NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

DIAGNOSTICS ASSESSMENT PROGRAMME

Evidence overview

Measurement of exhaled nitric oxide concentration in asthma: NIOX MINO, NIOX VERO and NObreath

This overview summarises the key issues for the Diagnostics Advisory Committee's consideration. It includes a brief description of the topic, a description of the analytical structure and model, a discussion of the analytical difficulties, and a brief summary of the results. It is not a complete summary of the diagnostics assessment report, and it is assumed that the reader is familiar with that document. This overview contains sections from the original scope and the diagnostics assessment report, as well as referring to specific sections of these documents.

1. Background

1.1 Introduction

The Medical Technologies Advisory Committee identified NIOX MINO (Aerocrine Ltd.) as potentially suitable for evaluation by the Diagnostics Assessment Programme on the basis of a briefing note. The technology measures the proportion of nitric oxide in exhaled breath with the aim to assist in the diagnosis, management and monitoring of asthma. The scope was extended to include another similar test (the NObreath test, Bedfont Scientific Ltd.), and outlines the approach for assessing the clinical and cost effectiveness components for exhaled nitric oxide concentration in asthma in both primary and secondary care.

During the assessment phase, Aerocrine Ltd. notified NICE and the External Assessment Group (EAG) that a new technology, NIOX VERO, is being

launched to replace NIOX MINO. Aerocrine made data on NIOX VERO available to the assessment group and it was included in the Diagnostics Assessment Report (DAR).

1.2 The condition

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 recurring 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 mite.

In people with asthma, cellular inflammation of the airways with eosinophils and neutrophils is considered to be a characteristic feature relevant to the pathogenesis of the disease. Eosinophilic asthma is a distinct phenotype of asthma that is associated pathologically by a thickening of the basement membrane zone. The symptoms of eosinophilic asthma can be controlled with corticosteroids. In contrast, neutrophilic asthma generally does not respond to treatment with corticosteroids.

In humans, nitric oxide, which is produced in the lungs and is present in exhaled breath, has been implicated in the pathophysiology of lung diseases, including asthma. It has been shown to act as a vasodilator, bronchodilator, neurotransmitter and inflammatory mediator in the lungs and airways. Over the years, fractional exhaled nitric oxide (FeNO) has been proposed as a non-invasive marker of eosinophilic airway inflammation in asthma. FeNO has been shown to be elevated in people with eosinophilic asthma and effective treatment with corticosteroids reduces the level of FeNO.

Asthma usually develops in childhood but may start at any age. 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 person 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.

Diagnostic and care pathways

Diagnosis

Asthma is diagnosed on the basis of symptoms and objective tests of lung function. Lung function measurements include peak expiratory flow rate (PEF), forced expiratory volume in the first second (FEV1) and percentage predicted FEV1 (calculated as a percentage of the predicted FEV1 for a person of the same height, sex and age without diagnosed asthma). Variability of PEF and FEV1, 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).

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.

Diagnosis in children

The diagnosis of asthma 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 Diagnosis of Asthma (see figure 3, page 43 of the DAR), 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 Thoracic Society/Scottish Intercollegiate Guidelines Network (BTS/SIGN),Guidelines, 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. For those 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. In such cases, specialist referral should always be considered.

Diagnosis in adults

The diagnosis of asthma in adults is 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. Spirometry is the preferred initial test to assess the presence and severity of airflow obstruction.

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. See figure 4, page 44 of the DAR).

Monitoring

For both children and adults, asthma is monitored 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.

Management

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 figures 5 and 6, pages 46-47 of the DAR). Management options include interventions with or without the use of drugs.

According to the guideline, there is evidence that suggests reducing parental smoking and encouraging breast-feeding may both reduce the chance of developing asthma. In addition, allergen avoidance, dietary manipulations and modified infant milk formulae have all shown inconsistent effects. Other non-drug management options include smoking cessation, weight reduction, allergen avoidance, immunotherapy (where there is a clinically significant and identified allergen that cannot be avoided) and dietary modifications.

In children, The BTS/SIGN guidelines 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.

The British Guideline on the Management of Asthma describes a stepwise approach to drug treatment in adults and children over 12. Mild intermittent asthma (step 1) is treated with inhaled short-acting Beta-2 agonists, as required. The introduction of regular preventer therapy with inhaled corticosteroids (step 2) should be considered when a person has had exacerbations of asthma in the previous 2 years, is using inhaled short-acting Beta-2 agonists three times a week or more, is symptomatic three times a week or more, or is waking at night once a week because of asthma. Add-on therapy (step 3) involves the introduction of an additional therapy, the first choice of which is an inhaled long-acting beta-2 agonist. Alternatives include orally administered leukotriene receptor antagonists, theophyllines and slow-release beta-2agonist tablets, or increasing the dose of inhaled cortico-steroids. At step 4, further interventions may be considered if control remains inadequate on a dose of inhaled corticosteroids that is equivalent to 800 micrograms per day of beclometasonedipropionate in combination with a long-acting Beta-2 agonist, or following an unsuccessful trial of long-acting Beta-2 agonists. Options include increasing the dose of the inhaled corticosteroids to 2000 micrograms beclometasonedipropionate equivalent per day or adding a leukotriene antagonist, a theophylline or a slow-release Beta-2 agonist tablet. At step 5, continuous or frequent courses of oral corticosteroids are introduced. The majority of people with asthma are treated at steps 1, 2 or 3.

The population

For the 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.

For the 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 inhaled corticosteroids or are being checked for compliance to treatment.

2. The technologies

NIOX MINO

NIOX MINO is a diagnostic and monitoring device that analyses a breath sample using an electrochemical sensor to determine exhaled nitric oxide concentration. The technology is designed to aid in the identification of patients whose airway inflammation will respond to treatment with inhaled corticosteroids. It can also be used to predict the onset of asthma symptoms or loss of asthma control and also to monitor compliance to corticosteroid therapy and the effectiveness of treatment.

NIOX MINO determines ENO 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 ENO 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 stand-alone or connected to a PC for monitoring with the NIOX MINO Data Management Program and for use with Electronic Medical Record systems. The device 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.

NIOX VERO

During the assessment phase, Aerocrine began launching NIOX VERO, a new FeNO device which is intended to replace NIOX MINO. The new device is a battery powered device which features a longer operational life and extended test volume life than NIOX MINO. The device is anticipated to obtain a CE mark by the time draft guidance goes out for consultation.

NObreath

NObreath (Bedfont Scientific Ltd.) is diagnostic monitoring device that measures ENO 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.

2.1 The Comparator

Current NHS practice

Scoping workshop attendees indicated that the SIGN and British Thoracic Society's British Guideline on the Management of Asthma are considered an appropriate comparator for people with asthma. Testing for airway hyperresponsiveness forms a part of the current NHS practice in diagnosis and management of asthma. This is synonymous with bronchial hyperresponsiveness which is an indicator of asthma. It is 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.

3. The evidence

This section summarises data from the diagnostics assessment report compiled by the External Assessment Group.

3.1 Clinical effectiveness

The External Assessment Group conducted the following reviews to identify the clinical evidence relevant to the decision problem:

1. Review of equivalence of FeNO devices (analytic validity)
2. Systematic review of diagnostic accuracy of FeNO for asthma
3. Systematic review of the efficacy of FeNO-guided management of asthma

Review of equivalence of FeNO devices

The External Assessment Group undertook this review 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. As there was insufficient evidence from primary research studies which used NIOX MINO, NIOX VERO and NObreath, a review of equivalence to other FeNO chemiluminescent devices (including Niox, a precursor to NIOX MINO) was conducted.

The review identified 27 studies that compared NIOX MINO, NIOX VERO and NObreath to other devices. The External Assessment Group undertook three main comparisons for this purpose:

- Comparisons of means: this compared reported mean FeNO values as measured by each device in the same cohort.
- Correlation coefficients: these show whether measurements by the two devices are correlated, but not whether the actual values produced are the same.
- Bland-Altman analysis: the results of this analysis produce statistics that assess agreement between devices rather than just correlation.

The results of the equivalence review are summarised below.

a) NIOX MINO

Eight studies (table 12, page 82 of the DAR) compared NIOX MINO to Niox (an earlier version of NIOX MINO) in adults. Five of these studies were exclusively in adults and three were undertaken in a mixture of adults and other age groups. There was variability in correlation between the devices among the studies. Five studies showed largely similar mean values between NIOX MINO and Niox, whereas three studies showed NIOX MINO providing higher FeNO readings (range 0.5 to 9ppb). Small (non-significant) differences in the mean FeNO readings were observed between the devices when the cohort mean FeNO values were below 30ppb (as measured by Niox). When the mean FeNO value was above 35ppb, the differences in cohort means were larger and statistically significant. Correlation coefficients ranged from 0.73 to 0.998. The result of one study suggests that there may also be some variation between NIOX MINO devices themselves. Among the eight studies,

Bland-Altman analyses were not reported in a consistent way. Limits of agreement were in some cases 10ppb above and below the mean, and the studies with the largest mean differences did not report Bland-Altman statistics.

Three studies (table 13, page 84 of the DAR) compared NIOX MINO to Niox in children. Two studies reported statistically significantly higher mean FeNO values with NIOX MINO, whilst one study reported statistically significantly lower values. This study had low mean values (below 10ppb). All studies reported good correlation between the devices while Bland-Altman statistics reported in two studies 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.

Twelve studies (table 14, page 87 of the DAR) compared NIOX MINO to other stationary chemiluminescent devices in adults and/or children. Six studies were in adults, three in an unspecified group and three in children. The chemiluminescence devices used in each of the twelve studies were different. In the adults/unspecified age group, correlation coefficients ranged from 0.876 to 0.96, indicating a good level of correlation between devices while the mean FeNO levels and Bland-Altman statistics paint a different picture. In four studies, NIOX MINO gave higher readings than the comparator device and lower readings in only two studies. Two studies show the devices to be comparable. Bland-Altman statistics, reported in four studies suggest that mean differences were small, but the limits of agreement were much larger.

In children, correlation coefficients, between NIOX MINO and other chemiluminescent devices, ranged from 0.69 to 0.98 indicating variable correlation. The study with the poorer correlation had higher mean FeNO levels, suggesting that poorer correlation is due to the greater variability at higher FeNO values. However, the study authors stated that correlation improved at higher values. One study noted that the direction of disagreement was different in children aged over and under 12 years. The back-transformed

Bland-Altman statistics and range of ratios reported indicate a wide range of agreement, and suggest the devices are not interchangeable.

The External Assessment Group stated that the comparability of NIOX MINO to chemiluminescent devices appears to be influenced by several factors:

1. There may be some variability between NIOX MINO devices themselves.
2. There appears to be a lack of comparability between other chemiluminescent devices themselves which leads to heterogeneity in estimates of comparability between these devices and NIOX MINO.
3. It appears that at higher FeNO levels, equivalence between the devices is poorer.

The External Assessment Group concluded that the different estimates of equivalence between NIOX MINO and other devices make it unclear if equivalence can be assumed.

b) NIOX VERO

One study which compared NIOX VERO to NIOX MINO was submitted by the manufacturer. (Table 15, page 90 of the DAR). [redacted]

c) NObreath

Four studies (table 16, page 92 of the DAR) compared NObreath to three other chemiluminescent devices. Bland-Altman analysis was undertaken in only one study and in a healthy cohort with low FeNO values, showed a mean difference of -3.95ppb in comparison to the chemiluminescent device. Limits of agreement in this study were wide (-10.98 to 4.08). Another study reported an absolute mean difference in FeNO measurements of -3.81ppb. Comparisons with the third type of chemiluminescent device showed small

differences between mean FeNO values for the cohort, with NObreath giving lower values in some cohorts.

Two studies (table 17, page 95 of the DAR) which compared NObreath to NIOX MINO in adults found that NIOX MINO provided lower mean FeNO values than NObreath in most analyses. This contradicts the available evidence for comparisons of NIOX MINO to Niox and NObreath to Niox, which suggest that NIOX MINO should provide higher readings than NObreath. The two direct comparisons of NObreath and NIOX MINO are in small numbers of patients, and only one includes asthmatic patients, but does not provide a Bland-Altman analysis to assess agreement. The External Assessment Group concluded that it is unclear whether the NIOX MINO and NObreath are interchangeable, and if not in which direction the difference may be.

]The External Assessment Group concluded that, based on available evidence, any differences in absolute values between results from NObreath and other devices are relatively small, though derived cut-offs and maximum sensitivity and specificity may differ. The External Assessment Group however stated that the limited data available was too conflicting for any conclusions to be drawn regarding the concordance of NObreath results with results from other devices.

Systematic review of diagnostic accuracy of FeNO for 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 identified 24 studies that met the inclusion criteria; 20 were conducted with adults of all ages and 4 with children. The studies were classified according to the position of the patients in the UK pathway (see figure 8, page 58 of the DAR) and the reference standards used. These groups were:

a) Adults presenting with symptoms of asthma versus most of or the entire UK pathway

The review identified 4 studies in this group. Cut-off for the highest sum of sensitivity and specificity ranged from 20 parts per billion (ppb) to 47 ppb amongst the four studies in this group. Sensitivities ranged from 32% to 88%, and specificities from 75% to 93%. Due to the heterogeneity between results, study designs and devices used, the External Assessment Group concluded that it is difficult to draw any conclusion as to the optimal cut-off for sensitivity and specificity.

Not all of the studies reported a range of cut-offs 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 76 ppb, with sensitivities from 56% to 13% and specificities from 88% to 100%.

Values of specificities consistently had a smaller range and higher values than values of sensitivities reported. This suggests 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 correct or incorrect classification of patients.

b) A subset of adults presenting with symptoms of asthma versus airway hyper-responsiveness

Two studies were identified in this subgroup. The estimates of diagnostic accuracy were not noticeably different to those produced by studies which recruited a broader spectrum of patients. The cut-off for the highest sum of

sensitivity and specificity ranged from 27 to 36ppb, sensitivities from 78% to 87% and specificities from 60% to 92%.

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

- When selecting the cut-off with the highest sensitivity, this was 25ppb, sensitivity 100%, specificity 47%
- When selecting the cut-off with the highest specificity, this was 100 ppb, sensitivity 28%, specificity 100%.

c) Difficult to diagnose patients versus airway hyper-responsiveness

Three studies used some form of airway hyper-responsiveness as the reference standard. Estimates of sensitivity and specificity appeared comparable to the studies recruiting patients presenting to primary care with symptoms of asthma versus airway reversibility, response to a trial of treatment with inhaled corticosteroids and airway hyper-responsiveness. One study recruited a set of patients who were negative by a methacholine challenge test (MCT) and compared FeNO to an adenosine challenge test. This study produced 100% sensitivity (29% 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 40 ppb amongst the studies versus MCT, which is a little narrower than in the broader cohorts. Sensitivities ranged from 24% to 74%, and specificities from 73% to 99%, which is a similar range to the broader cohort. A range of cut-offs was not reported in these studies.

d) Patients with chronic cough who were difficult to diagnose, versus response to a trial of treatment with inhaled corticosteroids

Three studies recruited patients with chronic cough who had tested negative for other causes. All three studies used response to a trial of treatment with inhaled corticosteroids as a reference standard. 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 95% sensitivity with 76% specificity and 90% sensitivity with 85% specificity.

e) Children with symptoms of asthma versus various reference standards

Four studies which compared adult cohorts with a similar spectrum of patients and reference standards were identified. The cut-offs 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 21 pbb. 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 20 ppb (compared with 9 ppb to 15 ppb in adults), sensitivity from 89% to 94% (85% to 96% in adults), and specificity from 14.1% to 70% (13% to 48% in adults). When selecting the cut-off with the highest specificity, results were also similar to adult cohorts. Cut-offs were a little lower again, and ranged from 30 to 50 ppb (compared with 47 to 76ppb in adults), sensitivity ranged from 20% to 50% (13% to 56% in adults, specificity from 92% to 100% (88% to 100% in adults).

Due to high heterogeneity between studies, no meta-analysis was conducted in any group. 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.

Because of this, the External Assessment Group found it difficult to estimate the relative diagnostic accuracy of FeNO in any situation and at any given cut-off point. However, there did not appear to be a difference in the relative diagnostic accuracy of FeNO in the two settings (primary and secondary care): in comparison with the standard UK pathway (either entire or parts) or in comparison with airway hyper-responsiveness in patients who are difficult to diagnose. The large variation in estimates within groups may however obscure any true underlying differences in the accuracy of FeNO between groups and versus different reference standards.

Some limited observations were made by the External Assessment Group:

- 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 true positive, true negative, false positive and false negative.
- That FeNO cut-off points should probably be lower in children than in adults.
- 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.

f) Subgroups

No cohort studies were found that provided evidence relating to the subgroups of pregnant women, older people and smokers or people exposed to environmental tobacco. As such, lower levels of evidence were consulted.

Among smokers FeNO appeared to be able to distinguish people with asthma from people not with asthma 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.

Among older people, a case control study indicated that FeNO is unlikely to be a useful test in the diagnosis of asthma in the older population.

Among 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

The External Assessment Group reviewed evidence relating to outcomes in adults, children and subgroups of people as defined in the scope for this assessment. The outcomes included exacerbations, inhaled corticosteroid use and health-related quality of life.

a) FeNO-guided management of asthma in adults

Four studies (based in the UK, New Zealand, Sweden and USA) were included in this review. Quality of the four studies was assessed according to the Cochrane and Centre for Reviews and Dissemination (CRD) handbook. The study with the highest risk of bias was the unpublished study by Syk; this was due to the lack of participant personnel blinding, incomplete outcome data and selective reporting.

Table 27, page 157 of the DAR provides details of the study design and timelines of the four studies. All were RCT studies, two of which were single

blind, one was open label and one was described as “multiply blinded”. There was a high degree of heterogeneity in all aspects of study design across the four studies.

Table 28, page 159 of the DAR provides details of patient characteristics across the four studies. The number of participants ranged from 94 to 229 across the studies. Participants were recruited from primary care in three studies. One study was unclear whether participants were recruited from primary or secondary care settings.

Table 29, page 162 of the DAR provides details of the interventions used in each study. Three studies did not clearly report which device was used to measure FeNO.

Table 30, page 165 of the DAR provides details of the control interventions used in each study. The inclusion criteria, trial protocols and treatment doses varied across the studies in both. Only one study reported using UK guidelines in the comparator arm.

Exacerbations were reported in all four studies although definitions varied (table 31, page 170 of the DAR) and results were not consistent across the studies. All four studies reported a fall in exacerbation rates per person year, though it appeared that this was mostly driven by mild and moderate exacerbations.

For severe exacerbations, Syk paper 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 oral corticosteroid use and the composite outcomes which include less severe exacerbations was less pronounced. Oral corticosteroid 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. Rate ratios calculated

by the External Assessment Group for major/severe exacerbations ranged from 0.79 (95% CI 0.44 to 1.41) to 1.29 (95% CI 0.51 to 3.30) whilst rate ratios calculated by the External Assessment Group 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).

Despite the high level of between-study heterogeneity, an exploratory meta-analysis of the rates of major/severe exacerbations using fixed effects methods was conducted. The result (figure 14, page 168 of the DAR) showed no heterogeneity, with an I^2 statistic of 0%. The pooled estimate was 0.87 (95% CI 0.64 to 1.19, p=0.38). This indicates fewer major exacerbations in the intervention arm, but this did not reach statistical significance.

A sensitivity analysis was undertaken using the results of studies that reported the number of exacerbations resulting in the use of oral corticosteroids. The pooled estimate continued to show a statistically non-significant trend towards a positive effect for asthma management with FeNO with a risk ratio of 0.97 (95% CI 0.61 to 1.54). However, there were only two studies in this analysis, and one showed a trend towards a reduction in oral corticosteroids use whilst Syk conversely showed a trend towards an increase in oral corticosteroids 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 exacerbations, three studies reported composite outcomes which the External Assessment Group considered to be broadly similar and represent “treatment failure”. In two studies, FeNO guided groups showed numerically but not statistically significant lower rates of failure. In the Syk study, the improvement was statistically significant, with a rate of 0.22 in the intervention arm compared with 0.41 in the control arm (p=0.024). The rate ratio calculated by the External Assessment Group was 0.52 (95% CI 0.30 to 0.91).

A meta-analysis of these rates was conducted despite the high level of heterogeneity between study characteristics. The result (figure 16, page 169 of the DAR) showed a statistically significant effect with a rate ratio of 0.58 (95% CI 0.43 to 0.77). This represents a statistically significant effect in favour of using FeNO guided management in asthmatics for this outcome.

All studies reported some data on inhaled corticosteroid use. Two studies reported inhaled corticosteroid 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. The Syk study showed a small (non-significant) increase in inhaled corticosteroid use in the intervention arm (586µg (SE 454) versus 540µg (SE 317) in the control arm. One study 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 (figure 17, page 175 of the DAR) used standardised mean difference analysis because outcomes were not reported in a standardised way. This 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).

Two studies used versions of the Asthma Quality of Life Questionnaire (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. A meta-analysis of the overall scores (figure 18 page 177 of the DAR) shows no effect with a standardised mean of 0.00 (-0.20 to 0.20).

:

All four studies reported data for asthma control. In three studies, asthma control did not change. In the Syk study, there was a statistically significant

increase in asthma control between the two trial arms. Two studies reported no significant difference between groups for bronchodilator use, and the Syk study reported non-significant trends towards greater numbers and mean use of Leukotriene receptor antagonist and Short-acting beta-2 agonists (significance not reported) in the FeNO controlled arm.

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

b) FeNO-guided management of asthma in children

Five studies (based in Austria, USA, Italy, Netherland and Australia) which recruited children (plus adolescents and/or young adults) and compared FeNO-guided management were identified. The quality of the studies was assessed according to criteria proposed in the Cochrane Handbook and CRD Handbook. The study quality varied; with no one study scoring well in every item, and no item scoring well in every study.

There was a high degree of heterogeneity in all aspects of study design across 4 studies. No study reported using UK guidelines in the comparator arm. Two studies recruited patients who appeared to be poorly controlled. One study recruited patients who were mild to moderate persistent asthmatics and one study recruited patients who had a stable dose of inhaled corticosteroids for the previous 3 months, suggesting they were reasonably well controlled.

All five studies reported some data on asthma exacerbations although the definition of exacerbation was unclear in some cases. Two studies reported severe exacerbations in a way that allowed calculation of rates per person year. Both had lower rates in the intervention arm. In uncontrolled asthmatics rate in the intervention arm was 0.746 and 0.950 in the control arm. In patients who had been on a stable dose of inhaled corticosteroids for three months; rate in the intervention arm was 0.21 and 0.39 in the control arm. Both rates were calculated by the reviewer and the statistical significance is unclear.

For all definitions of exacerbations, four studies reported outcomes that were not defined as either major or minor and had different definitions to one another. All the studies showed trend in favour of fewer exacerbations in the intervention arm. The only study to report a significant between-group difference was a conference abstract which showed exacerbations (not clearly defined) occurred in 6/31 participants in the intervention group (19.4%), and 15/32 in the control group (46.9%, p=0.021).

Table 40, page 201 of the DAR provides details of inhaled corticosteroid use in each study. Overall, there was indication of between-group differences in inhaled corticosteroid use although there was variability between the studies. These differences could be attributed to the specifics of the step-up/step-down protocols and/or the characteristics of the patients selected.

Health Related Quality of Life was only reported in one study in abstract form and using an unknown tool. The External Assessment Group was not able to draw a definite conclusion from this data.

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.

One study 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.

3.2 Costs and cost effectiveness

Economic analysis

The economic analysis undertaken by the External Assessment Group sought 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?

A systematic review of existing economic analyses as well as the construction of de novo economic models were undertaken to answer the above questions.

Review of existing economic analyses

The External Assessment Group undertook a review to identify existing economic analyses of FeNO testing (using NIOX MINO, NIOX VERO or NObreath) for diagnosing and managing asthma. The review also sought to identify existing models and potentially relevant evidence sources to inform parameter values within the *de novo* economic models developed by the External Assessment Group.

Only one published UK cost-effectiveness model was identified for asthma diagnosis and one for management of asthma. Modified versions of these models were submitted to NICE by Aerocrine. The wider review identified a number of economic analyses which were subject to a number of methodological problems, questionable assumptions and weak evidence.

Detailed descriptions and critical appraisal of these economic analyses are available in section 6.3 of the DAR.

De novo cost-effectiveness model

a) Model aim

In order to resolve the problems and gaps in evidence identified in the systematic review of existing economic analyses, the External Assessment Group developed two *de novo* models to assess:

1. the expected cost-effectiveness of FeNO testing in addition to or in place of standard tests for the diagnosis of asthma and

2. The expected cost-effectiveness of FeNO plus standard guidelines versus standard guidelines for the management of patients with diagnosed asthma.

The two models, although distinct, share a number of parameter values and assumptions.

b) Model structure

The diagnostic model (figure 1 below) is structured in the form of a decision tree. The decision tree model was used to estimate the probability that a person with asthma will be correctly diagnosed (true positive) or incorrectly diagnosed (false negative); and the probability that a person without asthma will be correctly diagnosed (true negative) or incorrectly diagnosed (false positive) and the expected health outcomes and costs arising from this.

The management model (figure 2 below) is in the form of a Markov model with two states; alive with diagnosed asthma and dead.

Figure 1: Asthma diagnostic model

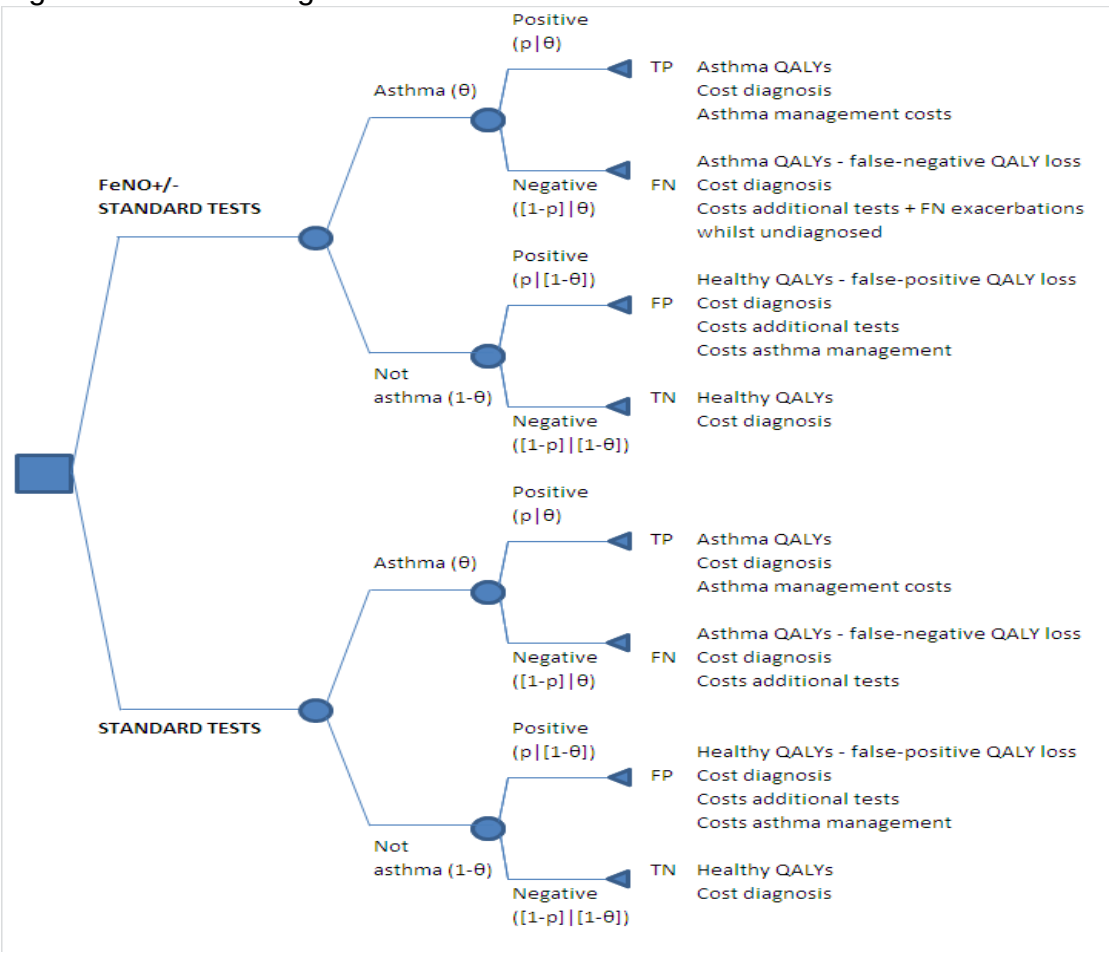
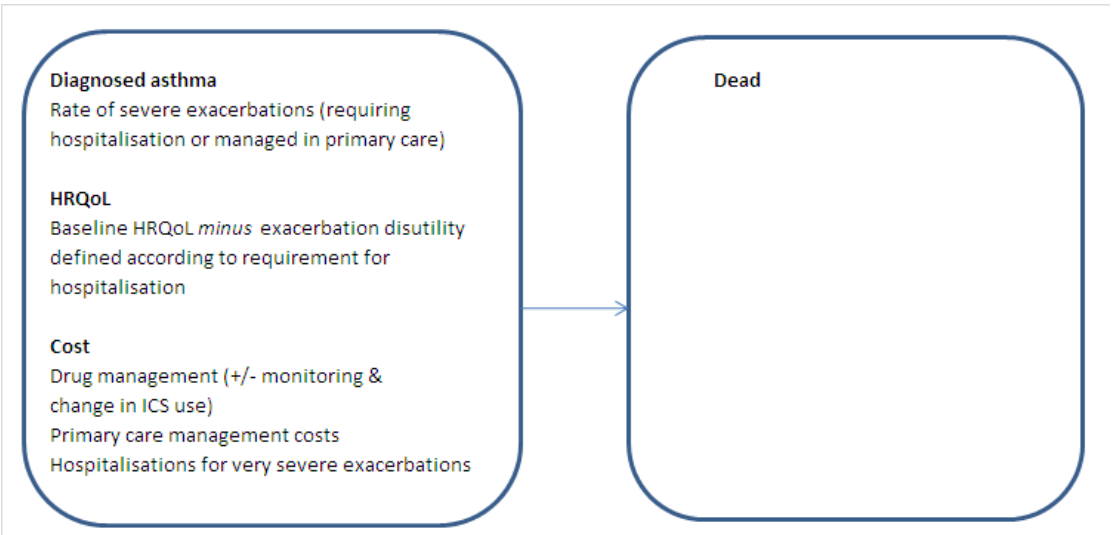


Figure 2: Asthma management model



c) Model inputs

Table 64 (page 283 of the DAR) presents the parameter values, distributions and evidence sources used to inform the two models.

Estimates of test accuracy for diagnostic tests were drawn from a number of separate studies based on the results of the systematic review for clinical effectiveness. Tables 65 and 66 in the DAR summarise the sources for these estimates and the values selected. The economic analyses included estimates of the sensitivity and specificity of individual tests as well as combinations of FeNO plus other standard tests. One study (Schneider et al) which used the NIOX MINO device was used to inform estimates of the sensitivity and specificity of FeNO alone. The EAG later found that this study had included whole body plethysmography as a comparator, a test that is not generally used in UK clinical practice. The EAG have therefore produced a DAR addendum using estimates of the sensitivity and specificity of individual tests as well as combinations of FeNO plus other standard tests from a different study.

The true pre-test probability of asthma in undiagnosed patients was estimated as a weighted mean of a number of cases of asthma and non-asthma in the studies used to inform the diagnostic test accuracy parameters. Across the included studies, 412 of 881 patients were diagnosed with asthma (probability = 0.47). The probability that the patient is male was estimated from two studies.

Health Related Quality of Life values (HRQoL) for people without asthma were estimated using a general population EQ-5D regression model. This (HRQoL) is common to all diagnostic comparator groups and does not therefore, have any effect on the estimates of incremental health gain for the diagnostic tests included in the economic analysis.

The disutility associated with asthma, estimated to be -0.0463 was taken from a catalogue of EQ-5D values reported by Sullivan *et al*. It is noted that this disutility was applied to all patients with asthma and to those who test false-

positive (until their misdiagnosis is corrected). This disutility is unlikely to fully reflect health losses associated with the delayed diagnosis of more serious pathology such as cancer or tuberculosis.

The disutility associated with poor asthma control was derived from a study (McTaggart Cowan) which reported EQ-5D estimates 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 External Assessment Group 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). This disutility is applied to all false negatives until the misdiagnosis is corrected.

Due to the lack of empirical evidence relating to the time required to resolve incorrect diagnoses, the External Assessment Group attempted to elicit these values from clinical experts (page 292 of the DAR). Based on the responses received, the External Assessment Group assumed that the time to resolve a false-negative diagnosis has a mean of 8 months (95% CI 4 to 12months) and the time to resolve a false positive diagnosis has a mean of 18 months (95% CI 12 to 24months). The External Assessment Group considered this estimates to be highly uncertain and tested them extensively the sensitivity analysis.

The following costs were included in the model:

i. Test costs: The marginal per-test cost for all 3 devices were calculated based on information provided by the manufacturers (Table 67, page 294 of the DAR). The calculation was complicated by the fact that the devices each have different lifetimes and test kits for each device are available at lower marginal costs if higher volumes of kits are purchased. These marginal per-test costs do not include any costs associated with education and training for NHS staff to use the devices.

The EAG assumed that the cost of maintenance of NObreath is free of charge to the NHS. The EAG assumed zero maintenance costs for NIOX MINO and NIOX VERO.

ii. GP costs: the EAG assumed that spirometry, reversibility testing and FeNO can be done in primary care and would require two GP visits and one nurse visit. The unit cost of a GP visit was based on published economic analyses that used 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.

For the management model, the External Assessment Group assumed that FeNO monitoring would be undertaken during routine GP visits and would require one additional nurse visit once every 3 months. The marginal cost of FeNO monitoring was applied as the per-test cost plus the cost of a primary care nurse appointment.

iii. Secondary care costs: the External Assessment Group 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. Secondary care attendance costs were based on the HRG for respiratory medicine attendances (£204.29). The cost of a laboratory visit was based on the HRG for simple bronchodilator studies (£203.29). The External Assessment Group assumed standard errors around these estimates were normally distributed with a standard error equal to 15% of the mean.

iv. Costs of asthma management: Estimates of the annual cost of combined inhalers were derived from two previous health technology assessment reports. The least expensive annual cost for combined inhalers was estimated to be £201 for children. For adults, the least expensive annual cost of the inhalers was estimated to be £231.

v. Costs associated with resolving misdiagnoses: A crude assumption was made that one additional primary care attendance, two additional secondary care attendances and one laboratory visit would be required to correctly diagnose false-positive and false negative results. This same assumption was made in previously published models.

vi. Costs associated with loss of control for false-negatives: The External Assessment Group assumed that false negatives would experience one exacerbation in each year they remain misdiagnosed. The model assumes that a proportion of these exacerbations will require hospitalisations.

vii. Costs of asthma management: Estimates of the annual costs of combined inhalers were derived from two previous health technology assessment reports. The least expensive annual cost for combined inhalers in children was estimated to be £201 in children and £231 in adults.

viii. Costs of managing exacerbations: the External Assessment Group assumed that a proportion of exacerbations would require hospitalisation whilst the remainder could be managed in primary care. Severe exacerbations which do not require hospitalisation would require one GP attendance (£43.00) plus oral steroids for 5 days (£1.73) based on an earlier HTA report. The cost of asthma hospitalisation was derived from current NHS Reference Costs (£1,266.72).

Base case results

In the base case, the model was evaluated probabilistically using Monte Carlo sampling techniques. Central estimates of cost-effectiveness are presented as incremental cost-effectiveness ratios (ICERs), cost-effectiveness planes and cost-effectiveness acceptability curves.

a) Diagnostic model results

For 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 is likely to be very small. Airway hyperresponsiveness (MCT) is expected to produce the greatest QALY gain (4.2834) as it has the highest sensitivity and specificity of all tests included in the analysis. FEV1/FVC is expected to produce the lowest QALY gain (4.2686).

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 of NObreath. All other options except FeNO (NObreath) + bronchodilator reversibility are expected to be dominated. The expected ICER of airway hyperresponsiveness (MCT) versus FeNO (NObreath) + bronchodilator reversibility is approximately £1.125 million per QALY gained. Table 69, page 306 of the DAR presents the central estimates of cost-effectiveness for diagnosis.

Uncertainty analysis was undertaken for the diagnostic options. Figure 28, page 307 of the DAR shows the cost-effectiveness acceptability curves. Assuming willingness-to-pay thresholds of £20,000 and £30,000 per QALY gained, FeNO (NObreath) + bronchodilator reversibility has the highest probability of producing the greatest amount of net benefit (probability= 0.98 and 0.95 respectively). Table 70, page 308 of the DAR summarises the probability of optimality values.

b) Management model results (children)

Table 77, page 316 of the DAR presents the central estimates of cost-effectiveness based on the probabilistic version of the management model in children. The results suggest that FeNO testing is expected to produce a small health benefit (0.05 QALYs) compared to guidelines. FeNO testing is also more costly (£8148.59 for NoBreath, £8,314.30 for NIOX VERO and £8,391.53 for NIOX MINO) than guidelines (£5,860.06) due to projected inhaled corticosteroid use for the FeNO groups.

NIOX MINO and NIOX VERO are expected to be dominated by NObreath due to their higher marginal per-test cost. The expected ICER of guidelines + NObreath versus guidelines alone is approximately £45,213 per QALY gained.

The cost effectiveness acceptability curve for management of asthma in children is presented in figure 30, page 316 of the DAR. Assuming a willingness to pay thresholds of £20,000 and £30,000 per QALY gained, guidelines alone has the highest probability of producing the greatest amount of net benefit (probability = 0.99 and 0.91 respectively).

c) Management model results (adults)

Table 80, page 319 of the DAR presents the central estimates of cost-effectiveness based on the probabilistic version of the management model in adults. The results suggest that FeNO testing is expected to produce a small health benefit (0.04 QALYs) compared to guidelines.

NIOX MINO and NIOX VERO are dominated by NObreath due to their higher marginal per-test cost. The ICER of guidelines + NObreath versus guidelines alone is approximately £2,146 per QALY gained. If dominance is ignored, the ICERs for the NIOX devices compared to guideline are £6,310 per QALY gained for NIOX VERO and £8,250 per QALY gained for NIOX MINO.

The cost effectiveness acceptability curve for management of asthma in adults is presented in figure 32, page 321 of the DAR. Assuming a willingness to pay thresholds of £20,000 and £30,000 per QALY gained, guidelines + NObreath strategy has the highest probability of producing the greatest amount of net benefit (probability = 0.82 and 0.87 respectively).

Deterministic sensitivity analysis

A large number of deterministic sensitivity analyses were undertaken for the diagnostic and management models.

a) Deterministic sensitivity analyses using the diagnostic model

Table 1 below provides details of the deterministic sensitivity analyses undertaken for the diagnostic model. The results of the analyses are presented in tables 71 to 76 (pages 309 to 314 of the DAR). The main indications from the results of the deterministic sensitivity analyses are:

- The cost-effectiveness frontier presented in the base case analysis (which includes only airway hyperresponsiveness and FeNO plus bronchodilator reversibility) is maintained across the majority of scenarios. 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.

Table 1: details of the deterministic sensitivity analyses undertaken for the diagnostic model

Parameter	Scenario	Description
Point estimates for parameters	DSA scenario D1	The model was evaluated using point estimate parameters instead of the expectation of the mean.
Alternative discount rates	DSA scenario D2	Model was evaluated using a 0% discount rate for costs and QALYs
	DSA scenario D3	Model was evaluated using a 6% discount rate for costs and QALYs
All tests undertaken in secondary care	DSA scenario D4	The model was run assuming all tests were undertaken in a secondary setting
Alternative asthma control disutilities for false-negatives	DSA scenario D5	The model was run using a disutility value of -0.10.
	DSA scenario D6	The model was run using a disutility value of -0.41.
Alternative disutilities for false positives	DSA scenario D7	Base case disutility applied to false-positives was doubled.
	DSA scenario D8	Base case disutility applied to false-positives was halved.
FeNO test costs	DSA scenario D9	Marginal per-test costs for all FeNO devices were doubled.
	DSA scenario D10	Marginal per-test costs for all FeNO devices were halved.
Alternative assumptions concerning NObreath device lifetime	DSA scenario D11	The maximum lifetime of NObreath is assumed to be equal that of NIOX MINO (3 years) resulting in a marginal per-test cost of £14.32.
	DSA scenario D12	The maximum lifetime of NObreath is assumed to be equal that of NIOX VERO (5 years) resulting in a marginal per-test cost of £8.88.
	DSA scenario D13	The maximum lifetime of NObreath is assumed to be 20 years (double that of the base case) resulting in a marginal per-test cost of £2.32.
Test visit costs	DSA scenario D14	All primary and secondary care visit costs (including costs of initial visits and subsequent visits to resolve misdiagnosis) were doubled.
	DSA scenario D15	All primary and secondary care visit costs (including costs of initial visits and subsequent visits to resolve misdiagnosis) were halved.
False-negative exacerbation rate	DSA scenario D16	Base case incremental exacerbation rate for false-negatives is doubled.
	DSA scenario D17	Base case incremental exacerbation rate for false-negatives is halved.
Asthma treatment costs	DSA scenario D18	Asthma treatment tests are doubled
	DSA scenario D19	Asthma treatment costs are halved
Time to resolve misdiagnosis	DSA scenario D20	The model was evaluated using time to resolve diagnosis 2x
	DSA scenario D21	The model was evaluated using time to resolve diagnosis 3x
	DSA scenario D22	The model was evaluated using time to resolve diagnosis 4x

	DSA scenario D23	The model was evaluated using time to resolve diagnosis 5x
	DSA scenario D24	The model was evaluated using time to resolve diagnosis 10x
	DSA scenario D25	The model was evaluated using time to resolve diagnosis 0.5x
Alternative sources for diagnostic accuracy of FeNO alone	DSA scenario D26	The model was evaluated using alternative estimates at a cutoff of 34ppb; sensitivity was 35% and specificity 95%.
	DSA scenario D27	The model was evaluated using alternative estimates at a cutoff of 40ppb; sensitivity was 74% and specificity 73%.
Alternative sources for diagnostic accuracy of non-FeNO comparators	DSA scenario D28	The model was evaluated using estimates for FEV_1/FVC, PEF and sputum induction from a different diagnostic study report.
"Rule out" diagnostic decision report	DSA scenario D29	Base case test characteristics for FeNO options.
	DSA scenario D30	Best sensitivity for FeNO options.
	DSA scenario D31	Best specificity for FeNO options
"Rule in" diagnostic decision report	DSA scenario D32	Base case test characteristics for FeNO options.
	DSA scenario D33	Best sensitivity for FeNO options.
	DSA scenario D34	Best specificity for FeNO options

b) Deterministic sensitivity analyses using the management model (children)

Table 79, page 318 of the DAR presents the scenarios and results for the deterministic sensitivity analyses for the management model in children. The results of the analyses indicate the following:

- 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 inhaled corticosteroid 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 inhaled corticosteroid 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.

c) Deterministic sensitivity analyses using the management model (adults)

Table 82, page 322 of the DAR, DAR presents the scenarios and results for the deterministic sensitivity analyses for the management model in adults. The results of the 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 inhaled corticosteroids 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 inhaled corticosteroid use for FeNO and guidelines at the last observed time point in the trials.
- The cost-effectiveness of FeNO monitoring is markedly less favourable when projected inhaled corticosteroid use is modelled according to the mean inhaled corticosteroid use observed in the trial reported by Shaw *et al*.

4. Issues for consideration

1. 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.

2. There is no clear evidence to suggest equivalence between the FeNO devices. The review undertaken by the external assessment group suggests that whilst there was often good correlation between FeNO measurement devices, equivalence of readings could not necessarily be assumed in all situations. Many studies concluded that the comparability of measurements between devices was within clinically acceptable limits; however, 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.

3. There was a high degree of correlation between measurements across all devices, suggesting that estimates of sensitivity and specificity are likely to be a reasonable indication of potential diagnostic accuracy of using FeNO to guide diagnosis and management. The derived cut-off points are however, not likely to be interchangeable between devices. The external assessment group assumed, for the purpose of this assessment, that sensitivities and specificities between devices are interchangeable. The cut-off points that should be used to achieve them however, cannot be assumed to be the same in each device. In addition, there is doubt as to whether the same diagnostic accuracy would be achievable in all devices.

4. From the limited data available, the External Assessment Group were able to observe 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.

5. The addition of another test to the diagnostic protocol using FeNO resulted in a change in diagnostic accuracy, but as this involved the usual trade-off between sensitivity and specificity the External Assessment Group concluded that it is difficult to tell if this represents an increase or decrease in clinical and cost-effectiveness.

6. There was limited evidence of in the subgroups defined a priori, namely pregnancy, the elderly and smokers/environmental tobacco exposure.

7. Four studies in adults were identified for FeNO guided management of asthma in adults. 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 did report 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. Despite the heterogeneity in results, and the lack of statistically significant findings the External Assessment Group were able to conclude that, in adults, FeNO-guided management 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 inhaled corticosteroid use. There was no evidence relating to whether these effects would be maintained over a longer time period.

8. For FeNO guided management of asthma in children, all included studies reported a decrease in exacerbations in the intervention arm, but only one reported a statistically significant reduction. The effects on inhaled corticosteroid use were heterogeneous. 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. The External Assessment Group, however, concluded that FeNO-guided management of is likely, during the first year of management, to result in non-statistically significant trends towards better management overall.

9. In both the diagnostic and management models 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. It appears that the costs of the devices appear to be the main driver of the ICERs, and the difference in cost is being driven largely by the lifetime of the device. NoBreath is assumed to have a lifetime of 10 years, NIOX VERO, 5 years and NIOX MINO, 3 years. The results of the sensitivity analysis show that when assuming different lifetimes for NoBreath (3, 5 and 20 years), NoBreath is always the most cost effective option.

10. The results of the analysis in the diagnostic model 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.

11. For the management model, the results in the children and adult subgroups are particularly sensitive to assumptions regarding changes in inhaled corticosteroids 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 inhaled corticosteroids use.

12. 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 study151 does however mean that all non-FeNO diagnostic options are assessed consistently within the same study.

13. 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.

14. 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.

5. Equality considerations

Asthma is more common in boys than girls, pre-puberty. The risk of persistent asthma is higher for girls in the transition from childhood to adulthood. Some people with severe asthma may be covered by equalities legislation, for example, where this affects activities of daily living. During scoping, it was suggested that the characteristics of asthma may be different in groups protected by equalities legislation, for instance women during pregnancy and older people.

6. Authors

Farouk Saeed
Topic Lead

Pall Jonsson
Technical Adviser

October 2013

Appendix A: Sources of evidence considered in the preparation of the overview

The diagnostics assessment report for this assessment was prepared by ScHARR, University of Sheffield:

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.* 2013.

The following organisations accepted the invitation to participate in this assessment as stakeholders. They were invited to attend the scoping workshop and to comment on the diagnostics assessment report.

I Manufacturers/sponsors:

- Aerocrine Ltd.
- Bedfont Scientific Ltd.
- Intermedical UK Ltd.

II Professional/specialist and patient/carer groups:

- Association of Respiratory Nurse Specialists (ARNS)
- Asthma UK
- Healthcare Improvement Scotland
- Research in Real Life Ltd.
- Royal College of Nursing
- The Royal College of Pathologists