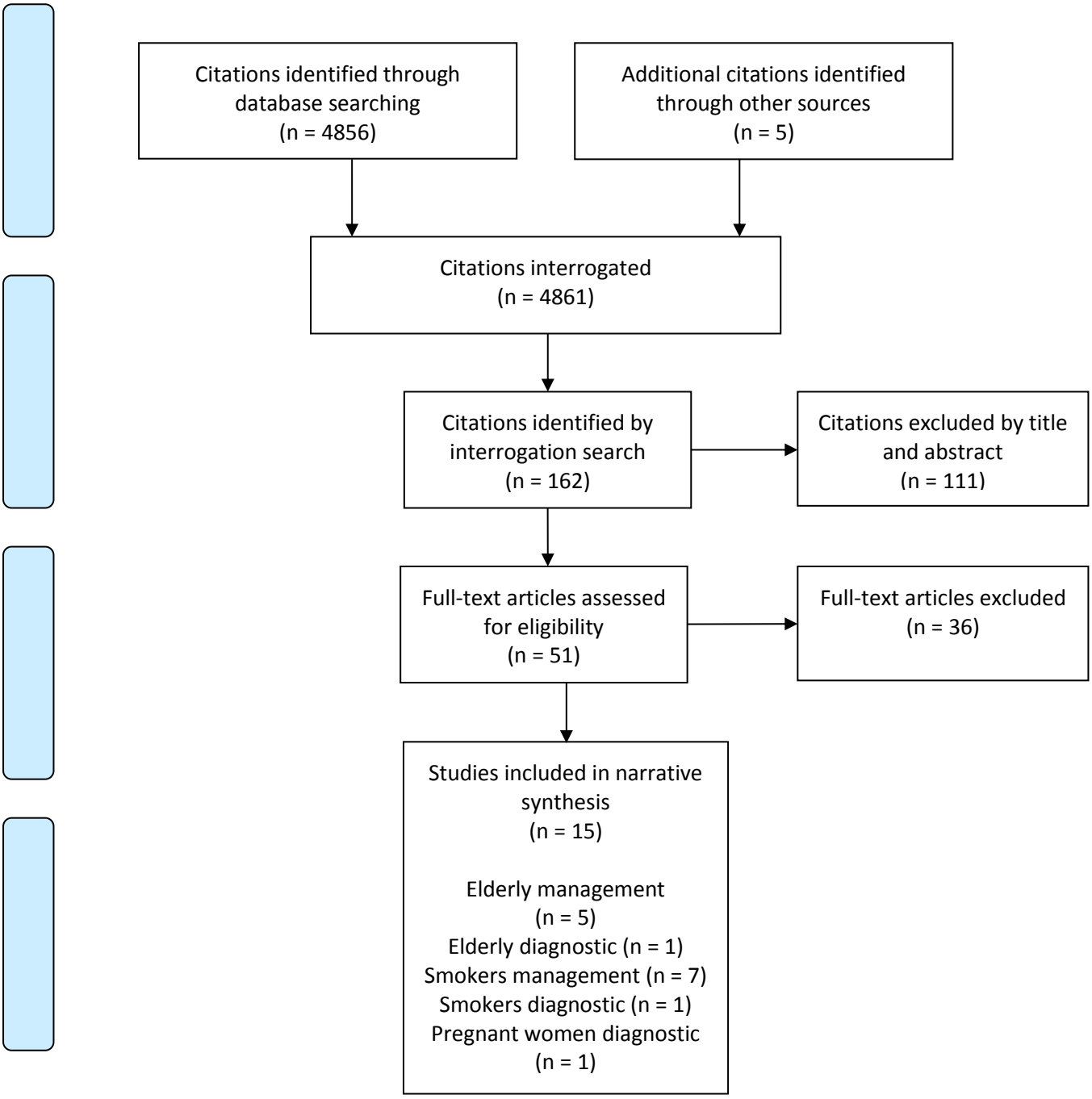


Table 89: Smokers management review – table of excluded studies with rationale

	Ref ID	Study	Reason for exclusion
1.	534	Baur 2005	Non asthmatics
2.	519	Berry 2005	Wrong flow rate
3.	2002	Bivins 2013	No data smoker v non-smoker
4.	502	Bohadana 2011	Not all asthmatics
5.	5694	Bommarito 2008	Diagnostic - not all asthmatic
6.	5780	de la Barra 2011	No data smoker v non-smoker
7.	6777	Dinakar 2005	Offline
8.	2962	Gaku 2010	Mean FeNO only
9.	2071	Gemicioglu 2009	Mean FeNO only
10.	3463	Gouvis-Echraghi 2012	Preschool
11.	331	Hillas 2011	Mean FeNO only
12.	462	Kostikas 2008	Diagnostic AUCs reported for smokers v non-smokers
13.	6006	Lehtimaki 2000	Wrong flow rate
14.	701	Mahut 2010	Adults mean FeNO only
15.	6044	Malinovschi 2009	Diagnostic adults
16.	6066	Matsunaga 2011	Diagnostic adults
17.	6065	Matsunaga 2011	Diagnostic adults
18.	5257	Munnik 2010	Diagnostic in adults, correction for smoking only
19.	4242	Nadif 2010	Mean FeNO only
20.	378	Nadif 2010	Mean FeNO only
21.	6126	Oshikata 2008	Diagnostic adults
22.	7475	Persson 1994	Healthy only
23.	2642	Perzanowski 2010	Wheeze not asthma
24.	5377	Rouhos 2010	Mean FeNO only
25.	626	Rutgers 1998	Mean FeNO only
26.	2269	Shimoda 2013	Mean FeNO only
27.	5444	Spears 2011	Wrong flow rate
28.	2713	Taylor 2011	No data smoker v non-smoker

Table 90: Elderly management and diagnostic review – table of excluded studies with rationale

	Ref ID	Study	Reason for exclusion
1.	4882	Bozek 2011	Population not asthmatic
2.	435	Hardaker 2011	Wrong flow rate
3.	3087	Hardaker 2010	Wrong flow rate
4.	5931	Jung 2012	Nasal NO

Table 91: Pregnancy diagnostic review – table of excluded studies with rationale

	Ref ID	Study	Reason for exclusion
1.	5029	Gibson 2011	Management study
2.	6003	Leblanc 2009	Not diagnostic study
3.	6070	McCallister 2013	Review
4.	5327	Powell 2011	Management study

Appendix 8: Table of study characteristics for non-relevant adult diagnostics

Study author, year	Study design, Funding	Country, setting, recruitment dates	Population	Device	Reference standard	n analysed/N recruited, reasons for withdrawals	Mean age in years (SD)	Gender, n male (%)	Severity FEV_1% or FEV_1/FVC	Mean FeNO	Smokers	Atopic
Fortuna, 2007[76]	Prospective, consecutive Funding NR	Spain, Secondary care, (outpatient clinic) Dates: October 2004 to November 2005	Symptoms suggestive of asthma (Position A)	SIR N-6008, Madrid, Spain	Lung function tests (spirometry and bronchodilator response) and methacholine challenge test following guidelines of the Global Initiative for Asthma (GINA)	57 recruited, 50 analysed N=7 receiving oral cortcosteroid treatment at the time of study.	Asthmatic: 37 (range: 18-68); Non-asthmatic: 38 (range: 18-64)	21/50 (42%)	Asthmatic: mean 94±19; Non-asthmatic: mean 99±10	Asthmatic: 40±31; non-asthmatic: 18±23	Asthmatic: n=3 smokers, n=4 ex-smokers; Non-asthmatic: n=4 smokers, n=3 ex-smokers.	NR
Fukuhara, 2011[53]	Prospecive Funding NR; authors reported no conflict of interest	Japan, Secondary care, Dates: May 2007 to June 2007	Symptoms suggestive of asthma (Position A)	NA623 (Chest MI, Tokyo, Japan)	1) at least 1 of the subjective symptoms of recurrent cough,wheezing, or dyspnea. 2) At least 2 of the 3 criteria of induced sputum eosinophilia, airway hyperresponsiveness, and reversible airway obstruction 3) exclusion of other lung diseases.	97 recruited, 61 analysed N=36 unable to complete all tests	55.6 (range: 17-81)	31/61 (50.8%)	96.1% (95% Cl, 90.1-102.0)	74.5 ppb (95% Cl, 56.2-92.8)	Current smokers n=6; ex-smokers n=13; non-smokers n=42	14/61 (23%)
Pizzimenti 2009[94]	prospective, consective Funding source NR	Italy Secondary care (outpatient clinic)	Patients with chronic cough (Position A)	NIOX MINO	Methacoline challenge (PD20 FEV-1 < 800 µ)	156 recruited, 156 analysed	NR	64/156 (41.0%)	NR	34.1 ppb (95% Cl: 28.5-39.5 ppb)	n=14/156 (9%)	74/156 (47.4%)

Study author, year	Study design, Funding	Country, setting, recruitment dates	Population	Device	Reference standard	n analysed/N recruited, reasons for withdrawals	Mean age in years (SD)	Gender, n male (%)	Severity FEV_1% or FEV_1/FVC	Mean FeNO	Smokers	Atopic
		Dates NR										
Mathew 2011[95] Sato 2008[81]	Prospective Funding NR	UK Secondary care Dates NR	Difficult to diagnose (Position B)	NR	Methcacholine Challenge Test	84 recruited, 84 analysed	NR	36/84 (42.9%)	NR	NR	NR	NR
Zhang 2011[93]	Prospective Funding NR	China, Secondary care. October 2009 to September 2010	Chronic cough with normal chest radiographs (Position A)	NIOX MINO	Diagnosis of CVA, EB and Other based on sputum cell counts, pulmonary function test, BHR, 24-h esophageal pH monitoring, SPT and serum IgE	106 recruited 106 analysed	NR	NR	NR	NR	NR	NR
Arora 2006[85]	Prospective, Funding: United States Air Force Surgeon General's Office	US Specialist care Dates: NR	Adult (military rtecruits) with symptoms suggestive of asthma (Position A)	Niox (Aerocrine AB, Stockholm, Sweden)	Included both a consistent history with recurrent respiratory symptoms of nonproductive cough, shortness of breath, chest tightness, or wheezing with exertion or at rest and positive histamine bronchoprovocation (20% fall in FEV_1 (PC20) of ≤8 mg/ml of histamine)	172 recruited 172 analysed	Asthmatic mean age 20 ± 2.7; Non-asthmatic 21 ± 2.7	85 (49.4%)	Asthmic 98±13 (69-133); Non asthmatic 107±14 (89-135)	Asthmatic ; 30ppb ±31; Non Asthmatic ; 19ppb ±11	0/138 (0%)	NR

Appendix 9: Table of highest sum of sensitivity and specificity, highest sensitivity and highest specificity for non-relevant studies

				Highest sum of sens and spec					Rule-out Sens					Rule-in Spec				
Study author, year	**Population**	**Device**	**Reference standard**	**cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**cut-off**	**Sens**	**Spec**	**PPV**	**NPV**
Position A versus whole pathway																		
Fortuna, 2007[76]	Adults Position A	SIR N-6008, Madrid, Spain	Lung function tests (spirometry and bronchodilator response) and methacholine challenge test following guidelines of the Global Initiative for Asthma (GINA)	≥20	77	64	62.96	78.26	-	-	-	-	-	-	-	-	-	-
Fukuhara, 2011[53]	Adults Position A	NA623 (Chest MI, Tokyo, Japan)	1) at least 1 of the subjective symptoms of recurrent cough,wheezing, or dyspnea. 2) At least 2 of the 3 criteria of induced sputum eosinophilia, airway hyperresponsiveness, and reversible airway obstruction 3) exclusion of other lung diseases.	40	78.6	89.5	94.28	65.38	-	-	-	-	-	-	-	-	-	-
Pizzimenti 2009[94]	Unspecified age group Position A	NIOX MINO	Methacoline challenge (PD20 FEV-1 < 800 µ)	55	10	67.2	39.28	97.66	-	-	-	-	-	-	-	-	-	-
Zhang 2011[93]	Unspecified age group Position A	NIOX MINO	Diagnosis of CVA, EB and Other based on sputum cell counts, pulmonary function test, BHR, 24-h esophageal pH monitoring, SPT and serum IgE	40	75%	86%	76.31	85.29	-	-	-	-	-	-	-	-	-	-
Arora 2006[85]	Adults	Niox	Included both a consistent history with	>17	63	58.8	86.14	28.17	>6	96.4	0	79.64	0.00	>46	16.7	100	100	22.81

				Highest sum of sens and spec					Rule-out Sens					Rule-in Spec				
Study author, year	**Population**	**Device**	**Reference standard**	**cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**cut-off**	**Sens**	**Spec**	**PPV**	**NPV**	**cut-off**	**Sens**	**Spec**	**PPV**	**NPV**
	Position A		recurrent respiratory symptoms of nonproductive cough, shortness of breath, chest tightness, or wheezing with exertion or at rest and positive histamine bronchoprovocation (20% fall in FEV_1 (PC20) of ≤8 mg/ml of histamine)															
Position B																		
Mathew 2011[95] Sato 2008[81]	Unspecified age group Position B	NR	Airway reversibility, positive response to ICS, airway hyper-responsiveness (MCT).	NR	NR	NR	8.7	70.49	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

Appendix 10: Table detailing the reference standards used in relevant adult diagnostic studies

Study author, year	Details of reference standard	Summarised as
Position A versus whole pathway		Position A versus whole pathway
Schneider 2013[75]	Measurements including spirometry were performed according to standard protocols (American Association for Respiratory Care, 2001) and reference values were adapted to sex, age, and height. Patients with FEV_1 < 80% predicted received salbutamol with an additional WBP investigation 20 min later. An obstructive airway disease was diagnosed if FEV_1/VC was ≤0.70. It was classified as asthma if clinical symptoms and history fitted and the change in FEV_1 was ≥12% compared to baseline and ≥200 mL and lung function returned to the predicted normal range. An incomplete bronchodilator response was stated if the response was ≤12% compared to baseline and ≤200 mL and lung volumes remained below predicted. It was classified as COPD, if clinical symptoms and history fitted and the bronchodilator response of FEV_1 after salbutamol was <12% compared to baseline and <200 mL. If there was no bronchial obstruction, bronchial provocation was performed to determine bronchial hyperresponsiveness (BHR) to methacholine according to the 1- concentration-4-step dosimeter protocol.25 This yields similar results as the ATS multi-concentration protocol but offers advantages in clinical practice. An "asthma" diagnosis required a 20% fall in FEV_1 from baseline after inhaling methacholine stepwise until the maximum concentration (16 mg/mL), alternatively a doubling of airway resistance (Raw) and its increase to ≥2.0 kPa*s. The responsible pneumologist was blinded to the FeNO results and made the diagnostic decision only on basis of medical history, physical examination, spirometry, WBP and bronchial provocation results.	FEV_1, FEV_1/FVC, airway reversibility, hyper-responsiveness (MCT)
Schneider 2009[77,78]	FEV_1/FVC <0.7 %/or FEV_1% <80%. If plus positive bronchodilator response = asthma. FEV/FVC>0.7 or FEV_1%>80%, plus +ve MTC = asthma	FEV_1, FEV_1/FVC, airway reversibility, airway hyper-responsiveness MCT
Smith 2005[87]	ATS 1987 diagnostic criteria, plus one or more of: 1. Positive response to bronchodilator (increase in FEV_1 of ≥12% from baseline 15mins after inhaled albuterol) 2.Positive response to ICS (increase in FEV_1 of ≥12%, or an increase in mean morning peak flow over previous 7 days of 15% or greater). 3. Positive test for airway hyperresponsiveness (defined as a provocative dose of methacholine,	Airway reversibility, positive response to ICS, airway hyper-responsiveness (MCT).

Study author, year	Details of reference standard	Summarised as
	resulting in a 20% reduction in FEV_1 of <8µmol.	
Smith 2004[90]	Relevant symptom history (ATS 1987 guidelines), AND a positive test for bronchial hyperresponsiveness (provocative dose of hypertonic saline resulting in a 15% fall in FEV_1 of less than 20ml) AND/OR a positive response to bronchodilator (increase in FEV_1 of 12% or greater from baseline 15 mins after inhaled albuterol).	Airway reversibility, positive response to ICS, airway hyper-responsiveness (MCT).
Position A versus airway reversibility		Position A versus airway reversibility
De La Barra 2011[88]	positive response to bronchodilator (increase in FEV_1 of 12% or greater from baseline 15 mins after inhaled albuterol)	Airway reversibility
Subset of Position A versus airway reversibility or airway hyper-responsiveness		Subset of Position A versus airway reversibility or airway hyper-responsiveness
Heffler 2006[86]	Asthma confirmed based on typical symptoms and > 12% improvement in FEV in response to salbutamol or methacholine PD20 FEV <800 µg	Airway hyper-responsiveness (MCT) or airway reversibility
Cordeiro 2011[91]	History of typical respiratory symptoms and FEV_1% improvement of >12% and >200 ml, or PC20 histamin of ≤8 mg/ml, according to GINA guidelines	Airway reversibility, airway hyper-responsiveness (histamine)
Difficult to diagnose versus airway hyper-responsiveness		Difficult to diagnose versus airway hyper-responsiveness
Schleich, 2012[83]	Asthma diagnosed based on airway hyperresponsiveness (via methacholine) provoking a 20% fall in FEV_1 of less than 16 mg⁄ml. Subjects were characterised as atopic if they had at least one positive skin prick test (wheal > 3 mm as compared with negative control) or specific IgE (> 0.35 KU⁄l; Phadia) for at least one common aeroallergen (cat, dog, house dust mites, grass pollen, tree pollen and a mixture of moulds).	Airway hyper-responsiveness (MCT)
Pedrosa 2010[89]	Consistent symptoms and a positive methacholine bronchial challenge (patients stopped asthma medication before test. Test performed according to ATS 1999 guidelines; considered positive when a decrease in FEV_1 from baseline of 20% or higher was obtained after MCh inhalation.)	Airway hyper-responsiveness (MCT)

Study author, year	Details of reference standard	Summarised as
Bobolea 2012[92]	Adenosine challenge test (PC20<400 mg/ml)	Adenosine challenge test
Suspected EIB versus Exercise challenge test		Suspected EIB versus Exercise challenge test
El Halawani 2003[84]	Exercise challenge test performed on a treadmill with incremental work rate. Up to 14 mins of symptom-limited exercise. Treadmill speed began @ 2miles/hour and increased 1 mile/hour every 2 min. Treadmill grade began @10%, increasing to 15% after 8 min.Targeted heart rate was 85% predicted maximum and maintained for 2 min. Spirometry performed every 5 min after exercise for total of 30 min. Pulmonary functioning discontinued when fall in FEV of 15% from baseline was demonstrated	Exercise challenge
Position F with chronic cough versus ICS responsiveness		Position F with chronic cough versus ICS responsiveness
Prieto 2009[82]	Responsiveness to FP was identified by a reduction of > 50% in the mean daily cough symptom scores during the 4 weeks of the FP trial compared with the baseline period.	ICS responsiveness
Hsu 2013[79]	complete improvement of cough upon ICS treatment with 250µg b.i.d. for at least 2 weeks	ICS responsiveness
Hahn 2007[80]	ICS responsiveness assessed 1 to 16 months after diagnostic tests.	ICS responsiveness

Appendix 11: Table detailing the inclusion and exclusion criteria of studies considered of most relevance to the review

Study author, year	Details of inclusion and exclusion criteria	Categorised as
Position A versus whole pathway		
Schneider 2013[75]	Patients presenting for the first time with symptoms such as dyspnoea, cough or phlegm for more than two months, leading to the clinical suspicion of obstructive or restrictive airway disease ('indicated population'). Patients were advised not to smoke at the day of investigation and not to use inhaler medication for twelve hours before lung function testing. Exclusions: Patients with respiratory tract infections within the last 6 weeks; previously established diagnosis of chronic obstructive airway disease; known contra-indications for bronchodilator reversibility testing or bronchial provocation, namely untreated hyperthyreosis, unstable coronary artery disease, and cardiac arrhythmia; pregnancy.	Position A
Schneider 2009[77,78]	Patients presenting to their GP for the first time with complaints suggestive of obstructive airway disease (OAD). Presentation of symptoms such as dyspnoea, coughing or expectoration for more than two months, thus leading to clinical suspicion of obstructive or restrictive airway disease as most important differential diagnoses ('indicated population'). GPs were advised to exclude patients with respiratory tract infections preceding the evaluation by 6 weeks. Patients with previously established diagnosis of OAD were excluded. Other exclusion criteria related to well known contra-indications for bronchodilator reversibility testing or bronchial provocation, namely untreated hyperthyreosis, unstable coronary artery disease, and cardiac arrhythmia. Pregnancy also led to exclusion.	Position A
Smith 2005[87]	Patients referred by GP for investigation of persistent, undiagnosed respiratory symptoms lasting at least 6 weeks. Exclusions: use of ICS or oral corticosteroids in the previous 4 weeks, respiratory tract infection in the previous 6 weeks, other established respiratory diagnosis, or significant comorbidity.	Position A

Study author, year	Details of inclusion and exclusion criteria	Categorised as
	Smokers were NOT excluded.	
Smith 2004[90]	Patients referred by GP for investigation of possible bronchial asthma with symptoms for a minimum of 6 weeks. No patient had been referred for specialist consultatiion. Exclusions: those who used oral or inhaled corticosteroids in previous 4 weeks, those with respiratory tract infection in previous 6 weeks.	Position A
Position A versus airway reversibility		
De La Barra 2011[88]	Inclusion: new undiagnosed symptoms of cough, wheeze, or dyspnoea of 6 weeks duration or longer	Position A
Subset of Position A versus airway reversibility or airway hyper-responsiveness		
Heffler 2006[86]	Nasal symptoms for >4 days / week over >8 weeks; asthma-like symptoms during past 2 months; Exclusion: use of steroids or any other anti-inflammatory drugs in last 2 months; current smoker (within past 12 months), previous diagnosis of asthma; respiratory infection within past 6 weeks	Position A
Cordeiro 2011[91]	All new patients who were referred to a general outpatient allergy clinic from January 2007 until September 2007. Patients using ICSs or oral corticosteriods within 6 weeks before the first visit were excluded from data analysis	Position A
Difficult to diagnose versus airway hyper-responsiveness		
Schleich, 2012[83]	Inclusion: Patients were addressed by their respiratory physician for a methacholine challenge to detect asthma. Subjects referred to methacholine challenge were those in whom the bronchodilating test failed	Difficult to diagnose

Study author, year	Details of inclusion and exclusion criteria	Categorised as
	to demonstrate reversible airways obstruction or those in whom baseline spirometric values were normal giving a low probability for a bronchodilating test to be significant. The patients studied here had either baseline $FEV_1 \geq 80\%$ predicted and FEV_1/FVC ratio $\geq 70\%$ or bronchodilation < 12% from baseline and 200 ml after 400 µg inhaled salbutamol in case of baseline FEV_1 was < 80% predicted or FEV_1/FVC ratio < 70%. Excluded: Patients already receiving inhaled corticosteroids.	
Pedrosa 2010[89]	Those reporting persistent symptoms consistent with asthma (shortness of breath, wheezing and/or cough) regardless of atopic status, who showed normal spirometry and had a negative bronchodilator test. Exclusions: as per ATS 1999 guidelines for bronchial challenge test.	Difficult to diagnose
Bobolea 2012[92]	Inclusion: Patient with suspected asthma who had normal spirometry, negative bronchodilator test, negative methacholine challenge (provocative concentration inducing a 20% fall in FEV_1 [PC20]>16 mg/ml)	Difficult to diagnose
Suspected EIB versus Exercise challenge test		
El Halawani 2003[84]	Inclusion: Patient with suspected asthma who had normal spirometry, negative bronchodilator test, negative methacholine challenge (provocative concentration inducing a 20% fall in FEV_1 [PC20]>16 mg/ml)	EIB
Position F with chronic cough versus ICS responsiveness		
Prieto 2009[82]	Chronic cough of at least 8 weeks with no evidence of any other lung disease, non-smokers, not currently being treated with angiotensin-converting enzyme inhibitors or b-blockers, ICS or oral corticosteroids or had respiratory tract infections in previous 4 weeks. FEV_1 of at least 80% predicted.	Difficult to diagnose with chronic cough

Study author, year	Details of inclusion and exclusion criteria	Categorised as
Hsu 2013[79]	Inclusion: patients with a history of chronic cough > 8 weeks and who didn't stop coughing after treatment for upper airway cough syndrome or gastro-oesophageal reflux disease. Exclusion: Obvious chest x-ray abnormalities, current smokers / smoking history of more than 10 pasck-years	Difficult to diagnose with chronic cough
Hahn 2007[80]	>18 years old, uncontrolled chronic cough (for >8 weeks), normal / nonlocalising chest radiograph, documented MCT results, and measurement of NO levels within 1 day of each other. Only patients who had started ICS therapy or who had their current ICS doeses altered were included. Current smokers and users of ACE inhibitors were excluded	Difficult to diagnose with chronic cough

Appendix 12: Table of results for all diagnostic studies in adults

Author, year	Prevalence of positive result by reference	FeNO cut-off ppb	True positive	False positive	False negative	True negative	Sensitivity asthma % (95% CI)	Specificity asthma % (95% ci)
Position A								
Schneider 2013[75]	154/393 (39.2%)	>9	146	209	7	31	96 (91, 98)	13 (9, 18)
		>12	135	167	23	68	85 (79, 90)	29 (23, 35)
		>16	105	128	46	114	70 (62, 76)	47 (41, 54)
		>20	91	89	62	151	60 (52, 67)	63 (57, 69)
		>25	75	59	79	180	49 (41, 57)	75 (69, 80)
		>35	50	29	104	210	33 (26, 40	88 (83, 91)
		>41	42	20	112	219	27 (21, 35)	92 (87, 94)
		>42	40	20	114	219	26 (20, 33)	92 (87, 94)
		>43	39	19	115	220	25 (19, 32)	92 (88, 95)
		>44	39	19	115	220	25 (19, 32)	92 (88, 95)
		>45	38	19	116	220	23 (17, 31)	92 (88, 95)
		>46	38	17	116	222	27 (21, 35)	92 (87, 94)
		>71	27	7	127	232	18 (12, 24)	97 (94, 99)
Schneider 2009[77,78]	75/160 (46.9%)	>20	48	36	27	49	64 (53, 74)	58 (47, 77)
		>12	64	65	11	20	85 (76, 92)	24 (16, 34)
		>16	52	40	23	45	69 (58, 79)	53 (42, 63)
		>35	24	14	51	71	32 (25, 42)	84 (74, 90)
		>46	24	6	51	79	32 (23, 43)	93 (85, 97)
		>76	10	0	65	85	13 (7, 23)	100 (96, 100)
Smith 2005[87]	27/52 (51.92%)	≥15	22	13	5	12	81.5	48
		>47	15	2	12	23	55.6	92
		<15	5	12	22	13	18.5	52
De La Barra 2011[88]	NR	25	10	17	2	23	83.3	57.5
		40	9	12	3	28	75	70
		50	7	8	5	32	58.3	80
		70	5	5	7	35	41.7	87.5
		90	5	3	7	37	41.7	92.5
		110	3	2	9	38	25	95
		130	2	2	10	38	16.7	95
		150	2	2	10	38	16.7	95
Smith 2004[90]	17/47 (36.2%)	>20	14	6	2	22	88	79
Fortuna, 2007[76]	Induced Sputum (Eos%) 16/50 (32.0%); Bronchodilator test 13/50 (26.0%); FEV_1 <80% 5/50 (10.0%)	≥20 ppb	17	10	5	18	77	64
Fukuhara,	42/61 (68.85%)	40 ppb	33	2	9	17	78.6	89.5

2011[53]								
Subset of Position A								
Cordeiro 2011[91]	42/114 (36.8%)	27	33	6	9	66	78	92
Heffler 2006[86]	18/48	>10 >15 >20 >25 >30 >34 >36 >40 >45 >50 >55 >60 >65 >75 >80 >85 >100	18 18 18 18 14 14 14 11 11 10 9 9 8 8 7 5 5	29 26 20 16 15 14 12 11 8 7 6 4 4 3 1 1 0	0 0 0 0 4 4 4 7 7 8 9 9 10 10 11 13 13	1 4 10 14 15 16 18 19 22 23 24 26 26 27 29 29 30	100 100 100 100 77.8 77.8 77.8 61.1 61.1 55.6 50 50 44.4 44.4 38.9 27.8 27.8	3.3 13.3 33.3 46.7 50 53.3 60 63.3 73.3 76.7 80 86.7 86.7 90 96.7 96.7 100
Pizzimenti 2009[94]	14/156 (9.0%)	55	11	17	3	125	78	88
Difficult to diagnose								
Schleich, 2012[83]	82/174 (47.1%)	34	29	4	53	88	35.4	95.4
Pedrosa 2010[89]	35/114 (30.7%)	40	26	22	9	57	74.3	72.5
Bobolea 2012[92]	6/30 (20.0%)	>30	6	17	0	7	100	29.2
Schneider 2009[77,78]	49/101 (48.5%) subjects with unsuspicious spirometric results	>46 >15	17 38	5 29	32 11	47 23	35 (23, 49) 78 (63, 89)	90 (79, 96) 45 (34, 57)
Mathew 2011[95]	20/84 (23.8%)	NR	2	21	18	43	10	67.2
Difficult to diagnose with chronic cough								
Hsu 2013[79]*	38/81 (46.9%)	33.9 30	36 37	9 14	2 1	34 29	94.7 97.4	76.3 65.8
Hahn 2007[80]*	38/64 (59.4%)	35 38	36 34	5 4	2 4	21 22	95 (83, 99) 90 (76, 96)	80 (62, 92) 85 (76, 96)
Prieto 2009[82]*	19/43 (44%)	20	10	9	9	15	53	63

Sato 2008[81]	48/71 (67.6%)	38.8	38	2	10	21	79.2	91.3
Zhang 2011[93]	39/106 (36.8%)	40 36*	29 32	9 5	10 7	58 62	75 82	86 93
EIB								
El Halawani 2003[84]	7/49 (14.3%)**	<12	7	29	0	13	100	31
Other								
Arora 2006[85]	138/172 (80.2%)	>6 >7 >8 >9 >10 >11 >12 >13 >14 >15 >16 >17 >18 >19 >20 >25 >30 >40 >46	133 131 130 127 119 115 113 110 102 98 92 87 83 78 73 56 45 32 23	34 33 31 30 28 26 25 21 19 19 17 14 14 13 11 7 7 3 0	5 7 8 11 19 23 25 28 36 40 46 51 55 60 65 82 93 106 115	0 1 3 4 6 8 9 13 15 15 17 20 20 21 23 27 27 31 34	96.4 94.6 94.2 92 86.2 83.3 81.9 79.7 73.9 71 66.7 63 60.1 56.5 52.9 40.6 32.6 23.2 16.7	0 2.9 8.8 11.8 17.6 23.5 26.5 38.2 44.1 44.1 50 58.8 58.8 61.8 67.6 79.4 79.4 91.2 100

*: Data are for ICS responsiveness
**: Test for exercise induced bronchoconstriction

Appendix 13: Table of results for all diagnostic studies in children

Author, year	Prevalence of positive result by reference	FeNO cut-off ppb	True positive	False positive	False negative	True negative	Sensitivity asthma %	Specificity asthma %
Linkosalo 2012[97]	18/30 (60%)*	10	16	8	2	4	89	33
		20	13	2	5	10	72	83
		30	9	1	9	11	50	92
		40	7	1	11	11	39	92
		50	2	1	16	11	11	92
Ramser 2008[98]	105/169 (62.13%)	10	75	24	45	25	76	36
		20	49	50	17	53	49	76
		30	33	66	12	58	33	83
		40	23	76	7	63	23	90
		50	20	79	5	65	20	93
Sivan 2009[99]	106/150 (70.67%)	15	62	13	7	31	90	70
		18	87	7	19	37	82	84
		19	59	5	10	39	86	89
		25	52	5	17	39	75	89
		>20 or <15	58	4	7	32	89	88
Woo 2012[100]	167/245 (68.2%)	>50	24	0	143	78	14.4	100
		>45	29	0	138	78	17.4	100
		>41	39	0	128	78	23.4	100
		>40	41	1	126	77	24.6	98.7
		>35	54	1	113	77	32.3	98.7
		>30	71	4	96	74	42.5	94.9
		>25	83	6	84	72	49.7	92.3
		>24	84	7	83	71	50.3	91
		>23	86	7	81	71	51.5	91
		>22	90	10	68	68	53.9	87.2
		>21	95	10	68	68	56.9	87.2
		>20	101	15	63	63	60.5	80.8
		>15	120	26	52	52	71.9	66.7
		>10	134	43	35	35	80.2	44.9
		>5	157	67	11	11	94	14.1
		22	93	9	51	51	72.1	85

Appendix 14: Medline strategies for searches for the Economic Review

1. NIOX MINO/NObreath in either diagnosis or management of asthma (30th May 2013)

1	niox mino.mp.
2	aerocrine.mp.
3	(niox adj5 (monitor$ or chemiluminescence or analyser$ or sensor or device$ or desktop)).mp.
4	nobreath.mp.
5	bedfont.mp.
6	or/1-5
7	exp "Costs and Cost Analysis"/
8	Economics/
9	exp Economics, Hospital/
10	exp Economics, Medical/
11	Economics, Nursing/
12	exp models, economic/
13	Economics, Pharmaceutical/
14	exp "Fees and Charges"/
15	exp Budgets/
16	budget$.tw.
17	ec.fs.
18	cost$.ti.
19	(cost$ adj2 (effective$ or utilit$ or benefit$ or minimi$)).ab.
20	(economic$ or pharmacoeconomic$ or pharmaco-economic$).ti.
21	(price$ or pricing$).tw.
22	(financial or finance or finances or financed).tw.
23	(fee or fees).tw.
24	(value adj2 (money or monetary)).tw.
25	quality-adjusted life years/
26	(qaly or qalys).af.
27	(quality adjusted life year or quality adjusted life years).af.
28	or/7-28
29	6 and 28

2. Models of Asthma and FENO (30th May 2013)

1	niox mino.mp.
2	aerocrine.mp.
3	(niox adj5 (monitor$ or chemiluminescence or analyser$ or sensor or device$ or desktop)).mp.
4	nobreath.mp.
5	bedfont.mp.
6	or/1-5
7	exp cough/
8	cough$.mp.
9	phlegm.mp.
10	sputum.mp.
11	mucus.mp.
12	wheez$.mp.
13	chest pain/
14	chest pain$.mp.
15	(chest adj5 tight$).tw.
16	((lower respiratory or lrt) adj5 symptom$).tw.
17	(lower airway adj5 symptom$).tw.
18	((trache$ or wind pipe or lung$ or bronch$) adj3 symptom$).tw.
19	exp lung/ or trachea/
20	symptom$.tw.
21	19 and 20
22	or/7-18,21

23	exp asthma/
24	asthma$.mp.
25	exp respiratory hypersensitivity/
26	exp bronchial hyperreactivity/
27	bronchial spasm/
28	bronchospas$.mp.
29	exp Bronchoconstriction/
30	bronchoconstric$.mp.
31	(bronch$ adj3 constrict$).mp.
32	(bronch$ adj5 spas$).mp.
33	(airway$ adj5 (obstruct$ or inflammation$)).mp.
34	((bronchial$ or respiratory or airway$ or lung$) adj3 (hypersensitiv$ or hyperreactiv$ or allerg$ or insufficiency)).mp.
35	or/23-34
36	Nitric Oxide/
37	nitric oxide.mp.
38	36 or 37
39	(exhal$ or expir$ or alveolar or fractional).mp.
40	38 and 39
41	exhaled NO.mp.
42	eno.mp.
43	fe?no$.mp.
44	(fractional adj2 NO).mp.
45	or/40-44
46	22 and 45
47	35 and 45
48	6 or 46 or 47
49	exp "Costs and Cost Analysis"/
50	Economics/
51	exp Economics, Hospital/
52	exp Economics, Medical/
53	Economics, Nursing/
54	exp models, economic/
55	Economics, Pharmaceutical/
56	exp "Fees and Charges"/
57	exp Budgets/
58	budget$.tw.
59	ec.fs.
60	cost$.ti.
61	(cost$ adj2 (effective$ or utilit$ or benefit$ or minimi$)).ab.
62	(economic$ or pharmacoeconomic$ or pharmaco-economic$).ti.
63	(price$ or pricing$).tw.
64	(financial or finance or finances or financed).tw.
65	(fee or fees).tw.
66	(value adj2 (money or monetary)).tw.
67	quality-adjusted life years/
68	(qaly or qalys).af.
69	(quality adjusted life year or quality adjusted life years).af.
70	or/49-69
71	48 and 70

3. Asthma management models (3rd June 2013)

1. exp asthma/
2. asthma$.mp.
3. exp respiratory hypersensitivity/
4. exp bronchial hyperreactivity/
5. bronchial spasm/
6. bronchospas$.mp.
7. exp Bronchoconstriction/
8. bronchoconstric$.mp.
9. (bronch$ adj5 spas$).mp.
10. (airway$ adj5 (obstruct$ or inflammation$)).mp.
11. ((bronchial$ or respiratory or airway$ or lung$) adj3 (hypersensitiv$ or hyperreactiv$ or allerg$ or insufficiency)).mp.
12. exp models, economic/
13. (cost$ adj2 (effective$ or utilit$ or benefit$ or minimi$)).ti.
14. ((cost$ or economic) adj5 model$).ti.
15. or/1-11
16. or/12-14
17. 15 and 16

4. Asthma diagnostic models (7th June 2013)

1	exp asthma/
2	asthma$.mp.
3	exp respiratory hypersensitivity/
4	exp bronchial hyperreactivity/
5	bronchial spasm/
6	bronchospas$.mp.
7	exp Bronchoconstriction/
8	bronchoconstric$.mp.
9	(bronch$ adj3 constrict$).mp.
10	(bronch$ adj5 spas$).mp.
11	(airway$ adj5 (obstruct$ or inflammation$)).mp.
12	((bronchial$ or respiratory or airway$ or lung$) adj3 (hypersensitiv$ or hyperreactiv$ or allerg$ or insufficiency)).mp.
13	or/1-12
14	exp "Costs and Cost Analysis"/
15	Economics/
16	exp Economics, Hospital/
17	exp Economics, Medical/
18	Economics, Nursing/
19	exp models, economic/
20	Economics, Pharmaceutical/
21	exp "Fees and Charges"/
22	exp Budgets/
23	budget$.tw.
24	ec.fs.
25	cost$.ti.
26	(cost$ adj2 (effective$ or utilit$ or benefit$ or minimi$)).ab.
27	(economic$ or pharmacoeconomic$ or pharmaco-economic$).ti.
28	(price$ or pricing$).tw.
29	(financial or finance or finances or financed).tw.
30	(fee or fees).tw.
31	(value adj2 (money or monetary)).tw.
32	quality-adjusted life years/
33	(qaly or qalys).af.

34	(quality adjusted life year or quality adjusted life years).af.
35	or/14-34
36	exp "Sensitivity and Specificity"/
37	sensitivity.tw.
38	specificity.tw.
39	((pre-test or pretest) adj probability).tw.
40	post-test probability.tw.
41	predictive value$.tw.
42	likelihood ratio$.tw.
43	diagnostic$.ti,ab.
44	or/36-43
45	13 and 35 and 44