National Guideline Centre

Final version

Asthma

Asthma: diagnosis and monitoring of asthma in adults, children and young people

NICE guideline NG80

Methods, evidence and recommendations

November 2017

Final for publication

Commissioned by the National Institute for Health and Care Excellence

Royal College of Physicians

The Royal College of Surgeons of England

RCGP Royal College of General Practitioners

Royal College of Nursing

NICE accredited

Disclaimer
Healthcare professionals are expected to take NICE clinical guidelines fully into account when exercising their clinical judgement. However, the guidance does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of each patient, in consultation with the patient and/or their guardian or carer.

Funding
National Institute for Health and Care Excellence

© NICE 2017. All rights reserved. Subject to Notice of rights.

Contents

Guideline committee members

Name	Role
John Alexander	Consultant in Paediatrics & Paediatric Intensive Care, University Hospitals of North Midlands NHS Trust
Tara Burn	Patient/carer member
Erol Gaillard	Senior Lecturer & Honorary Consultant in Paediatric Respiratory Medicine, University of Leicester & University Hospitals of Leicester NHS Trust
Catherine Lawlor	Advanced Nurse Practitioner, Surrey Docks Health Centre (until December 2014); Senior Lecturer in Adult Nursing, University of Greenwich (from December 2014)
Val Hudson	Patient member
Angela Key	Chief Respiratory Physiologist, Aintree University Hospital NHS Foundation Trust
Matthew Masoli	Consultant Respiratory Physician, Plymouth Hospitals NHS Trust
Melanie McFeeters	Consultant Nurse for Children's Respiratory Diseases, University Hospitals of Leicester NHS Trust (until April 2014); Clinical Service Specialist for Specialised Commissioning in the East Midlands (from April 2014 to April 2015) and Quality Lead for Specialised Commissioning (from April 2015), NHS England
Andrew Menzies-Gow (GC Chair)	Consultant Respiratory Physician, Royal Brompton & Harefield NHS Foundation Trust
Georgina Russell	RCP Clinical Fellow (until January 2014)
Tahmina Siddiqui	GP Partner, Whaddon Medical Centre, Milton Keynes
Michael Thomas	Professor of Primary Care Research, University of Southampton

NGC technical team members

Name	Role
Emily Davies	Senior Research Fellow (from September 2013)
Elisabetta Fenu	Health Economics Lead (until November 2013 and from September 2014)
Alexander Haines	Health Economist (from October 2013)
Bernard Higgins	Clinical Director
Emma Madden	Research Fellow (from June 2014)
Paul Miller	Senior Information Scientist (until September 2014)
Rachel O'Mahony	Senior Research Fellow (until May 2014)
Jill Parnham	Operations Director (until September 2014)
Amelia Unsworth	Project Manager (from November 2013)
Giulia Zuodar	Document Editor/Process Assistant (from June 2014)

Acknowledgements

The development of this guideline was greatly assisted by the following people:

- Shaleen Ahmad, Clinical Research Assistant, Cochrane Airways Group
- Alex Allen, Information Scientist, NGC
- Joanna Ashe, Senior Information Scientist, NGC
- Katie Broomfield, Document Editor/Process Assistant, NGC
- Annie Burden, Senior Statistician, Research in Real Life (RiRL) UK
- Hannah Carré, Project Administrator, NGC
- Christopher Cates, Senior Clinical Research Fellow, Cochrane Airways Group
- Saskia Cheyne, Project Manager, NGC
- Jill Cobb, Information Scientist, NGC
- Louise Craig, Research Fellow, NIHR Evaluation, Trials and Studies Coordinating Centre (NETSCC)
- Jackie Donovan, Principal Biochemist, Royal Brompton & Harefield NHS Foundation Trust
- Katharina Dworzynski, Senior Research Fellow, NGC
- Andreas Freitag, Senior Research Fellow, NGC
- Saoussen Ftouh, Senior Research Fellow/Project Manager, NGC
- Lina Gulhane, Joint Head of Information Science & Knowledge Manager, NGC
- Martin Harker, Senior Health Economist, NGC
- Kate Kelley, Operations Director, NGC
- Kayleigh Kew, Research Assistant, Cochrane Airways Group
- Grace Marsden, Senior Health Economist, NGC
- Ben Mayer, Research Fellow, NGC
- Jacoby Patterson, Systematic Reviewer
- Elizabeth Pearton, Information Scientist, NGC
- Nancy Pursey, Senior Project Manager, NGC
- Julie Robinson, Information Scientist, NGC
- Carlos Sharpin, Joint Head of Information Science/Research Fellow, NGC
- Thomas Strong, Document Delivery Assistant, NGC
- Sharon Swain, Associate Director, NGC
- Simon Ward, Head of Lung Function, Royal Brompton & Harefield NHS Foundation Trust
- Maggie Westby, Clinical Effectiveness Lead, NGC
- David Wonderling, Head of Health Economics, NGC

1 Introduction

Asthma is a chronic inflammatory respiratory disease. It can affect people of any age, but often starts in childhood. Asthma is a variable disease which can change throughout a person's life, throughout the year and from day to day. It is characterised by attacks (also known as exacerbations) of breathlessness and wheezing, with the severity and frequency of attacks varying from person to person. The attacks are associated with variable airflow obstruction and inflammation within the lungs, which if left untreated can be life-threatening, however with the appropriate treatment can be reversible.

In 2013, the World Health Organization estimated that 235 million people had asthma worldwide. It is the most common chronic condition to affect children, and in the UK approximately 5.4 million people (1.1 million children and 4.3 million adults) currently get treatment for asthma[8,67,68].

The causes of asthma are not well understood. A number of risk factors are associated with the condition, often in combination. These influences can be genetic (the condition clusters in families) and/or environmental (such as inhalation of allergens or chemical irritants). Occupational causes of asthma in adults are often under-recognised.

There is currently no gold standard test available to diagnose asthma; diagnosis is principally based on a thorough history taken by an experienced clinician. Studies of adults diagnosed with asthma suggest that up to 30% do not have clear evidence of asthma[1,95,104,155,169]. Some may have had asthma in the past, but it is likely that many have been given an incorrect diagnosis. Conversely, other studies suggest that asthma may be underdiagnosed in some cases. This guideline will improve patient outcomes and is cost-effective to the NHS in the long-term; NICE's cost impact assessment projects a saving of approximately £12 million per year in England, before implementation costs.

The typical wheeze found in a person with asthma is a continuous, polyphonic whistling sound produced in the airways during expiration and is related to obstruction of the airways on breathing out. Expiratory polyphonic wheeze is a characteristic clinical symptom and sign in people with asthma or other obstructive airways diseases.

Initial clinical assessment should include questions about symptoms (wheezing, cough, breathing and chest problems) and any personal or family history of allergies, atopic disorders or asthma. Various tests can be used to support a diagnosis, but there is no single test that can definitively diagnose asthma.

A number of methods and assessments are available to determine the likelihood of asthma. These include measuring airflow obstruction (spirometry and peak flow) and assessment of reversibility with bronchodilators, with both methods being widely used in current clinical practice. However, normal results do not exclude asthma and abnormal results do not always mean it is asthma, as they could be indicators of other respiratory diseases or spurious readings.

Testing for airway inflammation is increasingly used as a diagnostic strategy in clinical practice. This includes measuring fractional exhaled nitric oxide (FeNO).

Other diagnostic strategies include blood or skin prick tests to detect allergic reactions to environmental influences, exercise tests to detect evidence of bronchoconstriction, and measures of airway hyper-reactivity such as histamine/methacholine or mannitol challenge tests. However, it is debatable which test or measure, or combination of them, is the most effective to accurately diagnose asthma.

It is recognised that asthma control is suboptimal in many people with asthma. This has an impact on their quality of life, their use of healthcare services and the associated costs. Asthma control can be

monitored by measuring airway obstruction or inflammation and by using validated questionnaires, but the most effective monitoring strategy is unclear.

The aim of this guideline is, therefore, to determine the most clinical and cost-effective way to effectively diagnose people with asthma and determine the most effective monitoring strategy to ensure optimum asthma control.

The scope of this guideline covers the diagnosis for people presenting with new symptoms of suspected asthma and the monitoring of asthma and excludes other aspects of management. It is not intended to be used to re-diagnose every person with an asthma diagnosis in England.

This guideline covers infants and young children 0–5 years old, children 5–16 years old and adults and young people over the age of 16 who are being investigated for suspected asthma, or who have been diagnosed with asthma and are having their condition monitored. The guideline applies to all primary, secondary and community care settings in which NHS-funded care is provided for people with asthma.

This guideline does not cover the diagnosis and monitoring of people with severe, difficult to control asthma.

This guideline offers best practice advice on the care of people with suspected asthma presenting with respiratory symptoms, and monitoring asthma control in people with a confirmed diagnosis of asthma. Chapters 6 to 10 review the diagnostic accuracy of the initial clinical assessment questions for the diagnosis of asthma in people with suspected asthma presenting with respiratory symptoms. Chapters 11 to 20 review the diagnostic test accuracy of objective tests for the diagnosis of asthma in people with suspected asthma presenting with respiratory symptoms. The diagnostic pathway can be found in section 4.1. Chapters 23 to 30 review the clinical and cost-effectiveness of interventions used to monitor asthma control.

Patients and healthcare professionals have rights and responsibilities as set out in the NHS Constitution for England – all NICE guidance is written to reflect these. Treatment and care should take into account individual needs and preferences. Patients should have the opportunity to make informed decisions about their care and treatment, in partnership with their healthcare professionals. If the patient is under 16, their family or carers should also be given information and support to help the child or young person to make decisions about their treatment. Healthcare professionals should follow the Department of Health's advice on consent. If someone does not have capacity to make decisions, healthcare professionals should follow the code of practice that accompanies the Mental Capacity Act and the supplementary code of practice on deprivation of liberty safeguards. In Wales, healthcare professionals should follow advice on consent from the Welsh Government.

NICE has produced guidance on the components of good patient experience in adult NHS services. All healthcare professionals should follow the recommendations in Patient experience in adult NHS services.

1.1 Phased implementation

NICE is recommending objective testing with spirometry and FeNO for most people with suspected asthma. This is a significant enhancement to current practice, which will take the NHS some time to implement, with additional infrastructure and training needed in primary care. New models of care, being developed locally, could offer the opportunity to implement these recommendations. This may involve establishing diagnostic hubs to make testing efficient and affordable. They will be able to draw on the positive experience of NICE's primary care pilot sites, which trialled the use of FeNO.

The investment and training required to implement the new guidance will take time. In the meantime, primary care services should implement what they can of the new guidelines, using currently available approaches to diagnosis until the infrastructure for objective testing is in place.

2 Development of the guideline

2.1 What is a NICE clinical guideline?

NICE clinical guidelines are recommendations for the care of individuals in specific clinical conditions or circumstances within the NHS – from prevention and self-care through primary and secondary care to more specialised services. We base our clinical guidelines on the best available research evidence, with the aim of improving the quality of health care. We use predetermined and systematic methods to identify and evaluate the evidence relating to specific review questions.

NICE clinical guidelines can:

- provide recommendations for the treatment and care of people by health professionals
- be used to develop standards to assess the clinical practice of individual health professionals
- be used in the education and training of health professionals
- help patients to make informed decisions
- improve communication between patients and health professionals.

While guidelines assist the practice of healthcare professionals, they do not replace their knowledge and skills.

We produce our guidelines using the following steps:

- Guideline topic is referred to NICE from the Department of Health.
- Stakeholders register an interest in the guideline and are consulted throughout the development process.
- The scope is prepared by the National Guideline Centre (NGC).
- The NGC establishes a Guideline committee.
- A draft guideline is produced after the group assesses the available evidence and makes recommendations.
- There is a consultation on the draft guideline.
- The final guideline is produced.

The NGC and NICE produce a number of versions of this guideline:

- the 'full guideline' contains all the recommendations, plus details of the methods used and the underpinning evidence
- the 'NICE guideline' lists the recommendations
- 'information for the public' is written using suitable language for people without specialist medical knowledge
- NICE Pathways brings together all connected NICE guidance.

This version is the full version. The other versions can be downloaded from NICE at www.nice.org.uk.

2.2 Remit

NICE received the remit for this guideline from the Department of Health. They commissioned the NGC to produce the guideline.

The original remit for this guideline was:

'to prepare a guideline on the diagnosis and management of asthma'.

It was later decided during the scoping process to cover 'diagnosis and monitoring of asthma' in this guideline, and to cover 'management of asthma' as the subject of a separate guideline.

2.3 Who developed this guideline?

The group includes health professionals and researchers as well as lay members.

A multidisciplinary Guideline committee (GC) comprising health professionals and researches as well as lay members developed this guideline (see the list of Guideline committee members and the acknowledgements).

The National Institute for Health and Care Excellence (NICE) funds the National Guideline Centre (NGC) and thus supported the development of this guideline. The GC was convened by the NGC and chaired by Dr Andrew Menzies-Gow in accordance with guidance from NICE.

The group met every 5-6 weeks during the development of the guideline. At the start of the guideline development process all GC members declared interests including consultancies, fee-paid work, share-holdings, fellowships and support from the healthcare industry. At all subsequent GC meetings, members declared arising conflicts of interest.

Members were either required to withdraw completely or for part of the discussion if their declared interest made it appropriate. The details of declared interests and the actions taken are shown in Appendix B.

Staff from the NGC provided methodological support and guidance for the development process. The team working on the guideline included a project manager, systematic reviewers, health economists and information scientists. They undertook systematic searches of the literature, appraised the evidence, conducted meta-analysis and cost-effectiveness analysis where appropriate and drafted the guideline in collaboration with the GC.

(a) What this guideline covers

This guideline covers adults, children and young people who are being investigated for suspected asthma, or who have been diagnosed with asthma and are having their condition monitored. For further details please refer to the scope in Appendix A and the review questions in chapters 6 to 30.

(b) What this guideline does not cover

This guideline does not cover:

- severe or difficult-to-control asthma
- treating asthma.

For further details please refer to the scope in Appendix A.

(c) Relationships between the guideline and other NICE guidance

Related NICE clinical guidelines:

- Bronchiolitis in children. NICE guideline 9 (2015).
- Patient experience in adult NHS services. NICE clinical guideline 138 (2012).
- Chronic obstructive pulmonary disease (updated). NICE clinical guideline 101 (2009).
- Medicines adherence. NICE clinical guideline 76 (2009).
- Respiratory tract infections. NICE clinical guideline 69 (2008).

Related NICE diagnostic assessment guidance:

- Measuring fractional exhaled nitric oxide concentration in asthma: NIOX MINO, NIOX VERO and NObreath. NICE diagnostics guidance 12 (2014).

Related NICE interventional procedures guidance:

- Bronchial thermoplasty for severe asthma. NICE interventional procedure guidance 419 (2012).

Related NICE quality standards:

- Quality standard for asthma. NICE quality standard 25 (2013).

Related NICE technology appraisals:

- Roflumilast for treating chronic obstructive pulmonary disease. NICE technology appraisal guidance 461 (2017).
- Omalizumab for the treatment of severe persistent allergic asthma in children aged 6 and over and adults (review of TA133 and TA201) NICE technology appraisal guidance 278 (2013).
- Inhaled corticosteroids for the treatment of chronic asthma in adults and in children aged 12 years and over. NICE technology appraisal guidance 138 (2008).
- Inhaled corticosteroids for the treatment of chronic asthma in children under the age of 12 years. NICE technology appraisal guidance 131 (2007).
- Inhaler devices for routine treatment of chronic asthma in older children (aged 5–15 years). NICE technology appraisal guidance 38 (2002).
- Guidance on the use of inhaler systems (devices) in children under the age of 5 years with chronic asthma. NICE technology appraisal guidance 10 (2000).

Related NICE guidance currently in development:

- Asthma management NICE guideline. Publication expected 2017.

3 Methods

This chapter sets out in detail the methods used to review the evidence and to generate the recommendations that are presented in subsequent chapters. This guidance was developed in accordance with the methods outlined in the NICE guidelines manual 2012[114].

3.1 Developing the review questions and outcomes

Review questions were developed in a PICO framework (patient, intervention, comparison and outcome) for intervention reviews in a framework of population, index tests, reference standard and target condition for reviews of diagnostic test accuracy; and using population, presence or absence of factors under investigation (for example prognostic factors) and outcomes for prognostic reviews.

This use of a framework guided the literature searching process, critical appraisal and synthesis of evidence, and facilitated the development of recommendations by the Guideline committee (GC). The review questions were drafted by the NGC technical team and refined and validated by the GC. The questions were based on the key clinical areas identified in the scope (Appendix A).

A total of 23 review questions were identified.

Full literature searches, critical appraisals and evidence reviews were completed for all the specified review questions.

Table 1: Review questions

Chapter	Type of review	Review questions	Outcomes
6	Diagnostic	In people under investigation for asthma, what is the diagnostic accuracy of each of the following signs and symptoms? • wheezing • cough • breathlessness • nocturnal symptoms • diurnal and seasonal variations	Critical outcomes: • Sensitivity (%) and specificity (%)
7	Diagnostic	In people under investigation for asthma, what is the diagnostic accuracy of taking a personal/family history of atopic disorders?	Critical outcomes: • Sensitivity (%) and specificity (%)
8	Diagnostic	In people under investigation for asthma, what is the diagnostic accuracy of a clinical history of symptoms in response to exercise?	Critical outcomes: • Sensitivity (%) and specificity (%)
9	Diagnostic	In people under investigation for asthma, what is the diagnostic accuracy of a clinical history of symptoms after taking the following drugs: a) in adults - beta blockers, aspirin, or other NSAIDs b) in children – ibuprofen?	Critical outcomes: • Sensitivity (%) and specificity (%)
10	Diagnostic	In adults under investigation for occupational asthma, what is the diagnostic accuracy for case identification, of asking whether their symptoms are better away from work?	Critical outcomes: • Sensitivity (%) and specificity (%)
11	Diagnostic	In people under investigation for asthma,	Critical outcomes:

Chapter	Type of review	Review questions	Outcomes
		what is the diagnostic test accuracy and cost-effectiveness of spirometry / flow volume loop measures?	• Sensitivity (%) and specificity (%) at pre-specified thresholds
12	Diagnostic	In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of bronchodilator response (using PEF or FEV1)?	Critical outcomes: • Sensitivity (%) and specificity (%) at pre-specified thresholds
13	Diagnostic	In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of peak expiratory flow (PEF) variability?	Critical outcomes: • Sensitivity (%) and specificity (%) at pre-specified thresholds
14	Diagnostic	In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of skin prick tests?	Critical outcomes: • Sensitivity (%) and specificity (%) at pre-specified thresholds
15	Diagnostic	In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of total and specific serum IgE measures?	Critical outcomes: • Sensitivity (%) and specificity (%) at pre-specified thresholds
16	Diagnostic	In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of fractional exhaled nitric oxide (FeNO) measures?	Critical outcomes: • Sensitivity (%) and specificity (%) at pre-specified thresholds
17	Diagnostic	In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of eosinophil blood count measures?	Critical outcomes: • Sensitivity (%) and specificity (%) at pre-specified thresholds
18	Diagnostic	In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of airway hyper-reactivity (non-specific bronchial challenge) with histamine and methacholine?	Critical outcomes: • Sensitivity (%) and specificity (%) at pre-specified thresholds
19	Diagnostic	In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of airway hyper-reactivity (non-specific bronchial challenge) with mannitol?	Critical outcomes: • Sensitivity (%) and specificity (%) at pre-specified thresholds
20	Diagnostic	In people under investigation for asthma, what is the diagnostic accuracy of bronchoconstriction in response to an exercise challenge?	Critical outcomes: • Sensitivity (%) and specificity (%) at pre-specified thresholds
23	Intervention	In people with asthma, what is the clinical and cost-effectiveness of using symptom scores / diaries or validated questionnaires measuring symptom control (eg ACT, ACQ, CACT, RCP 3 questions) and / or health related quality of life (eg AQLQ, PAQLQ) to monitor asthma?	Critical outcomes • Mortality • Unscheduled healthcare utilisation (UHU) • Exacerbations (defined as need for course of oral steroids) • Asthma control questionnaires • QoL

Chapter	Type of review	Review questions	Outcomes
			Important outcomes • Lung function • Symptoms • Dose of regular preventer medication (ICS dose) • Rescue medication (SABA use) • Time off school or work
24	Intervention	In people with asthma, what is the clinical and cost-effectiveness of using measures of pulmonary function assessing asthma control (for example, spirometry and peak expiratory flow) to monitor asthma?	Critical outcomes • Mortality • UHU • Exacerbations (defined as need for course of oral steroids) • Asthma control questionnaires • QoL Important outcomes • Lung function • Symptoms • Dose of regular preventer medication (ICS dose) • Rescue medication (SABA use) • Time off school or work
25	Intervention	In people with asthma, what is the clinical and cost-effectiveness of using fractional exhaled nitric oxide (FeNO) measures for monitoring asthma control?	Critical outcomes • Mortality • UHU • Exacerbations (defined as need for course of oral steroids) • Asthma control questionnaires • QoL Important outcomes • Lung function • Symptoms • Dose of regular preventer medication (ICS dose) • Rescue medication (SABA use) • Time off school or work
26	Intervention	In people with asthma, what is the clinical and cost-effectiveness of using the peripheral blood eosinophil count for monitoring asthma control?	Critical outcomes • Mortality • UHU • Exacerbations (defined as need for course of oral steroids)

Chapter	Type of review	Review questions	Outcomes
			• Asthma control questionnaires • QoL Important outcomes • Lung function • Symptoms • Dose of regular preventer medication (ICS dose) • Rescue medication (SABA use) • Time off school or work
27	Intervention	In people with asthma, what is the clinical and cost-effectiveness of using indirect challenge tests with mannitol or direct challenge tests with histamine or methacholine for monitoring asthma control?	Critical outcomes • Mortality • UHU • Exacerbations (defined as need for course of oral steroids) • Asthma control questionnaires • QoL Important outcomes • Lung function • Symptoms • Dose of regular preventer medication (ICS dose) • Rescue medication (SABA use) • Time off school or work
28	Intervention	In people with asthma, what is the clinical and cost-effectiveness of monitoring adherence to treatment?	Critical outcomes • Mortality • UHU • Exacerbations (defined as need for course of oral steroids) • Asthma control questionnaires • QoL • Adherence Important outcomes • Lung function • Symptoms • Dose of regular preventer medication (ICS dose) • Rescue medication (SABA use) • Time off school or work
29	Intervention	In people with asthma, what is the optimal frequency and method for monitoring inhaler	Critical outcomes • Mortality

Chapter	Type of review	Review questions	Outcomes
		technique?	• UHU • Exacerbations (defined as need for course of oral steroids) • Asthma control questionnaires • QoL Important outcomes • Lung function • Symptoms • Dose of regular preventer medication (ICS dose) • Rescue medication (SABA use) • Time off school or work
30	Intervention	In people with asthma, what is the clinical and cost-effectiveness of tele-healthcare to monitor asthma control?	Critical outcomes • Mortality • Exacerbations requiring hospitalisation • Exacerbations (defined as need for course of oral steroids) • UHU • QOL • Asthma Control Questionnaires Important outcomes • Lung function • Symptoms

3.2 Searching for evidence

3.2.1 Clinical literature search

Systematic literature searches were undertaken to systematically identify all published clinical evidence relevant to the review questions. Searches were undertaken according to the parameters stipulated within The guidelines manual 2012.[114] Databases were searched using relevant medical subject headings, free-text terms and study-type filters where appropriate. Studies published in languages other than English were not reviewed. Where possible, searches were restricted to articles published in English. All searches were conducted in MEDLINE, Embase, and The Cochrane Library. All searches were updated on 1 October 2014. No papers published after this date were considered.

Search strategies were quality assured by cross-checking reference lists of highly relevant papers, analysing search strategies in other systematic reviews, and asking GC members to highlight any additional studies. The questions, the study types applied, the databases searched and the years covered can be found in Appendix F.

The titles and abstracts of records retrieved by the searches were sifted for relevance, with potentially significant publications obtained in full text. These were assessed against the inclusion criteria.

During the scoping stage, a search was conducted for guidelines and reports on the websites listed below and on organisations relevant to the topic. Searching for grey literature or unpublished literature was not undertaken. All references sent by stakeholders were considered.

- Guidelines International Network database (www.g-i-n.net)
- National Guideline Clearing House (www.guideline.gov)
- National Institute for Health and Care Excellence (NICE) (www.nice.org.uk)
- NICE evidence search (www.evidence.nhs.uk).

3.2.2 Health economic literature search

Systematic literature searches were also undertaken to identify health economic evidence within published literature relevant to the review questions. The evidence was identified by conducting a broad search relating to asthma in the NHS Economic Evaluation Database (NHS EED), the Health Technology Assessment database (HTA) and the Health Economic Evaluations Database (HEED) with no date restrictions. Additionally, the search was run on MEDLINE and Embase using a specific economic filter, from 2012 to ensure recent publications that had not yet been indexed by the economic databases were identified. Studies published in languages other than English were not reviewed. Where possible, searches were restricted to articles published in English.

The health economic search strategies are included in Appendix F. All searches were updated on 1 October 2014. No papers published after this date were considered.

3.3 Updated searches 2017

The systematic literature searches for all the review questions were rerun in March 2017. The titles and abstracts of records retrieved by the searches were sifted for relevance, with potentially significant publications obtained in full text. These were assessed against the protocol inclusion criteria for appropriate review questions and in total 9 studies were identified. Appendix R: 'Summary of evidence from 2017 update of Asthma: diagnosis and monitoring' outlines the studies that were identified and the questions and recommendations they were relevant to. The update summary details the impact the identified evidence has on the guideline recommendations; the GC considered that none of the studies had a significant impact on the evidence base and would not lead to a change in recommendations. The studies were therefore not used to update the guideline evidence base.

3.4 Evidence of effectiveness

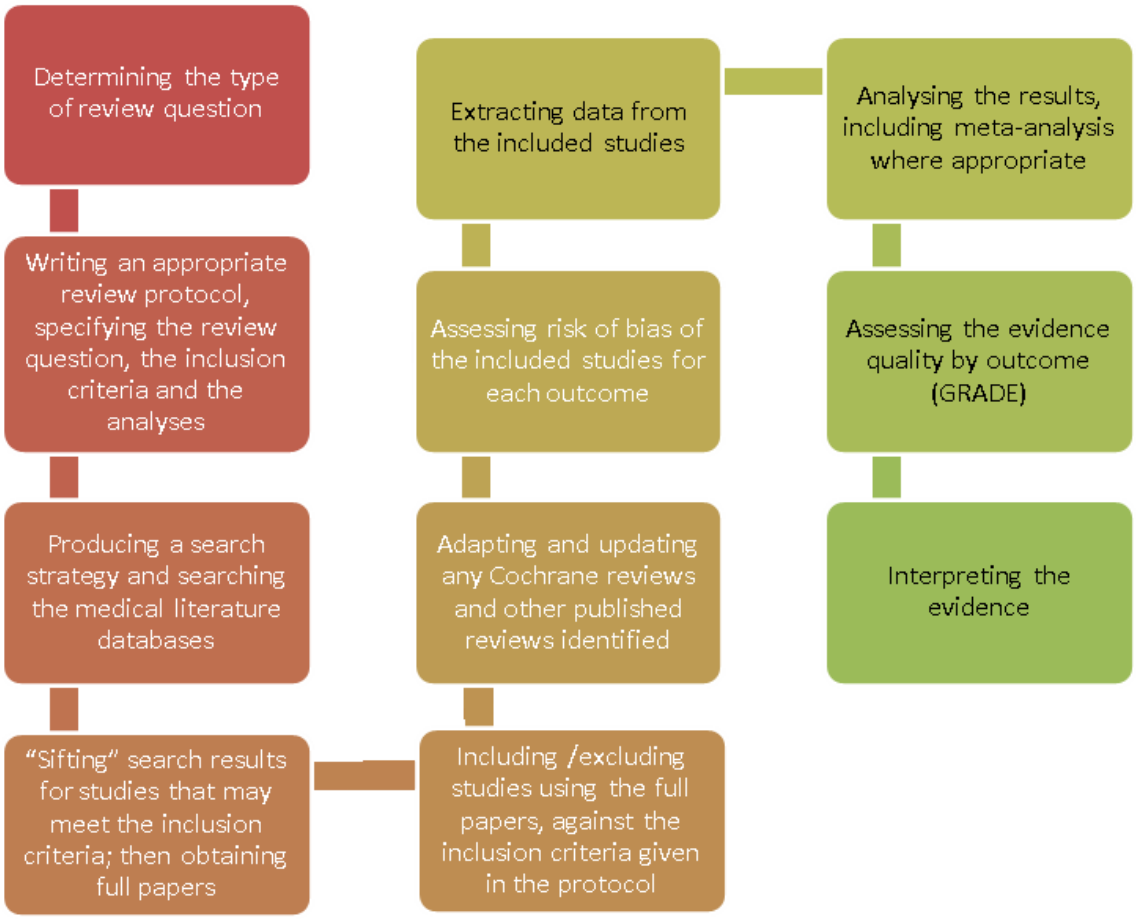
The evidence was reviewed following the steps shown schematically in Figure 1:

- Potentially relevant studies were identified for each review question from the relevant search results by reviewing titles and abstracts. Full papers were then obtained.
- Full papers were reviewed against pre-specified inclusion and exclusion criteria to identify studies that addressed the review question in the appropriate population (review protocols are included in Appendix C).
- Relevant studies were critically appraised using the appropriate checklist as specified in The guidelines manual.[114] For diagnostic questions, the QUADAS-2 checklist was followed (http://www.bris.ac.uk/quadas/quadas-2/).
- Key information was extracted on the study's methods, PICO factors and results. These were presented in summary tables (in each review chapter) and evidence tables (in Appendix G).
- Summaries of evidence were generated by outcome (included in the relevant review chapters) and were presented in GC meetings:

- Randomised studies: data were meta-analysed where appropriate and reported in GRADE profiles (for intervention reviews).
- Observational studies: data were presented as a range of values in GRADE profiles.
- Prognostic studies: data were presented as a range of values, usually in terms of the relative effect as reported by the authors.
- Diagnostic studies were presented as measures of diagnostic test accuracy (sensitivity, specificity, positive and negative predictive value). Coupled values of sensitivity and specificity were summarised in Receiver Operating Curves (ROC) to investigate heterogeneity more effectively, where evidence was available from five or more studies for any one index test. A meta-analysis of the summary operating point, (i.e. summary values for sensitivity and specificity) could not be conducted for any of the index tests, because the studies reported data at various thresholds and because data at any one threshold were not available from five or more studies.
- Qualitative studies: each study was summarised in a table where possible, otherwise presented in a narrative.

A 20% sample of each of the above stages of the reviewing process was quality assured by a second reviewer to eliminate any potential of reviewer bias or error.

Figure 1: Step-by-step process of review of evidence in the guideline



3.4.1 Inclusion and exclusion criteria

The inclusion and exclusion of studies was based on the review protocols, which can be found in Appendix C. Excluded studies by review question (with the reasons for their exclusion) are listed in Appendix K. The GC was consulted about any uncertainty regarding inclusion or exclusion.

Randomised trials, non-randomised trials, and observational studies (including diagnostic or prognostic studies) were included in the evidence reviews as appropriate.

For diagnostic reviews, the guideline population was defined as people with suspected asthma (presenting with respiratory symptoms). The GC agreed that general population studies, or studies using a questionnaire to identify people with symptoms in the general population, should be excluded unless there was no other evidence. This is because the diagnostic tests under investigation would be performed in people with suspected asthma presenting to their GP, not as screening tests in the general population.

For diagnostic reviews, the reference standard was defined as physician diagnosis of asthma based on symptoms plus an objective test. The GC agreed that studies using a reference standard of physician diagnosis only (without an objective test), or studies using an affirmative answer on a questionnaire to the question 'Has your doctor ever diagnosed you with asthma?' should be excluded. This is due to concerns about the over-diagnosis of asthma and the accuracy of a reference standard test that does not include an objective test. The GC specified the following objective tests and cut-off values and prioritised studies using a reference standard that included one of these:

- peak flow variability (cut-off value of more than 20% variability as indication of a positive test);
- bronchodilator reversibility (cut-off value of an improvement in FEV1 of more than or equal to 12%, and an increase in volume of more than or equal to 200mls as indication of a positive test);
- bronchial hyper-reactivity (histamine or methacholine challenge test, cut-off value of PC20 less than or equal to 8mg/ml as indication of a positive test)

For reviews including one of the above three objective tests as the index test under investigation, the respective tests were excluded from the reference standard for that review.

Where no evidence was available using the cut-off values specified above, evidence was included from studies using a reference standard of physician diagnosis with an objective test using an alternative threshold. Where no evidence was available from studies using physician diagnosis and an objective test, evidence was included from studies using physician diagnosis based on symptoms alone, or patient report of a previous physician diagnosis.

In children aged 1-<5 years, objective tests cannot be performed so the reference standard was defined as physician diagnosis based on recurrent and persistent wheezing.

For the monitoring reviews, the GC agreed that the most appropriate type of study design is one that involves a test and treat approach, comparing two strategies in a randomised design. People with asthma are randomised to receive the monitoring intervention plus appropriate change in treatment versus the comparator plus appropriate change in treatment, and the impact on patient outcomes is investigated.

For the monitoring reviews, the guideline population was people with asthma (defined as physician diagnosis with an objective test). The GC acknowledged that individual studies may not provide details on how the asthma population were diagnosed. Therefore, evidence was also included from studies in people with asthma (where the diagnosis criteria was unclear) or from studies including people on asthma medication.

Severe, difficult to control asthma is excluded from the scope of this guideline. Therefore, for monitoring reviews severe asthma was an exclusion criteria, defined as in the International ERS/ATS guidelines[35] and summarised below:

- Definition of severe asthma for patients aged 6 years or more: asthma which requires treatment with guidelines suggested medications for GINA steps 4–5 asthma (high dose ICS and LABA or leukotriene modifier/theophylline) for the previous year or systemic CS for 50% or more of the previous year to prevent it from becoming 'uncontrolled' or which remains 'uncontrolled' despite

this therapy. Uncontrolled asthma defined as at least one of the following: 1) poor symptom control: ACQ consistently >1.5, ACT <20 (or ''not well controlled'' by NAEPP/GINA guidelines); 2) frequent severe exacerbations: two or more bursts of systemic CS (>3 days each) in the previous year; 3) serious exacerbations: at least one hospitalisation, ICU stay or mechanical ventilation in the previous year; 4) airflow limitation: after appropriate bronchodilator withhold FEV1 <80% predicted (in the face of reduced FEV1/FVC defined as less than the lower limit of normal). Controlled asthma that worsens on tapering of these high doses of ICS or systemic CS (or additional biologics). Note: the definition of high dose ICS is age-specific.

Conference abstracts were not automatically excluded from the review but were initially assessed against the inclusion criteria and then further processed only if no other full publication was available for that review question, in which case the authors of the selected abstracts were contacted for further information. No clinical evidence was identified for two reviews (chapter 9: symptoms after using medication; and chapter 25: monitoring peripheral blood eosinophil count). However, no conference abstracts were identified which matched the review protocol. Therefore, no review included conference abstracts.

Literature reviews, posters, letters, editorials, comment articles, unpublished studies and studies not in English were excluded.

The review protocols are presented in Appendix C.

3.4.2 Methods of combining clinical studies

3.4.2.1 Data synthesis for intervention reviews

Where possible, meta-analyses were conducted to combine the results of studies for each review question using Cochrane Review Manager (RevMan5) software. Fixed-effects (Mantel-Haenszel) techniques were used to calculate risk ratios (relative risk) for the binary outcomes.

For continuous outcomes, measures of central tendency (mean) and variation (standard deviation) were required for meta-analysis. Data for continuous outcomes were analysed using an inverse variance method for pooling weighted mean differences and, where the studies had different scales, standardised mean differences were used. A generic inverse variance option in RevMan5 was used if any studies reported solely the summary statistics and 95% confidence interval (95% CI) or standard error; this included any hazard ratios reported. However, in cases where standard deviations were not reported per intervention group, the standard error (SE) for the mean difference was calculated from other reported statistics (p values or 95% CIs); meta-analysis was then undertaken for the mean difference and SE using the generic inverse variance method in RevMan5. When the only evidence was based on studies that summarised results by presenting medians (and interquartile ranges), or only p values were given, this information was assessed in terms of the study's sample size and was included in the GRADE tables without calculating the relative or absolute effects. Consequently, aspects of quality assessment such as imprecision of effect could not be assessed for evidence of this type. Where reported, time-to-event data were presented as a hazard ratio.

Stratified analyses were predefined for some review questions at the protocol stage when the GC identified that these strata are different in terms of biological and clinical characteristics and the interventions were expected to have a different effect on subpopulations. For example, objective tests of lung function used for both asthma diagnosis and for monitoring asthma control are known to perform differently in children, and three population strata were identified: children (1-<5 years old); children/ young people (5-16 years old); and adults (>16 years old).

Statistical heterogeneity was assessed by visually examining the forest plots, and by considering the chi-squared test for significance at $p<0.1$ or an I-squared inconsistency statistic (with an I-squared value of more than 50% indicating considerable heterogeneity). Where considerable heterogeneity

was present, we carried out predefined subgroup analyses – see protocols in appendix C. Assessments of potential differences in effect between subgroups were based on the chi-squared tests for heterogeneity statistics between subgroups. If no sensitivity analysis was found to completely resolve statistical heterogeneity then a random-effects (DerSimonian and Laird) model was employed to provide a more conservative estimate of the effect.

The means and standard deviations of continuous outcomes were required for meta-analysis. However, in cases where standard deviations were not reported, the standard error was calculated if the p values or 95% CIs were reported and meta-analysis was undertaken with the mean and standard error using the generic inverse variance method in RevMan5. Where p values were reported as 'less than', a conservative approach was undertaken. For example, if p value was reported as 'p≤0.001', the calculations for standard deviations will be based on a p value of 0.001. If these statistical measures were not available then the methods described in Section 16.1.3 of the Cochrane Handbook (September 2009) 'Missing standard deviations' were applied as the last resort.

For interpretation of the binary outcome results, differences in the absolute event rate were calculated using the GRADEpro software, for the median event rate across the control arms of the individual studies in the meta-analysis. Absolute risk differences were presented in the GRADE profiles and in clinical summary of findings tables, for discussion with the GC.

For binary outcomes, absolute event rates were also calculated using the GRADEpro software using event rate in the control arm of the pooled results.

3.4.2.2 Data synthesis for diagnostic test accuracy review

Data and outcomes

For the reviews of diagnostic test accuracy, a positive result on the index test was found if the patient had values of the measured quantity above a threshold value, and different thresholds could be used. Diagnostic test accuracy measures used in the analysis were: area under the Receiver Operating Characteristics (ROC) curve, and, for different thresholds, sensitivity, specificity, positive and negative predictive value. The threshold of a diagnostic test is defined as the value at which the test can best differentiate between those with and without the target condition and, in practice, it varies amongst studies. For this guideline, sensitivity and specificity were considered equally important. A high sensitivity (true positives) of a test can pick up the majority of the correct cases with asthma; conversely, a high specificity (true negatives) can correctly exclude people without asthma. The GC recognised that a test with a high sensitivity is important in order to not miss cases. However, the GC was also concerned about the over-diagnosis of asthma, and the importance of being able to correctly exclude people without asthma.

Data synthesis

Coupled forest plots of sensitivity and specificity with their 95% CIs across studies (at various thresholds) were produced for each test, using RevMan5. In order to do this, 2x2 tables (the number of true positives, false positives, true negatives and false negatives) were directly taken from the study if given, or else were derived from raw data.

To allow comparison between tests, summary ROC curves were generated for each diagnostic index test from the pairs of sensitivity and specificity calculated from the 2x2 tables, selecting 1 threshold per study. This was performed when sensitivity and specificity values were available from five or more studies for the index test, selecting only 1 threshold per study. A ROC plot shows true positive rate (sensitivity) as a function of false positive rate (1 minus specificity) and a summary ROC curve can be used to see how sensitivity and specificity trade-off with each other as thresholds vary. Data were entered into RevMan5 and summary ROC curves were fitted using the Moses Littenburg

approach. In this guideline, evidence was only available from enough studies to generate a summary ROC curve for the index test for FeNO (chapter 16). Therefore, it was not possible to plot two or more tests on the same graph in order to compare the performance of the diagnostic tests visually, and therefore a meta-analysis of the summary ROC curves was not performed. However, the GC was interested in the placement of the index tests in a diagnostic algorithm, and not just the performance of each diagnostic test in isolation. The paired sensitivity and specificity values from the diagnostic tests were used to inform the placement of index tests in the diagnostic algorithm.

Heterogeneity or inconsistency amongst studies was visually inspected in the forest plots where appropriate (only when there were similar thresholds). A diagnostic meta-analysis of the summary operating point, (i.e. summary values for sensitivity and specificity) was not conducted because data were not available from five or more studies at the same cut-off threshold for any of the index tests, in order to estimate a summary sensitivity and specificity point at a chosen threshold. Instead, at each threshold, the median sensitivity value and its corresponding specificity were presented, along with the range of values.

3.4.3 Type of studies

For monitoring reviews in this guideline, parallel randomised controlled trials (RCTs) were included because they are considered the most robust type of study design that could produce an unbiased estimate of the monitoring intervention effects. Crossover RCTs were not appropriate for the monitoring reviews due to the nature of the intervention, adjustment of treatment based on monitoring. Please refer to Appendix C for full details on the study design of studies selected for each review question.

For diagnostic reviews, cross-sectional and retrospective studies were included. Case–control studies were not included for reviews of diagnostic test accuracy.

3.4.4 Appraising the quality of evidence by outcomes

The evidence for outcomes from the included RCTs and, where appropriate, observational studies were evaluated and presented using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group (http://www.gradeworkinggroup.org/). The software developed by the GRADE working group (GRADEpro) was used to assess the quality of each outcome, taking into account individual study quality factors and the meta-analysis results. Results were presented in GRADE profiles ('GRADE tables'), which consist of 2 sections: the 'Clinical evidence profile' table includes details of the quality assessment while the 'Clinical evidence summary of findings' table includes pooled outcome data, where appropriate, an absolute measure of intervention effect and the summary of quality of evidence for that outcome. In this table, the columns for intervention and control indicate summary measures and measures of dispersion (such as mean and standard deviation or median and range) for continuous outcomes and frequency of events (n/N: the sum across studies of the number of patients with events divided by sum of the number of completers) for binary outcomes. Reporting or publication bias was only taken into consideration in the quality assessment and included in the 'Clinical evidence profile' table if it was apparent.

The evidence for each outcome was examined separately for the quality elements listed and defined in Table 2. Each element was graded using the quality levels listed in Table 3. The main criteria considered in the rating of these elements are discussed below (see Section 3.4.5 Grading of evidence). Footnotes were used to describe reasons for grading a quality element as having serious or very serious problems. The ratings for each component were summed to obtain an overall assessment for each outcome (Table 4).

The GRADE toolbox is currently designed only for randomised trials and observational studies but we adapted the quality assessment elements and outcome presentation for diagnostic accuracy studies.

Table 2: Description of the elements in GRADE used to assess the quality of intervention studies

Quality element	Description
Risk of bias ('Study limitations')	Limitations in the study design and implementation may bias the estimates of the treatment effect. High risk of bias for the majority of the evidence decreases confidence in the estimate of the effect
Inconsistency	Inconsistency refers to an unexplained heterogeneity of results
Indirectness	Indirectness refers to differences in study population, intervention, comparator and outcomes between the available evidence and the review question, or recommendation made, such that the effect estimate is changed
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect. Imprecision results if the confidence interval includes the clinically important threshold
Publication bias	Publication bias is a systematic underestimate or an overestimate of the underlying beneficial or harmful effect due to the selective publication of studies

Table 3: Levels of quality elements in GRADE

Level	Description
None	There are no serious issues with the evidence
Serious	The issues are serious enough to downgrade the outcome evidence by 1 level
Very serious	The issues are serious enough to downgrade the outcome evidence by 2 levels

Table 4: Overall quality of outcome evidence in GRADE

Level	Description
High	Further research is very unlikely to change our confidence in the estimate of effect
Moderate	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
Low	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
Very low	Any estimate of effect is very uncertain

3.4.5 Grading the quality of clinical evidence

After results were pooled, the overall quality of evidence for each outcome was considered. The following procedure was adopted when using GRADE:

1. A quality rating was assigned, based on the study design. RCTs start as High, observational studies as Low, and uncontrolled case series as Low or Very low.
2. The rating was then downgraded for the specified criteria: risk of bias (study limitations), inconsistency, indirectness, imprecision and publication bias. These criteria are detailed below. Evidence from observational studies (which had not previously been downgraded) was upgraded if there was: a large magnitude of effect, a dose–response gradient, and if all plausible confounding would reduce a demonstrated effect or suggest a spurious effect when results showed no effect. Each quality element considered to have 'serious' or 'very serious' risk of bias was rated down by 1 or 2 points respectively.
3. The downgraded or upgraded marks were then summed and the overall quality rating was revised. For example, all RCTs started as High and the overall quality became Moderate, Low or Very low if 1, 2 or 3 points were deducted respectively.

4. The reasons or criteria used for downgrading were specified in the footnotes.

The details of the criteria used for each of the main quality elements are discussed further in the following sections 3.4.6 to 3.4.9.

3.4.6 Risk of bias

Bias can be defined as anything that causes a consistent deviation from the truth. Bias can be perceived as a systematic error, for example, if a study was carried out several times and there was a consistently wrong answer, the results would be inaccurate.

The risk of bias for a given study and outcome is associated with the risk of over- or underestimation of the true effect.

Potential causes of bias are listed in Table 5.

A study with a poor methodological design does not automatically imply high risk of bias; the bias is considered individually for each outcome and it is assessed whether this poor design will impact on the estimation of the intervention effect.

The GC accepted that patient and investigator blinding in monitoring intervention studies was impossible to achieve in most situations. Nevertheless, open-label studies for monitoring were downgraded for subjective or patient reported outcomes to maintain a consistent approach in quality rating across the guideline, as these outcomes are highly subjected to bias in an open label setting.

Table 5: Risk of bias in randomised controlled trials

Risk of bias	Explanation
Allocation concealment	Those enrolling patients are aware of the group to which the next enrolled patient will be allocated (this is a major problem in 'pseudo' or 'quasi' randomised trials with, for example, allocation by day of week, birth date, chart number)
Lack of blinding	Patient, caregivers, those recording outcomes, those adjudicating outcomes, or data analysts are aware of the arm to which patients are allocated
Incomplete accounting of patients and outcome events	Missing data not accounted for and failure of the trialists to adhere to the intention-to-treat principle when indicated
Selective outcome reporting	Reporting of some outcomes and not others on the basis of the results
Other risks of bias	For example: • Stopping early for benefit observed in randomised trials, in particular in the absence of adequate stopping rules • Use of unvalidated patient-reported outcomes • Recruitment bias in cluster-randomised trials

3.4.6.1 Diagnostic studies

For diagnostic accuracy studies, the Quality Assessment of Diagnostic Accuracy Studies version 2 (QUADAS-2) checklist was used (see Appendix F in The guidelines manual[114]). Risk of bias and applicability in primary diagnostic accuracy studies in QUADAS-2 consists of 4 domains (see Figure 2):

- Patient selection
- Index test
- Reference standard

- Flow and timing

Figure 2: Summary of QUADAS-2 checklist

DOMAIN	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING
Description	Describe methods of patient selection: Describe included patients (prior testing, presentation, intended use of index test and setting):	Describe the index test and how it was conducted and interpreted:	Describe the reference standard and how it was conducted and interpreted:	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard:
Signalling questions (yes/no/unclear)	Was a consecutive or random sample of patients enrolled?	Were the index test results interpreted without knowledge of the results of the reference standard?	Is the reference standard likely to correctly classify the target condition?	Was there an appropriate interval between index test(s) and reference standard?
	Was a case-control design avoided?	If a threshold was used, was it pre-specified?	Were the reference standard results interpreted without knowledge of the results of the index test?	Did all patients receive a reference standard?
	Did the study avoid inappropriate exclusions?			Did all patients receive the same reference standard?
				Were all patients included in the analysis?
Risk of bias: High/low/unclear	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?
Concerns regarding applicability: High/low/unclear	Are there concerns that the included patients do not match the review question?	Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Are there concerns that the target condition as defined by the reference standard does not match the review question?	

Source: QUADAS-2 website, University of Bristol [182]

Optional domain, multiple test accuracy is applicable when a single study examined more than 1 diagnostic test (head-to-head comparison between 2 or more index tests reported within the same study). This optional domain contains 3 questions relating to risk of bias:

- Did all patients undergo all index tests or were the index tests appropriately randomised amongst the patients?
- Were index tests conducted within a short time interval?
- Are index test results unaffected when undertaken together on the same patient?

3.4.6.2 Additional considerations

The GC raised a number of issues that needed to be taken into consideration when assessing study quality and they are listed as follows:

Patient selection (concerns regarding applicability): the population was defined as people with suspected asthma (presenting with respiratory symptoms). The GC agreed that general population studies, or studies using a questionnaire to identify people with symptoms in the general population, should only be included if there was no other evidence, and downgraded for applicability. This is because the diagnostic tests under investigation would be performed in people with suspected asthma presenting to their GP, not as screening tests in the general population.

Index test: the GC thought that the interpretation of the index tests was unlikely to be influenced by the knowledge of the results of the reference standard, as they are not subjective tests.

Reference standard (concerns regarding applicability): the GC agreed that the reference standard should be physician diagnosis of asthma based on symptoms plus an objective test, as described in

section 3.3.1. Studies with a different reference standard were only included if there was no other evidence, and downgraded for applicability.

3.4.7 Inconsistency

Inconsistency refers to an unexplained heterogeneity of results. When estimates of the treatment effect across studies differ widely (that is, there is heterogeneity or variability in results), this suggests true differences in underlying treatment effect.

Heterogeneity in meta-analyses was examined and sensitivity and subgroup analyses performed as pre-specified in the protocols (Appendix C).

When heterogeneity exists (chi-squared $p<0.1$, I-squared inconsistency statistic of >50%, or evidence from examining forest plots), but no plausible explanation can be found (for example, duration of intervention or different follow-up periods), the quality of evidence was downgraded by 1 or 2 levels, depending on the extent of uncertainty to the results contributed by the inconsistency in the results. In addition to the I-squared and chi-squared values, the decision for downgrading was also dependent on factors such as whether the intervention is associated with benefit in all other outcomes or whether the uncertainty about the magnitude of benefit (or harm) of the outcome showing heterogeneity would influence the overall judgment about net benefit or harm (across all outcomes).

3.4.8 Indirectness

Directness refers to the extent to which the populations, intervention, comparisons and outcome measures are similar to those defined in the inclusion criteria for the reviews. Indirectness is important when these differences are expected to contribute to a difference in effect size, or may affect the balance of harms and benefits considered for an intervention.

For diagnostic review, indirectness was assessed using the applicability domains of the QUADAS II checklist (see Figure 2).

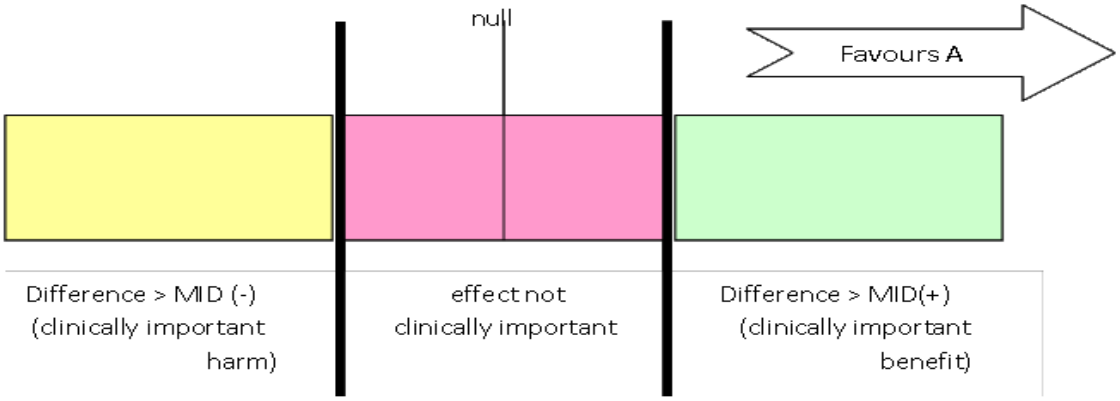
3.4.9 Imprecision

Imprecision in guidelines concerns whether the uncertainty (confidence interval) around the effect estimate means that it is not clear whether there is a clinically important difference between interventions or not. Therefore, imprecision differs from the other aspects of evidence quality, in that it is not really concerned with whether the point estimate is accurate or correct (has internal or external validity), rather it is concerned with uncertainty about what the point estimate is. This uncertainty is reflected in the width of the confidence interval.

The 95% confidence interval (95% CI) is defined as the range of values that contain the population value with 95% probability. The larger the trial, the smaller the 95% CI and the more certain the effect estimate.

Imprecision in the evidence reviews was assessed by considering whether the width of the 95% CI of the effect estimate is relevant to decision-making, considering each outcome in isolation. Figure 3 considers a positive outcome for the comparison of treatment A versus B. Three decision-making zones can be identified, bounded by the thresholds for clinical importance (minimal important difference – MID) for benefit and for harm. The MID for harm for a positive outcome means the threshold at which drug A is less effective than drug B by an amount that is clinically important to patients (favours B).

Figure 3: Illustration of precise and imprecise outcomes based on the confidence interval of outcomes in a forest plot



When the confidence interval of the effect estimate is wholly contained in one of the 3 zones (for example, clinically important benefit), we are not uncertain about the size and direction of effect (whether there is a clinically important benefit, or the effect is not clinically important, or there is a clinically important harm), so there is no imprecision.

When a wide confidence interval lies partly in each of 2 zones, it is uncertain in which zone the true value of effect estimate lies, and therefore there is uncertainty over which decision to make (based on this outcome alone). The confidence interval is consistent with 2 decisions and so this is considered to be imprecise in the GRADE analysis and the evidence is downgraded by 1 level ('serious imprecision').

If the confidence interval of the effect estimate crosses into 3 zones, this is considered to be very imprecise evidence because the confidence interval is consistent with 3 clinical decisions and there is a considerable lack of confidence in the results. The evidence is therefore downgraded by 2 levels in the GRADE analysis ('very serious imprecision').

Implicitly, assessing whether the confidence interval is in, or partially in, a clinically important zone, requires the GC to estimate an MID or to say whether they would make different decisions for the 2 confidence limits.

The GC was asked whether they were aware of any acceptable MIDs in the clinical community. The GC provided established MIDs which were used for the following outcomes:

- AQLQ, child AQLQ, carer AQLQ and mini AQLQ: 0.5[78]
- ACT: 3.0[150]
- FEV1 litres: 0.23L[148]
- PEF L/min: 20L/min[148]

Finally, the GC considered it clinically acceptable to use the GRADE default MID to assess imprecision: a 25% relative risk reduction or relative risk increase was used, which corresponds to clinically important thresholds for a risk ratio of 0.75 and 1.25 respectively. This default MID was used for all other outcomes in the monitoring evidence reviews.

3.4.10 Assessing clinical importance

The GC assessed the evidence by outcome in order to determine if there was, or potentially was, a clinically important benefit, a clinically important harm or no clinically important difference between interventions. To facilitate this, binary outcomes were converted into absolute risk differences (ARDs) using GRADEpro software: the median control group risk across studies was used to calculate the ARD and its 95% CI from the pooled risk ratio.

The assessment of benefit, harm, or no benefit or harm was based on the point estimate of absolute effect for intervention studies which was standardised across the reviews. The GC considered for most of the outcomes in the intervention reviews that if at least 100 participants per 1000 (10%) achieved (if positive) the outcome of interest in the intervention group compared to the comparison group then this intervention would be considered beneficial. The same point estimate but in the opposite direction would apply if the outcome was negative.

This assessment was carried out by the GC for each critical outcome, and an evidence summary table was produced to compile the GC's assessments of clinical importance per outcome, alongside the evidence quality and the uncertainty in the effect estimate (imprecision).

3.4.11 Evidence statements

Evidence statements are summary statements that are presented after the GRADE profiles, summarising the key features of the clinical effectiveness evidence presented. The wording of the evidence statements reflects the certainty or uncertainty in the estimate of effect. The evidence statements are presented by outcome and encompass the following key features of the evidence:

- the number of studies and the number of participants for a particular outcome
- a brief description of the participants
- an indication of the direction of effect (if one treatment is beneficial or harmful compared to the other, or whether there is no difference between the 2 tested treatments)
- a description of the overall quality of evidence (GRADE overall quality).

3.5 Evidence of cost-effectiveness

The GC is required to make decisions based on the best available evidence of both clinical and cost-effectiveness. Guideline recommendations should be based on the expected costs of the different options in relation to their expected health benefits (that is, their 'cost-effectiveness') rather than the total implementation cost.[114] Thus, if the evidence suggests that a strategy provides significant health benefits at an acceptable cost per patient treated, it should be recommended even if it would be expensive to implement across the whole population.

Evidence on cost-effectiveness related to the key clinical issues being addressed in the guideline was sought. The health economist:

- Undertook a systematic review of the published economic literature.
- Undertook new cost-effectiveness analysis in priority areas.

3.5.1 Literature review

The health economist:

- Identified potentially relevant studies for each review question from the economic search results by reviewing titles and abstracts. Full papers were then obtained.
- Reviewed full papers against pre-specified inclusion and exclusion criteria to identify relevant studies (see below for details).
- Critically appraised relevant studies using the economic evaluations checklist as specified in The guidelines manual.[114]
- Extracted key information about the studies' methods and results into evidence tables (included in Appendix H).
- Generated summaries of the evidence in NICE economic evidence profiles (included in the relevant chapter for each review question) – see below for details.

3.5.1.1 Inclusion and exclusion criteria

Full economic evaluations (studies comparing costs and health consequences of alternative courses of action: cost–utility, cost-effectiveness, cost–benefit and cost–consequence analyses) and comparative costing studies that addressed the review question in the relevant population were considered potentially includable as economic evidence.

Studies that only reported cost per hospital (not per patient), or only reported average cost-effectiveness without disaggregated costs and effects, were excluded. Literature reviews, abstracts, posters, letters, editorials, comment articles, unpublished studies and studies not in English were excluded.

Remaining studies were prioritised for inclusion based on their relative applicability to the development of this guideline and the study limitations. For example, if a high quality, directly applicable UK analysis was available, then other less relevant studies may not have been included. Where exclusions occurred on this basis, this is noted in the relevant section.

For more details about the assessment of applicability and methodological quality see the economic evaluation checklist (Appendix F of The guidelines manual.[114] and the health economics review protocol in Appendix C).

When no relevant economic studies were found from the economic literature review, relevant UK NHS unit costs related to the compared interventions were presented to the GC to inform the possible economic implications of the recommendations.

3.5.1.2 NICE economic evidence profiles

The NICE economic evidence profile has been used to summarise cost and cost-effectiveness estimates. The economic evidence profile shows an assessment of applicability and methodological quality for each economic evaluation, with footnotes indicating the reasons for the assessment. These assessments were made by the health economist using the economic evaluation checklist from The guidelines manual.[114] It also shows the incremental costs, incremental effects (for example, quality-adjusted life years [QALYs]) and incremental cost-effectiveness ratio for the base case analysis in the evaluation, as well as information about the assessment of uncertainty in the analysis. See Table 6 for more details.

If a non-UK study was included in the profile, the results were converted into pounds sterling using the appropriate purchasing power parity.[120]

Table 6: Content of NICE economic evidence profile

Item	Description
Study	First author name, reference, date of study publication and country perspective.
Applicability	An assessment of applicability of the study to the clinical guideline, the current NHS situation and NICE decision-making[(a)]: • Directly applicable – the study meets all applicability criteria, or fails to meet one or more applicability criteria but this is unlikely to change the conclusions about cost-effectiveness. • Partially applicable – the study fails to meet one or more applicability criteria, and this could change the conclusions about cost-effectiveness. • Not applicable – the study fails to meet one or more of the applicability criteria, and this is likely to change the conclusions about cost-effectiveness. Such studies would usually be excluded from the review.
Limitations	An assessment of methodological quality of the study[(a)]: • Minor limitations – the study meets all quality criteria, or fails to meet one or

Item	Description
	more quality criteria, but this is unlikely to change the conclusions about cost-effectiveness. • Potentially serious limitations – the study fails to meet one or more quality criteria, and this could change the conclusions about cost-effectiveness. • Very serious limitations – the study fails to meet one or more quality criteria, and this is highly likely to change the conclusions about cost-effectiveness. Such studies would usually be excluded from the review.
Other comments	Particular issues that should be considered when interpreting the study.
Incremental cost	The mean cost associated with one strategy minus the mean cost of a comparator strategy.
Incremental effects	The mean QALYs (or other selected measure of health outcome) associated with one strategy minus the mean QALYs of a comparator strategy.
Cost-effectiveness	Incremental cost-effectiveness ratio (ICER): the incremental cost divided by the incremental effects.
Uncertainty	A summary of the extent of uncertainty about the ICER reflecting the results of deterministic or probabilistic sensitivity analyses, or stochastic analyses of trial data, as appropriate.

(a) Applicability and limitations were assessed using the economic evaluation checklist in Appendix G of The guidelines manual (2012)[114]

3.5.2 Undertaking new health economic analysis

As well as reviewing the published economic literature for each review question, as described above, new economic analysis was undertaken by the health economist in selected areas. Priority areas for new health economic analysis were agreed by the GC after formation of the review questions and consideration of the available health economic evidence.

The GC identified the diagnosis of asthma in adults as the highest priority area for original economic modelling. Further details are available in Appendix M.

The following general principles were adhered to in developing the cost-effectiveness analysis:

- Methods were consistent with the NICE reference case.[115]
- The GC was involved in the design of the model, selection of inputs and interpretation of the results.
- Model inputs were based on the systematic review of the clinical literature supplemented with other published data sources where possible.
- When published data were not available GC expert opinion was used to populate the model.
- Model inputs and assumptions were reported fully and transparently.
- The results were subject to sensitivity analysis and limitations were discussed.
- The model was peer-reviewed by another health economist at the NGC.

Full methods for the cost-effectiveness analysis for diagnosis of asthma in adults are described in Appendix M.

3.5.3 Cost-effectiveness criteria

NICE's report 'Social value judgements: principles for the development of NICE guidance' sets out the principles that GCs should consider when judging whether an intervention offers good value for money.[113] In general, an intervention was considered to be cost-effective if either of the following criteria applied (given that the estimate was considered plausible):

- the intervention dominated other relevant strategies (that is, it was both less costly in terms of resource use and more clinically effective compared with all the other relevant alternative strategies), or
- the intervention cost less than £20,000 per QALY gained compared with the next best strategy.

If the GC recommended an intervention that was estimated to cost more than £20,000 per QALY gained, or did not recommend one that was estimated to cost less than £20,000 per QALY gained, the reasons for this decision are discussed explicitly in the 'Recommendations and link to evidence' section of the relevant chapter, with reference to issues regarding the plausibility of the estimate or to the factors set out in 'Social value judgements: principles for the development of NICE guidance'.[113]

3.5.4 In the absence of economic evidence

When no relevant published studies were found, and a new analysis was not prioritised, the GC made a qualitative judgement about cost-effectiveness by considering expected differences in resource use between options and relevant UK NHS unit costs, alongside the results of the clinical review of effectiveness evidence.

3.6 Developing recommendations

Over the course of the guideline development process, the GC was presented with:

- Evidence tables of the clinical and economic evidence reviewed from the literature. All evidence tables are in Appendices H and I.
- Summary of clinical and economic evidence and quality (as presented in chapters 6 to 30).
- Forest plots and summary ROC curves (Appendix J).
- A description of the methods and results of the cost-effectiveness analysis undertaken for the guideline (Appendix M).

Recommendations were drafted on the basis of the GC interpretation of the available evidence, taking into account the balance of benefits, harms and costs between different courses of action. This was either done formally in an economic model, or informally. Firstly, the net benefit over harm (clinical effectiveness) was considered, focusing on the critical outcomes. When this was done informally, the GC took into account the clinical benefits and harms when one intervention was compared with another. The assessment of net benefit was moderated by the importance placed on the outcomes (the GC's values and preferences), and the confidence the GC had in the evidence (evidence quality). Secondly, it was assessed whether the net benefit justified any differences in costs.

When clinical and economic evidence was of poor quality, conflicting or absent, the GC drafted recommendations based on their expert opinion. The considerations for making consensus-based recommendations include the balance between potential harms and benefits, the economic costs compared to the economic benefits, current practices, recommendations made in other relevant guidelines, patient preferences and equality issues. The consensus recommendations were agreed through discussions in the GC. The GC also considered whether the uncertainty was sufficient to justify delaying making a recommendation to await further research, taking into account the potential harm of failing to make a clear recommendation (see Section 3.6.1 below).

The wording of recommendations was agreed by the GC and focused on the following factors:

- The actions health professionals need to take.
- The information readers need to know.

- The strength of the recommendation (for example the word 'offer' was used for strong recommendations and 'consider' for weak recommendations).
- The involvement of patients (and their carers if needed) in decisions on treatment and care.
- Consistency with NICE's standard advice on recommendations about drugs, waiting times and ineffective interventions.

The main considerations specific to each recommendation are outlined in the 'Recommendations and link to evidence' sections within each chapter.

3.6.1 Research recommendations

When areas were identified for which good evidence was lacking, the GC considered making recommendations for future research. Decisions about inclusion were based on factors such as:

- the importance to patients or the population
- national priorities
- potential impact on the NHS and future NICE guidance
- ethical and technical feasibility.

3.6.2 Validation process

A draft of this guideline was subject to a 6-week public consultation during January-February 2015 as part of the quality assurance and peer review of the document. During consultation, some stakeholders suggested that a large investment in training and equipment would be needed to bring current practice in line with the guideline's diagnostic test recommendations, and that this was likely to be a major barrier to implementation. The concerns centred around the need for objective tests to confirm the diagnosis, whereas traditional management had relied in many cases on clinical history supplemented by examination findings and a trial of asthma treatment. This applied to some extent to all the objective tests covered in the draft guideline, but pre-eminently to the use of FeNO testing since this would be completely new to virtually every primary care group in England & Wales.

Guideline development was therefore paused in August 2015 to allow additional time to work with primary care professionals to assess the feasibility of adopting the diagnostic recommendations. An asthma feasibility project team was formed within the NICE Adoption and Impact team to work with 7 primary care sites across England, each of which agreed to implement the revised diagnostic recommendations and algorithms. The 7 sites were chosen to represent a cross-section (albeit small) of practices across the country with variation in size, geographical site and socio-economic profile of their patient lists. Outcome data was collected during a 6-month period May to October 2016. Further detail of the methods of this study, and its findings, are given in Appendix Q of this guideline.

The conclusions were important in determining the final recommendations in this guidance and are referred to in the relevant LETR sections in addition to the consideration of the standard evidence sources.

The guideline, including some of the diagnostic recommendations and the associated algorithms, was amended in the light of the results of the feasibility study. This amended guidance was subject to a second period of consultation in July 2017.

3.6.3 Updating the guideline

A formal review of the need to update a guideline is usually undertaken by NICE after its publication. NICE will conduct a review to determine whether the evidence base has progressed significantly to alter the guideline recommendations and warrant an update.

3.6.4 Disclaimer

Healthcare providers need to use clinical judgement, knowledge and expertise when deciding whether it is appropriate to apply guidelines. The recommendations cited here are a guide and may not be appropriate for use in all situations. The decision to adopt any of the recommendations cited here must be made by practitioners in light of individual patient circumstances, the wishes of the patient, clinical expertise and resources.

The National Guideline Centre disclaims any responsibility for damages arising out of the use or non-use of this guideline and the literature used in support of this guideline.

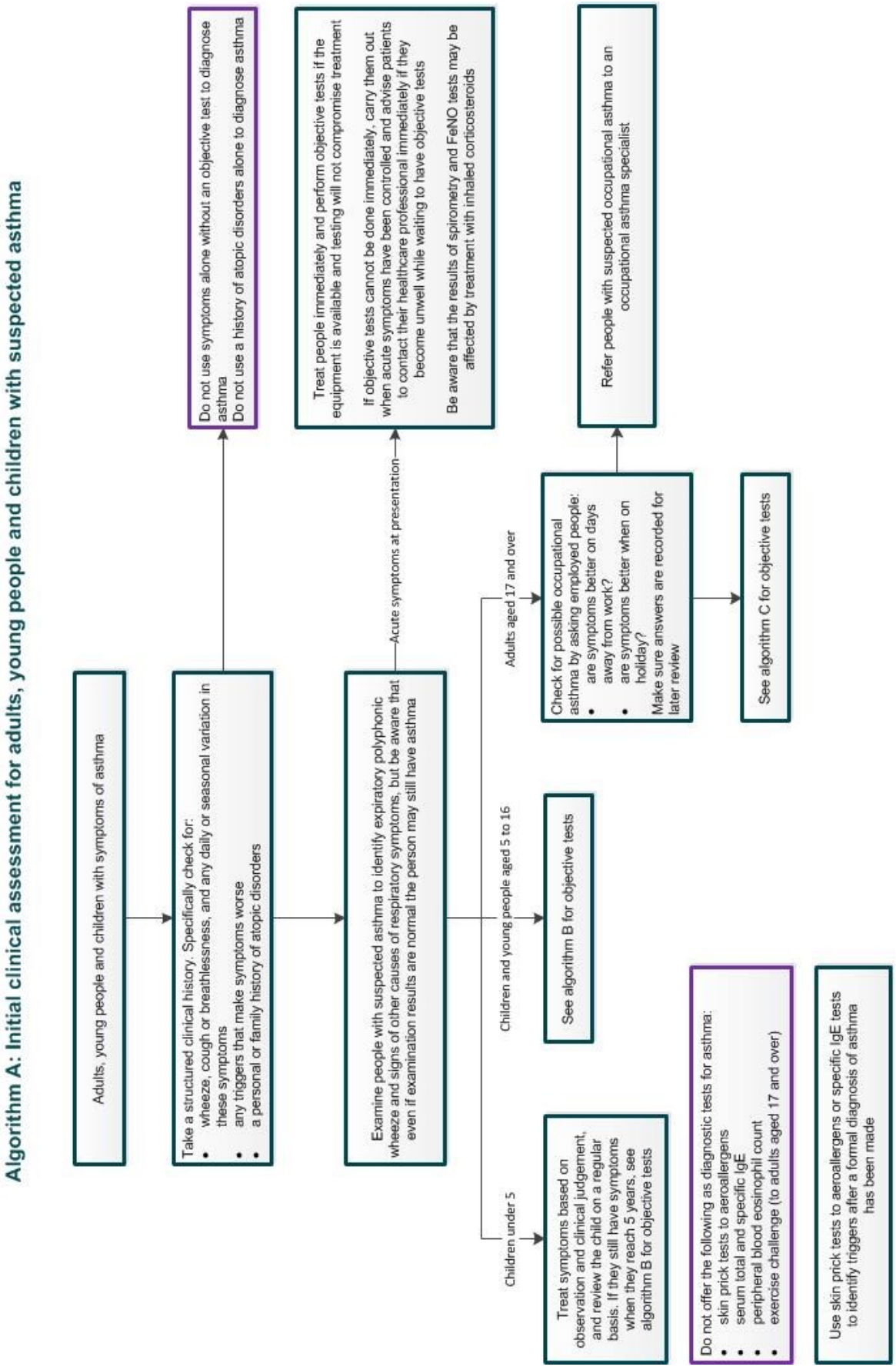
3.6.5 Funding

The National Guideline Centre was commissioned by the National Institute for Health and Care Excellence to undertake the work on this guideline.

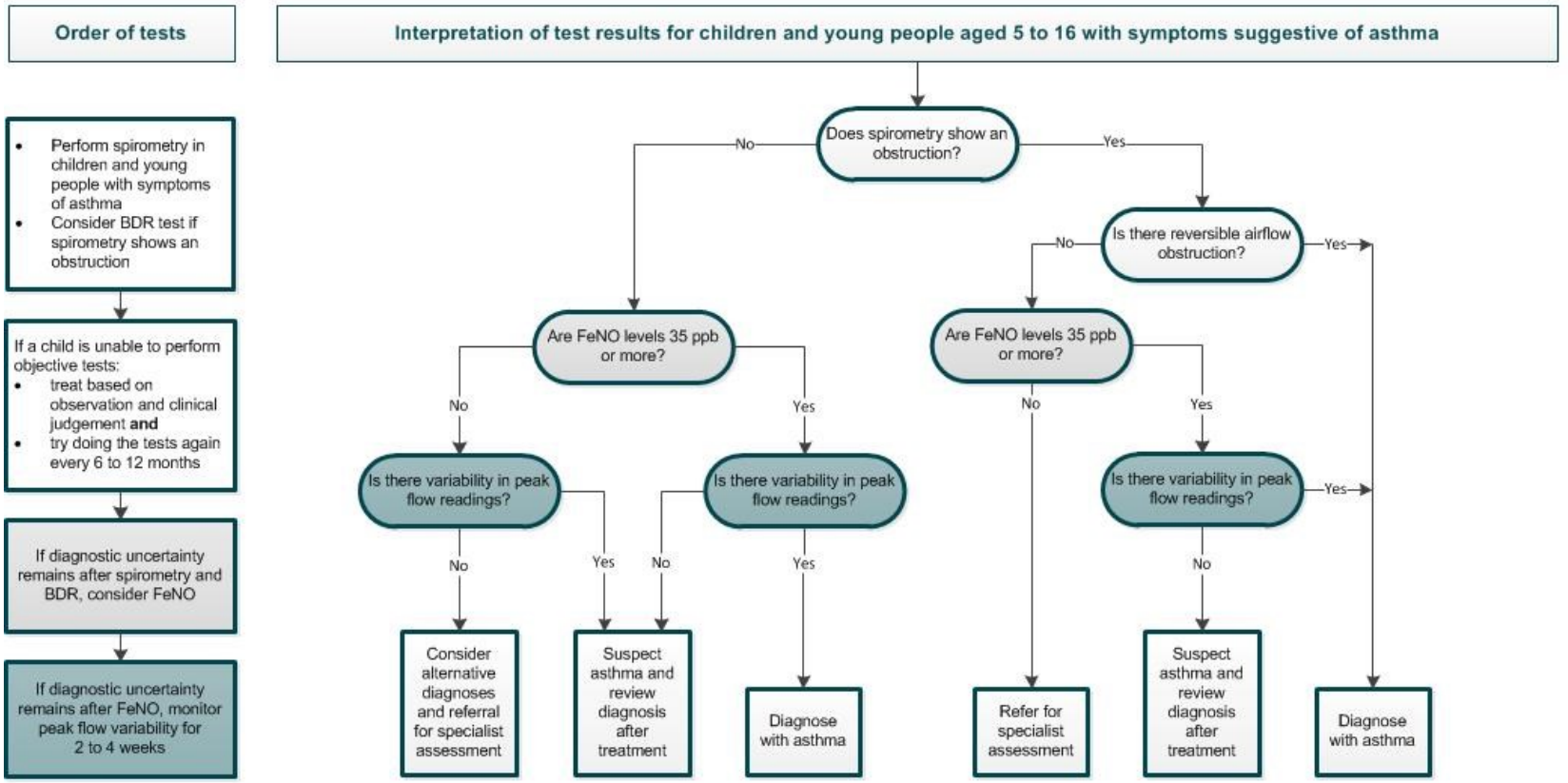
4 Guideline summary

4.1 Diagnostic algorithms

Algorithm A: Initial clinical assessment for adults, young people and children with suspected asthma



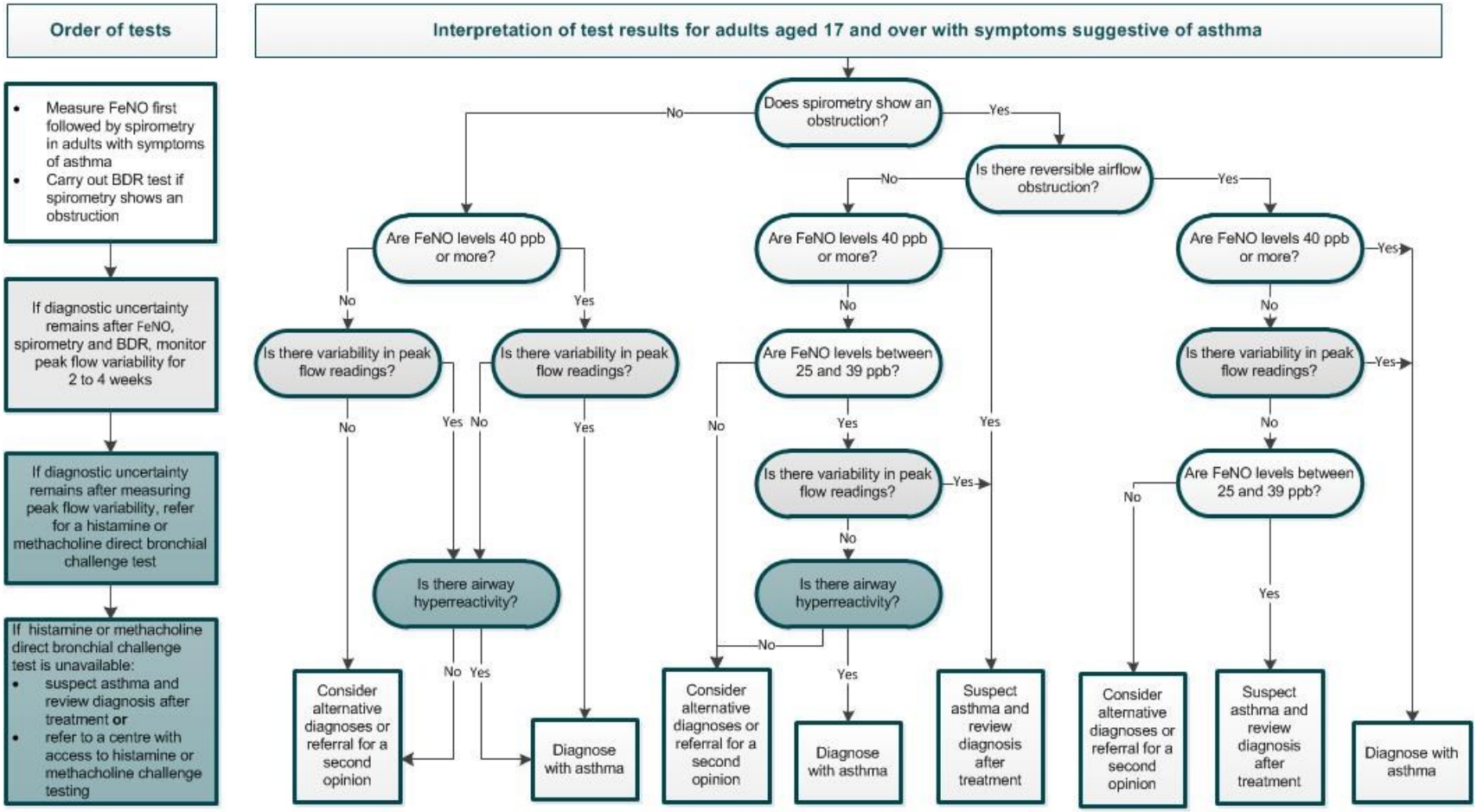
Algorithm B: Objective tests for asthma in children and young people aged 5 to 16



Positive test thresholds
Obstructive spirometry: FEV1/FVC ratio less than 70% (or below the lower limit of normal if available)
FeNO: 35 ppb or more
BDR: improvement in FEV1 of 12% or more
Peak flow variability: variability over 20%

Abbreviations:
FeNO, fractional exhaled nitric oxide
BDR, bronchodilator reversibility

Algorithm C: Objective tests for asthma in adults aged 17 and over



Positive test thresholds
Obstructive spirometry: FEV1/FVC ratio less than 70% (or below the lower limit of normal if available)
FeNO: 40 ppb or more
BDR: improvement in FEV1 of 12% or more and increase in volume of 200 ml or more
Peak flow variability: variability over 20%
Direct bronchial challenge test with histamine of methacholine: PC20 of 8 mg/ml or less

Abbreviations:
FeNO, fractional exhaled nitric oxide
BDR, bronchodilator reversibility

4.2 Full list of recommendations

Initial clinical assessment

See algorithm A for initial clinical assessment in adults, young people and children with suspected asthma

Clinical history

1. Take a structured clinical history in people with suspected asthma. Specifically, check for:

- wheeze, cough or breathlessness, and any daily or seasonal variation in these symptoms
- any triggers that make symptoms worse
- a personal or family history of atopic disorders.

2. Do not use symptoms alone without an objective test to diagnose asthma.
3. Do not use a history of atopic disorders alone to diagnose asthma.

Physical examination

4. Examine people with suspected asthma to identify expiratory polyphonic wheeze and signs of other causes of respiratory symptoms, but be aware that even if examination results are normal the person may still have asthma.

Initial treatment and objective tests for acute symptoms at presentation

5. Treat people immediately if they are acutely unwell at presentation, and perform objective tests for asthma (for example, fractional exhaled nitric oxide [FeNO], spirometry and peak flow variability) if the equipment is available and testing will not compromise treatment of the acute episode.
6. If objective tests for asthma cannot be done immediately for people who are acutely unwell at presentation, carry them out when acute synpotms have been controlled and advise people to contact their healthcare professional immediately if they become unwell while waiting to have objective tests.
7. Be aware that the results of spirometry and FeNO tests may be affected in people who have been treated empirically with inhaled corticosteroids.

Testing for asthma

8. Do not offer the following as diagnostic tests for asthma:

- skin prick tests to aeroallergens
- serum total and specific IgE
- peripheral blood eosinophil count
- exercise challenge (to adults aged 17 and over).

9. Use skin prick tests to aeroallergens or specific IgE tests to identify triggers after a formal diagnosis of asthma has been made.

Occupational asthma

10. Check for possible occupational asthma by asking employed people with suspected new-onset asthma, or established asthma that is poorly controlled:

- Are symptoms better on days away from work?
- Are symptoms better when on holiday[a]?

Make sure all answers are recorded for later review.

11. Refer people with suspected occupational asthma to an occupational asthma specialist.

Diagnosing asthma in young children

12. For children under 5 with suspected asthma, treat symptoms based on observation and clinical judgement, and review the child on a regular basis (see section 1.8 of the NICE guideline on asthma: diagnosis, monitoring and chronic asthma management). If they still have symptoms when they reach 5 years, carry out objective tests (see recommendations 14 to 35 and algorithm B).
13. If a child is unable to perform objective tests when they are aged 5:

- continue to treat based on observation and clinical judgement
- try doing the tests again every 6 to 12 months until satisfactory results are obtained
- consider referral for specialist assessment if the child repeatedly cannot perform objective tests and is not responding to treatment.

Objective tests for diagnosing asthma in adults, young people and children aged 5 and over

See also table 7 for a summary of objective test threshold levels.

Diagnostic hubs

14. Those responsible for planning diagnostic service support to primary care (for example, clinical commissioning groups) should consider establishing asthma diagnostic hubs to achieve economies of scale and improve the practicality of implementing the recommendations in this guideline.

Airway inflammation measures

Fractional exhaled nitric oxide

15. Offer a FeNO test to adults (aged 17 and over) if a diagnosis of asthma is being considered. Regard a FeNO level of 40 parts per billion (ppb) or more as a positive test.
16. Consider a FeNO test in children and young people (aged 5 to 16)[b] if there is diagnostic uncertainty after initial assessment and they have either:

- normal spirometry **or**
- obstructive spirometry with a negative bronchodilator reversibility (BDR) test.

Regard a FeNO level of 35 ppb or more as a positive test.

17. Be aware that a person's current smoking status can lower FeNO levels both acutely and cumulatively. However, a high level remains useful in supporting a diagnosis of asthma.

[a] 'Holiday' here means any longer time away from work than usual breaks at weekends or between shifts.
[b] Children at the lower end of the age range may not be able to do the FeNO test adequately. In these cases, apply the principles in recommendation 13.

Lung function tests

Spirometry

18. Offer spirometry to adults, young people and children aged 5 and over if a diagnosis of asthma is being considered. Regard a forced expiratory volume in 1 second/forced vital capacity (FEV1/FVC) ratio of less than 70% (or below the lower limit of normal if this value is available) as a positive test for obstructive airway disease (obstructive spirometry).

Bronchodilator reversibility

19. Offer a BDR test to adults (aged 17 and over) with obstructive spirometry (FEV1/FVC ratio less than 70%). Regard an improvement in FEV1 of 12% or more, together with an increase in volume of 200 ml or more, as a positive test.
20. Consider a BDR test in children and young people (aged 5 to 16) with obstructive spirometry (FEV1/FVC ratio less than 70%). Regard an improvement in FEV1 of 12% or more as a positive test.

Peak expiratory flow variability

21. Monitor peak flow variability for 2 to 4 weeks in adults (aged 17 and over) if there is diagnostic uncertainty after initial assessment and a FeNO test and they have either:

- normal spirometry **or**
- obstructive spirometry, reversible airways obstruction (positive BDR) but a FeNO level of 39 ppb or less.

Regard a value of more than 20% variability as a positive test.

22. Consider monitoring peak flow variability for 2 to 4 weeks in adults (aged 17 and over) if there is diagnostic uncertainty after initial assessment and they have:

- obstructive spirometry **and**
- irreversible airways obstruction (negative BDR) **and**
- a FeNO level between 25 and 39 ppb.

Regard a value of more than 20% variability as a positive test.

23. Monitor peak flow variability for 2 to 4 weeks in children and young people (aged 5 to 16) if there is diagnostic uncertainty after initial assessment and a FeNO test and they have either:

- normal spirometry **or**
- obstructive spirometry, irreversible airways obstruction (negative BDR) and a FeNO level of 35 ppb or more.

Regard a value of more than 20% variability as a positive test.

Airway hyper-reactivity measures

Direct bronchial challenge test with histamine or methacholine

24. Offer a direct bronchial challenge test with histamine or methacholine[c] to adults (aged 17 and over) if there is diagnostic uncertainty after a normal spirometry and either a:

[c] At the time of publication (November 2017), histamine and methacholine did not have UK marketing authorisation for this use. The healthcare professional should follow relevant professional guidance, taking full responsibility for the

- FeNO level of 40 ppb or more and no variability in peak flow readings, **or**
- FeNO level of 39 ppb or less with variability in peak flow readings.

Regard a PC20 value of 8 mg/ml or less as a positive test.

25. Consider a direct bronchial challenge test with histamine or methacholine[d] in adults (aged 17 and over) with:

- obstructive spirometry without bronchodilator reversibility **and**
- a FeNO level between 25 and 39 ppb **and**
- no variability in peak flow readings (less than 20% variability over 2 to 4 weeks).

Regard a PC20 value of 8 mg/ml or less as a positive test.

26. If a direct bronchial challenge test with histamine or methacholine is unavailable, suspect asthma and review the diagnosis after treatment, or refer to a centre with access to a histamine or methacholine challenge test.

Diagnosis in children and young people aged 5 to 16

See algorithm B for objective tests in young people and children aged 5 to 16.

27. Diagnose asthma in children and young people (aged 5 to 16) if they have symptoms suggestive of asthma and:

- a FeNO level of 35 ppb or more and positive peak flow variability **or**
- obstructive spirometry and positive bronchodilator reversibility.

28. Suspect asthma in children and young people (aged 5 to 16) if they have symptoms suggestive of asthma and:

- a FeNO level of 35 ppb or more with normal spirometry and negative peak flow variability, **or**
- a FeNO level of 35 ppb or more with obstructive spirometry but negative bronchodilator reversibility and no variability in peak flow readings, **or**
- normal spirometry, a FeNO level of 34 ppb or less and positive peak flow variability.

Do not rule out other diagnoses if symptom control continues to remain poor after treatment. Review the diagnosis after 6 weeks by repeating any abnormal tests and reviewing symptoms.

29. Refer children and young people (aged 5 to 16) for specialist assessment if they have obstructive spirometry, negative bronchodilator reversibility and a FeNO level of 34 ppb or less.

30. Consider alternative diagnoses and referral for specialist assessment in children and young people (aged 5 to 16) if they have symptoms suggestive of asthma but normal spirometry, a FeNO level of 34 ppb or less and negative peak flow variability.

---

decision to use this test. Informed consent should be obtained and documented. See the General Medical Council's Prescribing guidance: prescribing unlicensed medicines for further information.

[d] At the time of publication (November 2017), histamine and methacholine did not have UK marketing authorisation for this use. The healthcare professional should follow relevant professional guidance, taking full responsibility for the decision to use this test. Informed consent should be obtained and documented. See the General Medical Council's Prescribing guidance: prescribing unlicensed medicines for further information.

Diagnosis in adults aged 17 and over

See also algorithm C for objective tests in adults aged 17 and over.

31. Diagnose asthma in adults (aged 17 and over) if they have symptoms suggestive of asthma and:

- a FeNO level of 40 ppb or more with either positive bronchodilator reversibility or positive peak flow variability, or bronchial hyperreactivity, **or**
- a FeNO level between 25 and 39 ppb and a positive bronchial challenge test, **or**
- positive bronchodilator reversibility and positive peak flow variability irrespective of FeNO level.

32. Suspect asthma in adults (aged 17 and over) with symptoms suggestive of asthma, obstructive spirometry and:

- negative bronchodilator reversibility, and either a FeNO level of 40 ppb or more, **or**
- a FeNO level between 25 and 39 ppb and positive peak flow variability, **or**
- positive bronchodilator reversibility, a FeNO level between 25 and 39 ppb and negative peak flow variability.

Do not rule out other diagnoses if symptom control continues to remain poor after treatment. Review the diagnosis after 6 to 10 weeks by repeating spirometry and objective measures of asthma control and reviewing symptoms.

33. Consider alternative diagnoses, or referral for a second opinion, in adults (aged 17 and over) with symptoms suggestive of asthma, and:

- a FeNO level below 40 ppb, normal spirometry and positive peak flow variability, **or**
- a FeNO level of 40 ppb or more but normal spirometry, negative peak flow variability, and negative bronchial challenge test, **or**
- obstructive spirometry with bronchodilator reversibility, but a FeNO level below 25 ppb, and negative peak flow variability, **or**
- positive peak flow variability but normal spirometry, a FeNO level below 40 ppb, and a negative bronchial challenge test, **or**
- obstructive spirometry with negative bronchodilator reversibility, a FeNO level below 25 ppb, and a negative peak flow variability (if measured).

Diagnosis in people who are unable to perform an objective test

For young children who cannot perform objective tests, see recommendations 12 and 13.

34. If an adult, young person or child with symptoms suggestive of asthma cannot perform a particular test, try to perform at least 2 other objective tests. Diagnose suspected asthma based on symptoms and any positive objective test results.

Good clinical practice in asthma diagnosis

35. Record the basis for a diagnosis of asthma in a single entry in the person's medical records, alongside the coded diagnostic entry.

Summary of objective test results for adults, young people and children (over 5)

Algorithms have been produced that summarise objective testing for asthma in adults, young people and children (over 5). Table 7 summarises the objective test threshold levels.

Table 7: Positive test thresholds for objective tests for adults, young people and children (aged 5 and over)

Test	Population	Positive result
FeNO	Adults	40 ppb or more
	Children and young people	35 ppb or more
Obstructive spirometry	Adults, young people and children	FEV1/FVC ratio less than 70% (or below the lower limit of normal if this value is available)
Bronchodilator reversibility (BDR) test	Adults	Improvement in FEV1 of 12% or more and increase in volume of 200 ml or more
	Children and young people	Improvement in FEV1 of 12% or more
Peak flow variability	Adults, young people and children	Variability over 20%
Direct bronchial challenge test with histamine or methacholine	Adults	PC20 of 8 mg/ml or less
	Children and young people	n/a
Abbreviations: FeNO, fractional exhaled nitric oxide; FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity; PC20, provocative concentration of methacholine causing a 20% fall in FEV1.		

Monitoring asthma control

36. Monitor asthma control at every review. If control is suboptimal:

- confirm the person's adherence to prescribed treatment in line with the recommendations on assessing adherence in the NICE guideline on medicines adherence
- review the person's inhaler technique
- review if treatment needs to be changed
- ask about occupational asthma (see recommendation 10) and/or other triggers, if relevant.

37. Consider using a validated questionnaire (for example, the Asthma Control Questionnaire or Asthma Control Test) to monitor asthma control in adults (aged 17 and over).

38. Monitor asthma control at each review in adults, young people and children aged 5 and over using either spirometry or peak flow variability testing.

39. Do not routinely use FeNO to monitor asthma control.

40. Consider FeNO measurement as an option to support asthma management in people who are symptomatic despite using inhaled corticosteroids. (This recommendation is from NICE's diagnostics guidance on measuring fractional exhaled nitric oxide concentration in asthma.)

41. Do not use challenge testing to monitor asthma control.

42. Observe and give advice on the person's inhaler technique:

- at every consultation relating to an asthma attack, in all care settings
- when there is deterioration in asthma control

- when the inhaler device is changed
- at every annual review
- if the person asks for it to be checked.

4.3 Key research recommendations

Diagnosing asthma in children and young people aged 5 to 16

- What is the acceptability and diagnostic accuracy of objective tests that could be used to comprise a diagnostic pathway for asthma in children and young people aged 5 to 16 (for example, exercise challenge, direct bronchial challenge with histamine or methacholine, indirect bronchial challenge with mannitol and peripheral blood eosinophil count)?

Diagnosing asthma in adults (aged 17 and over)

- What is the clinical and cost effectiveness of using an indirect bronchial challenge test with mannitol to diagnose asthma in adults (aged 17 and over)?

Monitoring adherence to treatment

- What is the clinical and cost effectiveness of using electronic alert systems designed to monitor and improve adherence with regular inhaled maintenance therapy in people with asthma?

Monitoring inhaler technique

- What is the current frequency and the current method being used to check the inhaler technique of people with asthma? What is the optimal frequency and the best method of checking inhaler technique to improve clinical outcomes for people with asthma?

Monitoring asthma control using tele-healthcare

- What is the long-term (more than 12 months) clinical and cost effectiveness of using tele-healthcare as a means to monitor asthma control in adults, young people and children? Methods of tele-healthcare can include telephone interview (with healthcare professional involvement) and internet or smartphone-based monitoring support (no healthcare professional involvement).

5 Diagnosing asthma

5.1 Initial clinical assessment

Chapters 6 to 10 review the diagnostic accuracy of the initial clinical assessment questions for the diagnosis of asthma in people with suspected asthma presenting with respiratory symptoms.

5.2 Objective tests

Chapters 11 to 20 review the diagnostic test accuracy of objective tests for the diagnosis of asthma in people with suspected asthma presenting with respiratory symptoms.

6 Diagnosis: Signs and symptoms

6.1 Introduction

There are several signs and symptoms associated with (but not specific for) asthma. Although at one time these symptoms were likely to have been under-interpreted (leading to under-diagnosis or delayed diagnosis), now they are over-interpreted (leading to over-diagnosis if not supported by objective tests). It is important to identify asthma-related signs and symptoms from the history of presenting complaints. However, the diagnostic test accuracy of asking about asthma signs and symptoms is currently uncertain. Asthma signs and symptoms can vary from mild, moderate to severe. They can also vary throughout the year depending on the season or exposure to variable environmental triggers, such as viral infections, allergens and air pollution. They also vary with age; asthma is most common in younger age groups. The early identification of asthma-related signs and symptoms allows early diagnostic testing and appropriate treatment decisions.

6.2 Review question: In people under investigation for asthma, what is the diagnostic accuracy of each of the following signs and symptoms?

- **wheezing**
- **cough**
- **breathlessness**
- **nocturnal symptoms**
- **diurnal and seasonal variations**

For full details see review protocol in Appendix C.

Table 8: Characteristics of review question

Component	Description
Population / Target condition	People with suspected asthma (presenting with respiratory symptoms). All ages, stratified into the following 3 different groups: • Children (1-<5 years old) • Children/young people (5-16 years old) • Adults (>16 years old)
Index test	Signs and symptoms of asthma Each of the following symptoms alone or in combination: • Wheezing (current or persistent or triggered) • Cough (including nocturnal cough) • Breathlessness • Nocturnal symptoms • Diurnal and seasonal variations
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test In children 1-<5 years, objective tests cannot be performed so the reference standard will be physician diagnosis based on recurrent and persistent wheezing.
Outcomes	• Diagnostic accuracy (sensitivity and specificity)

6.3 Clinical evidence

We searched for cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of signs or symptoms of asthma to identify whether the condition is present (as indicated by the reference standard) in people under investigation for asthma.

Six studies[33,151-153,177,193] were included in the review (see Table 9 and Table 10). Evidence from these are summarised in the clinical evidence profile below (Table 11). See also the study selection flow chart in Appendix D, sensitivity / specificity forest plots in Appendix J, study evidence tables in Appendix G and exclusion list in Appendix K. All studies were conducted in adults, except for Weverhess 1999, which was in children <5 years old. No evidence was identified for the strata of children aged 5-16 years. Data have been separated for each age group.

A variety of index tests symptoms or combinations of symptoms were used (Table 9 and Table 10). The accuracy of individual symptoms in the diagnosis of asthma are analysed and reported separately (unless combined into a symptom score by the individual study) as the GC was interested in which individual symptoms indicate asthma.

None of the studies reported the diagnostic accuracy of signs.

The reference standard was physician's diagnosis of asthma with an objective test, with the exception of the study in children <5 years, as objective tests cannot be performed in this age group. A variety of objective tests and thresholds were used for the reference standard (see Table 9 and Table 10).

In anticipation of there being a large number of studies retrieved from the search, inclusion was limited to studies of populations in the UK, USA, Australia, Canada, New Zealand and Western Europe. These countries were expected to be similar to the UK in terms of how people report symptoms and the impact of language. If relevant studies were identified from other review questions reporting populations outside these countries, then these were included.

Summary of included studies

Table 9: Summary of studies included in the review: diagnostic accuracy of symptoms vs. physician Dx with an objective test (adults)

Study	Presentation	Target condition	Index test	Reference standard	Comments
CHOI 2007[33]	• Adults • Respiratory symptoms such as dyspnoea, cough or wheezing	Asthma	Symptoms (questionnaire)	Physician Dx with objective test (patients with an FEV1 >70% had MCT, all other patients had BDR to short-acting beta2-agonist). Definite Dx of asthma made using test (MCh PC20 <16mg/ml or BDR FEV1 increase >12% and 200ml)	
SCHLEICH 2012[151]	• Adults • Patients referred to chest physicians for methacholine challenge test for asthma diagnosis; bronchodilator test failed to show reversible airway obstruction or baseline spirometry normal.	Asthma	Symptoms (questionnaire)	Methacholine challenge (cut off PC20 <16mg/mL).	
SCHNEIDER 2009A[153]	• Adults • Visiting GP for the first time with complaints of suggested obstructive airway disease (OAD). Symptoms such as dyspnoea, coughing, or expectoration	Asthma	Symptoms (questionnaire)	Dx by respiratory physician based on whole-body plethysmography (FEV1/VC ≤70% or FEV1 <80%) followed by either BDR if obstruction is present (FEV1 ≥12% and ≥200ml) or methacholine if obstruction is not present (PC20 ≤16mg/ml or extreme increase in airway resistance accompanied by clinical symptoms in two patients)	
SCHNEIDER 2012[152]	• Adults **GPs:** first time visit with complaints of suggested OAD or RAD; symptoms for >2 months (data presented in Schneider 2009) **Respiratory physician:** 1st visit for Dx work-up to include or exclude OAD or	Asthma	Symptoms (questionnaire)	Dx by respiratory physician based on whole-body plethysmography (FEV1/VC ≤70% or FEV1 <80%) followed by either BDR if obstruction is present (FEV1 ≥12% and ≥200ml) or methacholine if obstruction is not present (PC20 ≤16mg/ml). Most asthma patients were identified by the BPT.	

Study	Presentation	Target condition	Index test	Reference standard	Comments
	RAD; other criteria as for GPs. **Hospital:** Patients with suspected OAD who were hospitalised for the first time.				
TOMITA 2013[177]	• Adults • Outpatients with non-specific respiratory symptoms including wheeze, shortness of breath, and cough	Asthma	Symptoms (questionnaire/ interview)	Relevant symptom history (all patients) and BDR (FEV1 <200ml and 12%) and/or BHR (methacholine PC20 <8mg/ml)	

Table 10: Summary of studies included in the review: diagnostic accuracy of symptoms vs. physician Dx with an objective test (children <5 years)

Study	Presentation	Target condition	Index test	Reference standard	Comments
WEVERHESS 1999[193]	• Children aged 0-4 years • Symptoms that were suggestive of asthma	Asthma	Symptoms (visit and questionnaire)	Dx taken from medical notes at follow-up (2 years later). Dx made by a paediatrician on clinical grounds, based on recurrence of symptoms and need for and response to therapy according to the guidelines for diagnosis of asthma in young children (follow up statement from the International Paediatric Asthma Consensus Group).	

Table 11: Clinical evidence profile: Symptoms vs Reference Standard (physician Dx and objective test where appropriate to the age group)

Index Test (Threshold)	No of studies	n	Risk of bias(a)	Inconsistency	Indirectness	Imprecision	Sensitivity % (range)	Specificity % (range)	Area Under Curve (range)	Quality
ADULTS >16 years										
Paroxsymal coughing	1	302	Serious risk of bias[(a)]	No serious inconsistency	No serious indirectness	n/a[(c)]	16	42	-	MODERATE
Dyspnoea without wheeze	1	302	Serious risk of bias[(a)]	No serious inconsistency	No serious indirectness	n/a[(c)]	11	71	-	MODERATE
Wheeze without dyspnoea	1	302	Serious risk of bias[(a)]	No serious inconsistency	No serious indirectness	n/a[(c)]	9	79	-	MODERATE
Diurnal cough	1	174	Serious risk of bias[(a)]	No serious inconsistency	Serious indirectness[(d)(e)]	n/a[(c)]	66	26	-	LOW
Nocturnal cough	1	174	Serious risk of bias[(a)]	No serious inconsistency	Serious indirectness[(d)(e)]	n/a[(c)]	37	65	-	LOW
Diurnal wheeze	1	174	Serious risk of bias[(a)]	No serious inconsistency	Serious indirectness[(d)(e)]	n/a[(c)]	57	62	-	LOW
Nocturnal wheeze	1	174	Serious risk of bias[(a)]	No serious inconsistency	Serious indirectness[(d)(e)]	n/a[(c)]	56	79	-	LOW
Dyspnoea	2	393	Serious risk of bias[(a)]	Serious inconsistency[(b)]	Serious indirectness[(e)]	n/a[(c)]	Range 61 – 73	Range 38 – 55	-	LOW
Wheeze	2	785	Serious risk of bias[(a)]	Serious inconsistency[(b)]	No serious indirectness	n/a[(c)]	Range 30 – 52	Range 53 – 87	-	LOW
Cough	1	219	Serious risk of bias[(a)]	No serious inconsistency	Serious indirectness[(e)]	n/a[(c)]	43	33	-	LOW
Nocturnal dyspnoea	1	219	Serious risk of bias[(a)]	No serious inconsistency	Serious indirectness[(e)]	n/a[(c)]	30	81	-	LOW

Index Test (Threshold)	No of studies	n	Risk of bias(a)	Inconsistency	Indirectness	Imprecision	Sensitivity % (range)	Specificity % (range)	Area Under Curve (range)	Quality
Diurnal symptoms	1	566	Serious risk of bias[(a)]	No serious inconsistency	No serious indirectness	n/a[(c)]	54	69	-	MODERATE
Total symptom score ≥5	1	302	Serious risk of bias[(a)]	No serious inconsistency	No serious indirectness	n/a[(c)]	74	48	-	MODERATE
Dyspnoea attacks	1	219	No risk of bias[(a)]	No serious inconsistency	Very serious indirectness[(e)(f)]	n/a[(c)]	40	78	-	LOW
Dyspnoea going upstairs	1	219	No risk of bias[(a)]	No serious inconsistency	Very serious indirectness[(e)(f)]	n/a[(c)]	47	49	-	LOW
Dyspnoea when walking	1	219	No risk of bias[(a)]	No serious inconsistency	Very serious indirectness[(e)(f)]	n/a[(c)]	4.8	93.2	-	LOW
Dyspnoea on minimal exercise	1	219	No risk of bias[(a)]	No serious inconsistency	Very serious indirectness[(e)(f)]	n/a[(c)]	2.5	94	-	LOW
CHILDREN 5-16 years										
No evidence identified	0									
CHILDREN <5 years										
Cough and wheeze	1	188	Serious risk of bias[(a)]	No serious inconsistency	No serious indirectness	n/a[(c)]	49	59	-	MODERATE
Dyspnoea	1	188	Serious risk of bias[(a)]	No serious inconsistency	No serious indirectness	n/a[(c)]	76	52	-	MODERATE
Wheeze	1	188	Serious risk of bias[(a)]	No serious inconsistency	No serious indirectness	n/a[(c)]	54	57	-	MODERATE
Cough	1	188	Serious risk of bias[(a)]	No serious inconsistency	No serious indirectness	n/a[(c)]	88	7	-	MODERATE

(a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.

(b) Inconsistency was assessed by inspection of the sensitivity / specificity values in each of the studies in relation to a high (≥75%), moderate (50-74%) or low (<50%) effect estimate. Inconsistency was considered high if the point estimates from individual studies were within two areas, and very high if they were in 3 areas or both high and low areas.
(c) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed.
(d) Population included people with a normal spirometry or a normal BDR.
(e) Reference standard objective test cut-off threshold did not match protocol.
(f) Population included people who had been hospitalised due to suspected obstructive airways disease. Index test was based on anamnestic data.

6.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow diagram in Appendix E.

6.5 Evidence statements

Clinical

- One study with 302 adults showed that symptoms of paroxsymal coughing has a sensitivity of 0.16 and a corresponding specificity of 0.42 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 302 adults showed that symptoms of dyspnoea without wheeze has a sensitivity of 0.11 and a corresponding specificity of 0.71 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 302 adults showed that symptoms of wheeze without dyspnoea has a sensitivity of 0.09 and a corresponding specificity of 0.79 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 174 adults showed that symptoms of diurnal cough has a sensitivity of 0.66 and a corresponding specificity of 0.26 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 174 adults showed that symptoms of nocturnal cough has a sensitivity of 0.37 and a corresponding specificity of 0.65 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 174 adults showed that symptoms of diurnal wheeze has a sensitivity of 0.57 and a corresponding specificity of 0.62 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 174 adults showed that symptoms of nocturnal wheeze has a sensitivity of 0.56 and a corresponding specificity of 0.79 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- Two studies with 393 adults showed that symptoms of dyspnoea has a sensitivity range of 0.61 to 0.73 and a corresponding specificity range of 0.38 to 0.55 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- Two studies with 785 adults showed that symptoms of wheeze has a sensitivity range of 0.30 to 0.52 and a corresponding specificity range of 0.53 to 0.87 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 219 adults showed that symptoms of cough has a sensitivity of 0.43 and a corresponding specificity of 0.33 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 219 adults showed that symptoms of nocturnal dyspnoea has a sensitivity of 0.30 and a corresponding specificity of 0.81 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 566 adults showed that diurnal symptoms has a sensitivity of 0.54 and a corresponding specificity of 0.69 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)

- One study with 302 adults showed that a total symptom score ≥5 has a sensitivity of 0.74 and a corresponding specificity of 0.48 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 219 adults showed that symptoms of dyspnoea attacks has a sensitivity of 0.40 and a corresponding specificity of 0.78 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 219 adults showed that symptoms of dyspnoea going upstairs has a sensitivity of 0.47 and a corresponding specificity of 0.49 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 219 adults showed that symptoms of dyspnoea when walking has a sensitivity of 0.05 and a corresponding specificity of 0.93 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 219 adults showed that symptoms of dyspnoea on minimal exercise has a sensitivity of 0.03 and a corresponding specificity of 0.94 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 188 children <5 years showed that symptoms of cough and wheeze has a sensitivity of 0.49 and a corresponding specificity of 0.59 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 188 children <5 years showed that symptoms of dyspnoea has a sensitivity of 0.76 and a corresponding specificity of 0.52 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 188 children <5 years showed that symptoms of wheeze has a sensitivity of 0.54 and a corresponding specificity of 0.57 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 188 children <5 years showed that symptoms of cough has a sensitivity of 0.88 and a corresponding specificity of 0.07 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- No evidence was identified in children aged 5-16 years.

Economic

- No relevant economic evaluations were identified.

6.6 Recommendations and link to evidence

Recommendations	**1. Take a structured clinical history in people with suspected asthma. Specifically, check for:** • **wheeze, cough or breathlessness, and any daily or seasonal variation in these symptoms** • **any triggers that make symptoms worse** • **a personal or family history of atopic disorders.** **2. Do not use symptoms alone without an objective test to diagnose asthma.** **3. Examine people with suspected asthma to identify expiratory polyphonic wheeze and signs of other causes of respiratory symptoms, but be aware that even if examination results are normal the person may still have asthma.**

	4. Treat people immediately if they are acutely unwell at presentation, and perform objective tests for asthma (for example, fractional exhaled nitric oxide [FeNO], spirometry and peak flow variability) if the equipment is available and testing will not compromise treatment of the acute episode. **5. If objective tests for asthma cannot be done immediately for people who are acutely unwell at presentation, carry them out when acute symptoms have been controlled and advise people to contact their healthcare professional immediately if they become unwell while waiting to have objective tests.** **6. Be aware that the results of spirometry and FeNO tests may be affected in people who have been treated empirically with inhaled corticosteroids.**
Relative values of different diagnostic measures and outcomes	The GC was interested in the diagnostic test accuracy of signs and symptoms in the diagnosis of asthma. There is evidence of significant misdiagnosis of asthma, with both false positive and false negative errors (see introduction to the guideline). The GC was interested in both the sensitivity and specificity of diagnostic tests and prioritised studies which reported these outcomes. Diagnostic studies were found comparing the sensitivity and specificity of symptoms (individual symptoms or total symptoms scores) for the diagnosis of asthma vs. the reference standard (physician's diagnosis based on symptoms plus an objective test). The symptoms covered by the studies included coughing, dyspnoea (breathlessness), wheeze, nocturnal symptoms, diurnal symptoms, as well as various combinations of these. None of the studies assessed seasonal variations in symptoms. No evidence was found on the signs of asthma.
Trade-off between clinical benefits and harms	There is no direct harm associated with asking symptom questions and the trade-off between benefit and harm will depend on their reliability as a means of indicating asthma. There was a large variation in the sensitivity and specificity values across the different individual symptoms. The sensitivities and specificities for the majority of the individual symptoms were moderate or low. None of the symptoms had a combination of both high sensitivity and high specificity. Evidence was available from one study for the diagnostic accuracy of a combined symptom score. Although this had a moderate sensitivity, the specificity was low. The sensitivity and specificity was not high enough for the GC to recommend using symptoms in isolation to diagnose asthma. In addition, the GC considered that many of the symptoms are also common to a variety of other respiratory conditions, and so to make a diagnosis of asthma based on symptoms alone would be both inaccurate and inappropriate.
Economic considerations	No economic evaluations were found on this question. The sensitivity and specificity of diagnosing asthma with symptoms alone was shown to be low for the majority of individual symptoms. Therefore, an asthma diagnosis based on individual symptoms alone would lead to a large number of false negatives and false positives. This would be of clinical harm to individuals who have asthma and who would go untreated for a period of time as well as those without asthma who would receive unnecessary treatment. Therefore, even though the cost of a history and examination is low, using it as a stand-alone diagnostic test would lead to more harm than good because it will lead to inappropriate over-treatment, waste of resources and lack of a correct diagnosis.

Quality of evidence	Most of the evidence found for the diagnostic value for each symptom was based on single studies. The quality of the evidence ranged from low to moderate in adults. There was limited evidence with the ideal reference standard, therefore evidence was included from studies using an alternative reference standard (four studies included a methacholine test with a cut-off value of 16mg/ml as part of the index test). This will affect the number of people diagnosed with asthma using the reference standard, and therefore the accuracy of the index test. Therefore the evidence quality was downgraded for indirectness. One study included people who had been hospitalised with suspected obstructive airways disease and reflects the diagnostic accuracy of symptoms in a different population. Again, this evidence quality was downgraded for indirectness and these limitations were taken into account by the GC when interpreting the evidence quality. The quality of the evidence in children aged <5 years was moderate, but evidence was only available from one study. The GC discussed the reference standard for this study and agreed it was an appropriate reference standard for the diagnosis of asthma in this age group. In children aged 5-16 years, no diagnostic studies were found.
Other considerations	The GC pointed out that all of the respiratory symptoms associated with asthma also occur in other conditions, and therefore an asthma diagnosis should not be made on the basis of symptoms alone without objective testing. However, a history of symptoms compatible with asthma is highly important in interpreting any subsequent tests. The GC also discussed the importance of taking a clinical history of symptoms at initial presentation and at each asthma review in order to document asthma control and to identify triggers. The GC agreed that this should be done using a structured template and referred to the recent NRAD recommendations[144] for a standard national asthma template to facilitate a structured, thorough asthma review. No formal evidence was identified on the diagnostic accuracy of clinical signs. However, the GC also agreed by consensus that it is important to physically examine people with suspected asthma since the identification of expiratory polyphonic wheeze is indicative of airflow obstruction when present; is good evidence of asthma when it varies over time; and because signs may indicate other causes of respiratory symptoms. The GC agreed that patients who are unwell at presentation, and regarded as probably having asthma as the cause of their symptom, should be treated for the presumptive asthma diagnosis while awaiting further tests. These patients may be at risk if treatment is delayed until after the diagnosis is confirmed. It was agreed that some tests could be performed immediately at presentation (including spirometry and FeNO). For other tests that may not be immediately available, treatment should commence and these objective tests should be performed when the acute symptoms have been controlled. However, the GC stressed that the diagnosis of asthma should not be made on the basis of response of symptoms alone to treatment, and recommended that a formal diagnosis of asthma should not be made until objective tests have been performed. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

7 Diagnosis: History of atopic disorders

7.1 Introduction

The term atopy refers to allergic conditions which include allergic rhinitis (hay fever), atopic dermatitis (eczema), allergic asthma and other specific and non-specific allergic problems like food allergies. There is considerable overlap between these conditions; however, the link between these different atopic disorders is not well understood. As these conditions often co-exist in the same individual and tend to cluster in families, it is of interest to know whether taking a personal or family history of atopic disorders is accurate in the diagnosis of asthma in people with asthma symptoms.

7.2 Review question: In people under investigation for asthma, what is the diagnostic accuracy of taking a personal/family history of atopic disorders?

For full details see review protocol in Appendix C.

Table 12: Characteristics of review question

Population	People with suspected asthma (presenting with respiratory symptoms). All ages, stratified into the following 3 different groups: • Children (1-<5 years old) • Children/young people (5-16 years old) • Adults (>16 years old)
Index test	Personal/family history of atopic disorders. • This is likely to be ascertained by a questionnaire. NOTE: personal history is defined as an individual who has had one of the atopic disorders listed below NOTE: family history is defined as: 1st degree relatives. NOTE: atopic disorders are defined as: eczema, hay fever, allergic rhinitis, food allergy, asthma.
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test In children 1-<5 years, objective tests cannot be performed so the reference standard will be physician diagnosis based on recurrent and persistent wheezing.
Statistical measures	Diagnostic accuracy (sensitivity and specificity)

7.3 Clinical evidence

We searched for cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of a taking a personal or family history of atopic disorders to identify whether the condition is present (as indicated by the reference standard) in people under investigation for asthma.

Five studies were included in the review[37,43,177,183,193] (see Table 13). Evidence from these studies is summarised in the clinical evidence profile below (Table 14). See also the study selection flow chart in Appendix D, sensitivity / specificity forest plots in Appendix J, study evidence tables in Appendix G and exclusion list in Appendix K.

All included studies looked at a personal/family history of atopic disorders in patients with signs and symptoms. The reference standard was physician's diagnosis of asthma with an objective test, with the exception of the study in children <5 years, as objective tests cannot be performed in this age group. In adults, evidence was available from three studies[37,43,177], in children aged 5-16 years evidence was available from one study[183] and in children aged <5 years evidence was available from one study[193].

Summary of included studies

Table 13: Summary of studies included in the review

Study	Population	Index test & cut-off	Reference standard
Index test vs Reference Standard			
CORDEIRO 2011[37]	N = 114 Adults and children/young people Referrals to allergy clinic with symptoms of nasal or ocular complaints; pulmonary complaints; skin complaints and general complaints.	Family history (unclear if first degree relatives)	History of typical respiratory symptoms and FEV1 improvement >12% and >200mL with salbutamol 400µg or PC20 histamine ≤8mg/mL according to GINA.
Deilami 2009[43]	N = 81 Adults Suffering from cough for at least 8 weeks and went to the pulmonary disease clinic. Normal spirometry	Personal history of allergy NB Family history of asthma sensitivity/specificity data were not extracted as was not first class relatives only	Only objective test (not physician Dx with objective test). Methacholine challenge test: PC20 ≤4mg/ml
TOMITA 2013[177]	N = 566 Adults Adult outpatients with non-specific respiratory symptoms including wheeze, shortness of breath, and cough.	Routine interview including following questions: a) Personal history: 'Have you had any medical history of allergic diseases such as asthma, atopic dermatitis, and allergic rhinitis?' b) Family history: 'Do you have any close relatives with allergic disease?'	Relevant symptom history (all patients) and BDR (FEV1 <200ml and 12%) and/or BHR (methacholine PC20 <8mg/ml) NB. 64/367 were diagnosed on the basis of reactivity to ICS without BDR or BHR)
WEVER-HESS 1999[193]	N = 188 Children (including aged 2-4yr subgroup only) Aged 0-4 years with symptoms that were suggestive of asthma	History taken at initial visit: a) Past or present rhinitis b) past or present eczema c) family history	Dx taken from medical notes at follow-up (2 years later). Dx made by a paediatrician on clinical grounds, based on recurrence of symptoms and need for and response to therapy according to the guidelines for diagnosis of asthma in young children.
VANDERMARK 2014[183]	N = 438 Children aged 1-5 years included in longitudinal	a) Family history of asthma (parents and/or siblings)	At age 6 years, spirometry and BHR obtained in children with wheezing, shortness of breath, recurrent coughing or

Study	Population	Index test & cut-off	Reference standard
	study (asthma Dx at age 6 years). Presented in primary care in the previous 12 months with current coughing (≥2 visits), wheezing (≥1 visits), and/or shortness of breath (≥1 visits)		use of asthma medication during the previous 12 months. Dx defined as having persistent symptoms and/or using asthma medication in the last year in combination with BHR (methacholine <8mg.ml) or BDR (>10% increase in FEV1).

Table 14: Clinical evidence profile: Index test vs Reference Standard (physician Dx and objective test)

Index Test (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity %	Specificity %	Area Under Curve (range)	Quality
ADULTS										
Personal history of atopic disorders	2	656	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	n/a[c]	Range 54.2 - 55	Range 67.8 - 73.7	-	MODERATE
Family history of atopic disorders	2	680	Serious risk of bias[a]	Serious inconsistency[b]	No serious indirectness	n/a[c]	Range 25.9 - 59.5	Range 55.6 - 82.9	-	LOW
CHILDREN 5-16 years										
Family history of asthma	1	438	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	n/a[c]	43.8	69.7	-	MODERATE
CHILDREN <5 years										
Family history of atopic disorders	1	188	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	n/a[c]	43.8	56.8	-	MODERATE
Personal history of atopic disorders – rhinitis only	1	188	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	n/a[c]	61.8	20.5	-	MODERATE
Personal history of atopic disorders – eczema only	1	188	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	n/a[c]	46.5	75.0	-	MODERATE

(a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.

(b) Inconsistency was assessed by inspection of the sensitivity / specificity values in each of the studies in relation to a high (≥75%), moderate (50-74%) or low (<50%) effect estimate. Inconsistency was considered high if the point estimates from individual studies were within two areas, and very high if they were in 3 areas or both high and low areas

(c) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed.

7.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow diagram in Appendix E.

7.5 Evidence statements

Clinical

- Two studies with 656 adults showed that a personal history of atopic disorders has a sensitivity range of 54.2-55% and a corresponding specificity range of 67.8-73.7% for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- Two studies with 680 adults showed that a family history of atopic disorders has a sensitivity range of 25.9-59.5% and a corresponding specificity range of 55.6-82.9% for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 438 children 5-16 years showed that a family history of asthma has a sensitivity of 43.8% and a corresponding specificity of 69.7% for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 188 children <5 years showed that a family history of atopic disorders has a sensitivity of 43.8% and a corresponding specificity of 56.8% for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 188 children <5 years showed that a personal history of rhinitis has a sensitivity of 61.8% and a corresponding specificity of 20.5% for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 188 children <5 years showed that a personal history of eczema has a sensitivity of 46.5% and a corresponding specificity of 75.0% for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)

Economic

- No relevant economic evaluations were identified.

7.6 Recommendations and link to evidence

Recommendations	**7. Ask about a personal or family history of atopic disorders, when taking a structured clinical history in people with suspected asthma. Record any triggers that make symptoms worse.** **8. Do not use a history of atopic disorders alone to diagnose asthma.**
Relative values of different outcomes	The GC was interested in the diagnostic test accuracy of a personal or family history of atopy in the diagnosis of asthma. There is evidence of significant misdiagnosis of asthma, with both false positive and false negative errors (see introduction to the guideline). The GC was interested in both the sensitivity and specificity of diagnostic tests and prioritised studies which reported these outcomes. The GC wished to know whether asking about a personal or family history of atopic disorders was an accurate method of diagnosing (or contributing to the diagnosis) of asthma. Atopic disorders are known to have a hereditary component and it has been

	suggested that taking a family history of atopic disorders in a person with suspected asthma may aid diagnosis. Allergic asthma is an IgE-driven disease similar to conditions like eczema and seasonal rhinitis. Together, these conditions form the atopic spectrum of disorders and these frequently co-exist in individuals with confirmed asthma. We therefore considered whether the presence or absence of eczema and allergic rhinitis in patients with recurrent respiratory symptoms predicts the presence or absence of asthma in these people. Atopic disorders were defined by the GC as eczema, hay fever, allergic rhinitis, food allergy and asthma. Evidence for these atopic disorders was considered separately for personal history and family history.
Trade-off between clinical benefits and harms	There is no direct harm associated with asking these questions (aside from the time taken) and the trade-off between benefit and harm will depend on the reliability of the clinical history questions as a means of indicating atopic asthma. The GC noted that the evidence demonstrates the following when asking questions about: **Personal history** Adults: • Low to moderate sensitivity and moderate to high specificity. Children: • Rhinitis had moderate sensitivity and low specificity. • Eczema had moderate sensitivity and high specificity. **Family history** Adults: • Low to moderate sensitivity and moderate to high specificity. Children: • Low sensitivity and moderate specificity. The GC acknowledged that the prevalence of atopy is high in people with asthma and the GC consensus opinion was that around 1 in 3 people in the UK are atopic. Overall, the GC agreed that the sensitivity and specificity of asking these questions was not high enough for the GC to recommend using these questions in isolation to diagnose atopic asthma. The GC discussed the importance of asking these questions as part of a diagnostic assessment that includes clinical history, taken in addition to performing other objective tests. The GC noted that these questions will only help to identify people with atopic asthma, and that taking a history of atopic disorders would be of little benefit in adults with occupational asthma or late onset or non-atopic asthma. Importantly, the group acknowledged that in young children the use of objective tests to diagnose atopic asthma is limited and until further research is available in relation to these tests, more weight would need to be placed on clinical history questions.
Economic considerations	No economic evaluations were found on this question. The sensitivity of diagnosing asthma using a history of atopic disorders was shown to be low for the majority of individual symptoms. Therefore, an asthma diagnosis based on a history of atopic disorders alone would lead to a large number of false negatives. This would be of clinical harm to individuals who have asthma and who would go untreated for a period of time. However as the cost of asking questions related to atopy are negligible and there

	may be some value in the information gained, the GC recognised this as a useful tool to inform management once the diagnosis is made.
Quality of evidence	In adults, evidence from three studies was included[37,43,177]. In one study[43], the reference standard was symptoms plus a positive methacholine test with a cut-off of 4mg/ml. The remaining studies used physician diagnosis plus an objective test as the reference standard. Although one study (Cordeiro 2011) had an indirect population with general allergic symptoms rather than respiratory symptoms, this did not affect the overall quality of the evidence as the study did not contribute the majority of the evidence. Evidence for a family history of atopic disorders was of low quality and for a personal history of atopic disorders was of moderate quality. In children aged 5-16 years evidence from one moderate quality study was included [183]. In children aged <5 years, evidence from one moderate quality study was included[193]. This study used a reference standard diagnosis made by a paediatrician on clinical grounds, based on recurrence of symptoms, and need for and response to therapy in accordance with guidelines for the diagnosis of asthma in young children (statement from an international paediatric asthma consensus group). This longitudinal study looked at prognostic factors for asthma, and the final diagnosis was made at two-year follow-up after initial assessment. The GC agreed that this study was appropriate to include given that in children <5 years of age the diagnosis is usually made at follow-up based on reoccurrence of symptoms and response to therapy (rather than objective lung function tests which young children are unable to perform accurately). The Wever-Hess et al. study provided evidence for a family history of atopic disorders and a personal history of rhinitis and/or eczema.
Other considerations	The GC discussed whether asking these questions to identify patients with atopic asthma would guide treatment decisions. Initial treatment would be similar regardless of atopic status; however, this information may be useful in guiding the monitoring of asthma, for example, avoidance of allergens. A key finding of the RCP National Review of Asthma Deaths (NRAD) report[144], published in May 2014, was that exacerbating factors or triggers, such as atopy, were documented in the records of less than half of the patients in the review. In light of this, the GC made a recommendation to address poor documentation (aligning with the report findings). The GC agreed that it is vitally important to ensure that a personal or family history of atopic disorders is accurately and comprehensively documented in the medical records of all people presenting with potential atopic asthma. Whilst this information is insufficient to make a diagnosis, it is important for trigger identification and subsequent monitoring of asthma. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

8 Diagnosis: Symptoms after exercise

8.1 Introduction

Symptoms of cough, wheeze and chest tightness suggestive of bronchoconstriction can occur after exercise in many people with asthma, and also in some people who do not have asthma. The symptoms are associated with prolonged exercise, such as long-distance running, rather than short bursts of intensive exercise. Classically, symptoms occur a few minutes after stopping exercise, or less commonly during exercise, and can vary from mild symptoms to an acute asthma attack. Treatment with beta-2-agonists prior to exercise reduces or eliminates symptoms of exercise-induced bronchoconstriction. However, the diagnostic test accuracy of asking about a history of symptoms after exercise to diagnose asthma is currently unclear.

8.2 Review question: In people under investigation for asthma, what is the diagnostic accuracy of a clinical history of symptoms in response to exercise?

For full details see review protocol in Appendix C.

Table 15: Characteristics of review question

Population	People with suspected asthma (presenting with respiratory symptoms). All ages, stratified into the following 3 different groups: • Children (1- <5 years old) • Children/young people (5-16 years old) • Adults (>16 years old)
Index test	• Clinical history of symptoms in response to exercise. NOTE: symptoms would be a combination of the following, or individual symptoms – wheezing, cough, breathlessness (over and above what you would expect during exercise).
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test In children 1-<5 years, objective tests cannot be performed so the reference standard will be physician diagnosis based on recurrent and persistent wheezing.
Statistical measures	Diagnostic accuracy (sensitivity, specificity)

8.3 Clinical evidence

We searched for cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of a clinical history of symptoms in response to exercise to identify whether the condition is present (as indicated by the reference standard) in people under investigation for asthma.

One study was included in the review[33] (see Table 16). Evidence from this study is summarised in the clinical evidence profile below (Table 17). See also the study selection flow chart in Appendix D, sensitivity / specificity forest plots in Appendix J, study evidence tables in Appendix G and exclusion list in Appendix K. The population consisted of adults only. No studies were identified in children or young people.

The included study[33] was a cross-sectional study, and looked at the diagnostic accuracy of a clinical history of symptoms after exercise in patients with signs and symptoms of asthma. The reference standard was physician's diagnosis of asthma with an objective test.

Summary of included studies

Table 16: Summary of studies included in the review

Study	Population	Index test & cut-off	Reference standard	Comparator test
Index test vs Reference Standard				
CHOI 2007[33]	N = 302 Adults Respiratory symptoms such as dyspnoea, cough or wheezing	Symptoms after exercise Questionnaire consisting of 11 questions regarding symptoms. Affirmative answer to Q3 = Have you had wheezing associated with dyspnoea (provoking factor – exercise)?	Physician Dx with objective test Patients with an FEV1 >70% had MCT, all other patients had BDR. Definite Dx of asthma made using test (MCh PC20 <16mg/ml or BDR FEV1 increase >12% and 200ml)	None

Table 17: Clinical evidence profile: Symptoms in response to exercise vs Reference Standard (physician Dx and objective test)

Index Test (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity %	Specificity %	Area Under Curve (range)	Quality
Symptoms in response to exercise - Adults										
Affirmative answer to questionnaire – exercise as a provoking factor	1	302	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[c]	n/a[d]	40.0	78.3	n/a	LOW

(a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.
(b) Inconsistency was assessed by inspection of the sensitivity / specificity values in each of the studies in relation to a high (≥75%), moderate (50-74%) or low (<50%) effect estimate. Inconsistency was considered high if the point estimates from individual studies were within two areas, and very high if they were in 3 areas or both high and low areas.
(c) The cut-off threshold for the reference standard objective test was not performed at the optimal cut-off for objective tests as determined by this guideline
(d) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed.

8.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow diagram in Appendix E.

8.5 Evidence statements

Clinical

- One study with 302 adults showed that a clinical history of symptoms after exercise has a sensitivity of 40.0% and a corresponding specificity of 78.3% for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)

Economic

- No relevant economic evaluations were identified.

8.6 Recommendations and link to evidence

Recommendations	9. Do not use an isolated clinical history of symptoms after exercise to diagnose asthma.
Relative values of different outcomes	The GC was interested in the diagnostic test accuracy of symptoms after exercise in the diagnosis of asthma. There is evidence of significant misdiagnosis of asthma, with both false positive and false negative errors (see introduction to the guideline). The GC was interested in both the sensitivity and specificity of diagnostic tests and prioritised studies which reported these outcomes. The outcome of interest is the diagnostic test accuracy of taking a clinical history of symptoms after exercise. Exercise is known to exacerbate asthma in a proportion of patients and it has been suggested that taking a history of symptoms after exercise in a patient with suspected asthma may aid diagnosis. A number of studies were excluded as they reported the diagnostic accuracy of taking a history of symptoms after exercise in the general population as a screening tool for asthma. The GC was only interested in studies where the population consisted of patients presenting with respiratory signs and symptoms suggesting possible asthma.
Trade-off between clinical benefits and harms	There is no direct harm associated with taking a clinical history of symptoms after exercise and the trade-off between benefit and harm will depend on the reliability of the question as a means of indicating asthma. Taking a clinical history of symptoms after exercise to diagnose asthma had a low sensitivity and a high specificity. The GC discussed that symptoms of breathlessness (over and above what one would normally expect) after exercise can occur for reasons other than asthma, and a diagnosis of asthma should not be made on the basis of a positive history of symptoms after exercise in isolation. The GC also noted the low sensitivity and the high number of false negatives suggesting the question should not be used as a 'rule out' test.
Economic considerations	No relevant economic evaluations were identified.

	The clinical evidence showed that symptoms after exercise had a low diagnostic accuracy. Therefore, even though the costs of asking about this symptom are very low, using it as a stand-alone diagnostic test would lead to more harm than good because it will lead to inappropriate under-treatment, waste of resources and lack of a correct diagnosis.
Quality of evidence	• In children aged <5 years, no studies were identified. • In children aged 5-16 years, no studies were identified. • In adults >16 years, one study was identified[33] using the reference standard (physician diagnosis plus an objective test) that addressed the diagnostic test accuracy of a history of symptoms after exercise in the diagnosis of asthma. The included study used a symptom questionnaire, and in this case an affirmative answer to the question 'Have you had wheezing associated with dyspnoea with exercise as the provoking factor (within 1 year)' was taken as a positive index question. The evidence was downgraded for indirectness as the cut-off threshold for the reference standard objective test was not performed at the optimal cut-off for objective tests as determined by this guideline. Evidence was of low quality.
Other considerations	The GC recommendation is based on review of the evidence in adults and consensus opinion of the GC in children. The GC did not make a future research recommendation. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

9 Diagnosis: Symptoms after using medication

9.1 Introduction

Certain drugs are known to exacerbate asthma in a proportion of patients. The two classes of drugs that are commonly associated with worsening of underlying asthma are non-steroidal anti-inflammatories (NSAIDs), including aspirin and ibuprofen, and beta-blockers.

Cross-sectional studies suggest that less than 10% of people with asthma have worsening of their respiratory symptoms after ingestion of NSAIDs. The exact mechanism by which NSAIDs exacerbate asthma is uncertain, but it is believed to be related to their effect on the metabolism of inflammatory mediators, known as leukotrienes.

Beta-blockers are currently contra-indicated in asthma because of the potential to block the beta-receptors in the smooth muscle within the airways. Stimulation of the beta-receptors normally leads to dilatation of the bronchi and blockade of these receptors may worsen the bronchoconstriction that is commonly associated with asthma, as well as preventing beta2 agonists, such as salbutamol, from dilating the asthmatic airway.

Associated respiratory symptoms following ingestion of these drugs can sometimes indicate a person may have asthma, this raises the question of whether taking a clinical history of symptoms after using medication is a good diagnostic test, in the diagnosis of asthma in people presenting with symptoms.

9.2 Review question: In people under investigation for asthma, what is the diagnostic accuracy of a clinical history of symptoms after taking the following drugs:

a) in adults - beta blockers, aspirin, or other NSAIDs

b) in children – ibuprofen?

For full details see review protocol in Appendix C.

Table 18: PICO characteristics of review question

Population	People with suspected asthma (presenting with respiratory symptoms). All ages, stratified into the following 3 different groups: • Children (1-<5 years old) - for ibuprofen only • Children/young people (5-16 years old) – for ibuprofen only • Adults (>16 years old) – for beta blockers, aspirin or other NSAIDs
Index test	• Clinical history of symptoms after taking drugs. NOTE: drugs of interest for the adult population are aspirin and NSAIDs, beta blockers. For children – ibuprofen. NOTE: symptoms would be a combination of the following, or individual symptoms – wheezing, cough, breathlessness, nocturnal symptoms, diurnal and seasonal variations.
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test
Outcomes	• Diagnostic accuracy (sensitivity, specificity)

9.3 Clinical evidence

There were no relevant clinical studies identified of clinical history of symptoms after taking aspirin, NSAIDs or beta blockers compared with the reference standard of physician diagnosis or other objective tests in adults. There were also no studies found in children for ibuprofen.

The majority of the evidence did not address the review question. Most studies compared the diagnostic accuracy of the index test with aspirin challenge tests in the diagnosis of aspirin-sensitive asthma in people with a confirmed diagnosis of asthma. No studies compared the diagnostic accuracy of the index test with the reference standard in the diagnosis of asthma in patients with suspected asthma (see excluded studies in Appendix K).

9.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow diagram in Appendix E.

9.5 Evidence statements

Clinical

- No evidence identified.

Economic

- No relevant economic evaluations were identified.

9.6 Recommendations and link to evidence

Recommendations	No clinical recommendation.
Relative values of different outcomes	The GC was interested in the diagnostic test accuracy of taking a clinical history of symptoms after using medication in adults (beta-blockers, aspirin or other NSAIDs) and children (ibuprofen only). These drugs are known to exacerbate asthma in a proportion of patients and it has been suggested that taking a history of symptoms after using medication in a patient with suspected asthma may aid diagnosis. There is evidence of significant misdiagnosis of asthma, with both false positive and false negative errors (see introduction to the guideline). The GC was interested in both the sensitivity and specificity of diagnostic tests and prioritised studies which reported these outcomes.
Trade-off between clinical benefits and harms	No clinical evidence was identified in order to assess the diagnostic test accuracy (sensitivity and specificity) of a clinical history of symptoms after taking aspirin, NSAIDs or beta blockers compared with the reference standard in the diagnosis of asthma. There were also no studies found in children for ibuprofen. The majority of the evidence assessed the sensitivity and specificity of a clinical history of symptoms after using medication in the diagnosis of aspirin-sensitive asthma in people with a confirmed diagnosis of asthma. Although the GC looked for all evidence of ibuprofen use in children, they acknowledged that children under 12 are not routinely exposed to either beta-blockers or NSAID (due to concern about Reyes Syndrome) and thus no evidence was expected in this age range.

Economic considerations	No relevant economic evaluations were identified. As there were no clinical studies identified there was no way of identifying whether symptoms after certain drugs was a cost-effective tool to aid in the diagnosis of asthma.
Quality of evidence	No clinical evidence was identified.
Other considerations	The GC acknowledged that the utility of this question is hampered by the information that patients are given by pharmacists or healthcare professionals. People with asthma are often told they should avoid NSAIDs and are allergic to them, without evidence of a reaction to drugs. This is particularly pertinent when considering treatment for osteoarthritis in adults where NSAIDs are a mainstay of treatment. The GC suggested that the lack of evidence derives from the fact that taking a clinical history of symptoms after using medication is not routinely used in the diagnosis of asthma; rather, it is used to characterise a particular asthma phenotype in order to guide management, e.g. for the avoidance of certain drugs. The GC stated that around 1 in 12 people with severe asthma have a response to drugs and further research may be beneficial to determine the diagnostic test accuracy of taking a clinical history of symptoms after using medication. Anecdotally, clinicians may find a history of respiratory symptoms in response to specific drug exposure useful in increasing the suspicion of a diagnosis of asthma; however, mandating to ask or not ask the question is not possible based on current evidence. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

10 Diagnosis: Occupational asthma

10.1 Introduction

Occupational asthma is a form of asthma attributable to a particular exposure in the workplace and not due to stimuli encountered outside the workplace. The true frequency is unknown, but there are concerns that it is under-reported. Published evidence estimates that occupational asthma may account for between 9 and 15% of adult onset asthma. Occupational asthma is the commonest industrial lung disease in the developed world with over 400 reported causes.

Occupational asthma should therefore be considered in all workers with adult onset asthma. The current BTS/SIGN guidelines for asthma state that all adults with airflow obstruction should be asked whether they are better on days away from work or on holiday. However, currently there is a lack of certainty about the diagnostic accuracy of asking about symptoms away from work.

10.2 Review question: In adults under investigation for occupational asthma, what is the diagnostic accuracy for case identification, of asking whether their symptoms are better away from work?

For full details see review protocol in Appendix C.

Table 19: PICO characteristics of review question

Population / Target condition	Adults (>16 years old) with suspected occupational asthma.
Index test	Symptoms are better away from work. NOTE: symptoms are defined as – wheezing, cough, breathlessness, nocturnal symptoms, diurnal variations
Reference standard	Physician's diagnosis of occupational asthma supported by an objective test (e.g. specific inhalation challenge)
Outcomes	• Diagnostic accuracy (sensitivity, specificity)

10.3 Clinical evidence

We searched for cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of asking whether symptoms are better away from work to identify whether occupational asthma is present (as indicated by the reference standard) in people under investigation for occupational asthma. The reference standard for occupational asthma was physician's diagnosis supported by an objective test (e.g. specific inhalation challenge).

Four studies were included in the review[16,103,185,186]. Evidence from these are summarised in Table 20 and the clinical evidence profile (Table 21) below. See also the study selection flow chart in Appendix D, forest plots in Appendix J, study evidence tables in Appendix G and exclusion list in Appendix K.

All the studies were cross-sectional diagnostic studies in adults[16,103,185,186] and included symptomatic samples (i.e. referred for possible occupational asthma). The reference standard was a physician's diagnosis of occupational asthma supported by an objective test (e.g. specific inhalation challenge). The GC was interested in the diagnostic accuracy of asking this question for all causal agents

combined. Due to heterogeneity in the results, they also considered the accuracy of asking this question for individual causal agents.

Summary of included studies

Table 20: Summary of studies included in the review

Study	N	Index test/reference standard	Index test/reference standard cut-off for positivity	Population	Age
Asking whether symptoms are better away from work vs. reference standard (physician diagnosis)					
BAUR 1998[16]	62 healthcare workers (airborne latex; 12 asthma) 28 bakers (flour, baking enzymes; 7 asthma) 114 isocyanate workers (isocyanates; 21 asthma)	ATS-DLD questionnaire and experienced physician medical and occupational case history Bronchial challenge with occupational agents	Reversible airways narrowing (SOB, wheeze) causally related to exposure in the working environment occurred repeatedly Specific conductance (sG_{aw}) dropped ≥40% from baseline and absolute value ≤0.5$(kPa*s)^{-1}$	Healthcare workers with contact with latex gloves, bakers or isocyanate workers presenting with suspected occupational asthma (excluded if challenge tests contra-indicated or declined)	Healthcare workers 31 (8.1); bakers 32 (11.9); isocyanate workers 39 (11.1) years
Malo 1991[103]	162 (75 occupational asthma)	Questionnaire/ chest physician (SOB, cough, wheezing or chest tightness present and timings) Final diagnosis including specific inhalation challenges, serial monitoring of peak flow at work and away from work or both	Whether symptoms worse during or after work and improved during weekends and holidays Fall in FEV1 > 20% (or ≥15% in late component of dual reactions) on specific challenge or patterns	Consecutive cases referred for possible occupational asthma	Mean 39.6 (11.8) years

Study	N	Index test/reference standard	Index test/reference standard cut-off for positivity	Population	Age
			suggestive of work-related asthma using graphs of individual, mean, maximum and minimum daily values using Burge criteria		
Vandenplas 2001[185]	45 (31 occupational asthma)	Questionnaire/ physician Clinical diagnosis including objective test: SICs with natural rubber latex gloves;	Symptoms present only on work days SICs with natural rubber latex gloves; FEV1 fell by more than 20%	Consecutive patients referred for investigation of possible OA caused by latex; exposed at work to airborne natural rubber latex (NRL) allergens from NRL gloves.	Mean 33.6 years
Vandenplas 2005[185]	212 (72 occupational asthma)	Questionnaire Specific inhalation challenge	a) Improve-ment or dis-appearance of symptoms at weekends; b) Improve-ment or dis-appearance of symptoms during vacations A sustained fall in FEV_1 of 20%	Prospectively assessed in outpatient clinics of four hospital centres and who underwent objective testing with specific inhalation challenges	38.8 (10.7) years

OA = occupational asthma

Table 21: Clinical evidence profile: Question of whether symptoms better away from work vs. Physician Dx of occupational asthma with objective test

Question whether symptoms better away from work (Yes/No)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Median Sensitivity % (range)	Median Specificity % (range)[e]	Area Under Curve (range)	Quality
Question whether symptoms better away from work (all causal agents)	4	623	Serious risk of bias[b]	Very serious inconsistency[a]	No serious indirectness	N/A[c]	range 0.48 to 1.0	range 0.32 to 0.71	-	VERY LOW
Improvement or disappearance of symptoms during weekend (many causal agents)	1	212	No risk of bias	No serious inconsistency	No serious indirectness	N/A[c]	0.76	0.54	-	HIGH
Improvement or disappearance of symptoms during vacations (many causal agents)	1	212	No risk of bias	No serious inconsistency	No serious indirectness	N/A[c]	0.74	0.57	-	HIGH
Symptoms better away from work (flour).	1	28	Serious risk of bias[b]	No serious inconsistency	No serious indirectness	N/A[c]	1.00	0.62	-	MODERATE
Symptoms better away from work (isocyanate).	1	114	Serious risk of bias[b]	No serious inconsistency	No serious indirectness	N/A[c]	0.67	0.66	-	MODERATE
Symptoms better away from work (latex).	2	107	Serious risk of bias[b]	Very serious inconsistency[a]	No serious indirectness	N/A[c]	Range 0.48 to 0.92	Range 0.32 to 0.71	-	VERY LOW
Symptoms better away from work (many causal agents).	2	374	No risk of bias	Serious inconsistency[a]	No serious indirectness	N/A[c]	Range 0.74 to 0.87	Range 0.55 to 0.57	-	MODERATE

a) *Inconsistency was assessed by inspection of the sensitivity / specificity values in each of the studies in relation to a high (≥75%), moderate (50-74%) or low (<50%) effect estimate. Inconsistency was considered high if the point estimates from individual studies were within two areas, and very high if they were in 3 areas or both high and low areas.*
b) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.
c) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed.

10.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow diagram in Appendix E.

10.5 Evidence statements

Clinical

- Four studies with 623 adults showed that question of whether symptoms are better away from work (question: whether symptoms better away from work, all causal agents) has a sensitivity range of 0.48 to 1.0 and a corresponding specificity range of 0.32 to 0.71 for diagnosing occupational asthma in people presenting signs and symptoms of possible occupational asthma. (VERY LOW QUALITY)
- One study with 212 adults showed that question of whether symptoms are better away from work (question: improvement or disappearance of symptoms during the weekend, many casual agents) has a sensitivity of 0.76 and a corresponding specificity of 0.54 for diagnosing occupational asthma in people presenting signs and symptoms of possible occupational asthma. (HIGH QUALITY)
- One study with 212 adults showed that question of whether symptoms are better away from work (question: improvement or disappearance of symptoms during vacations, many casual agents) has a sensitivity of 0.74 and a corresponding specificity of 0.57 for diagnosing occupational asthma in people presenting signs and symptoms of possible occupational asthma. (HIGH QUALITY)
- One study with 28 adults showed that question of whether symptoms are better away from work (question: symptoms better away from work, causal agent flour) has a sensitivity of 1.00 and a corresponding specificity of 0.62 for diagnosing occupational asthma in people presenting signs and symptoms of possible occupational asthma. (MODERATE QUALITY)
- One study with 114 adults showed that question of whether symptoms are better away from work (question: symptoms better away from work, causal agent isocyanate) has a sensitivity of 0.67 and a corresponding specificity of 0.66 for diagnosing occupational asthma in people presenting signs and symptoms of possible occupational asthma. (MODERATE QUALITY)
- Two studies with 107 adults showed that question of whether symptoms are better away from work (question: symptoms better away from work, latex) has a sensitivity range of 0.48 to 0.92 and a corresponding specificity range of 0.32 to 0.71 for diagnosing occupational asthma in people presenting signs and symptoms of possible occupational asthma. (VERY LOW QUALITY)
- Two studies with 374 adults showed that question of whether symptoms are better away from work (question: symptoms better away from work, many casual agents) has a sensitivity range of 0.74 to 0.87 and a corresponding specificity range of 0.55 to 0.57for diagnosing occupational asthma in people presenting signs and symptoms of possible occupational asthma. (MODERATE QUALITY)

Economic

- No relevant economic evaluations were identified.

10.6 Recommendations and link to evidence

Recommendations	**10. Check for possible occupational asthma by asking employed people with suspected new-onset asthma, or established asthma that is poorly controlled:** • **Are symptoms better on days away from work?** • **Are symptoms better when on holiday[e]?** **Make sure all answers are recorded for later review.** **11. Refer people with suspected occupational asthma to an occupational asthma specialist.**
Relative values of different outcomes	The GC was interested in the sensitivity and specificity of two simple questions which have been previously put forward as a good means of indicating people whose asthma might be of occupational aetiology. The diagnosis of occupational asthma is of considerable significance for an individual both for health and economic reasons, and would need expert confirmation with further, occasionally complex, assessment (beyond the scope of this guideline). The GC was aware that the questions may pick up people whose asthma is worse at work because of non-specific irritants to which they might be exposed in the workplace, but this circumstance should be distinguishable once the patient has undergone further assessment. They were therefore particularly interested in the sensitivity of the questions (rather than the specificity) since it would be important not to miss potential cases.

Trade off between clinical benefits and harms	• There is no direct harm associated with asking these questions. The trade-off between benefit and harm will depend on their reliability as a means of indicating possible cases of occupational asthma.
	• **Are symptoms better away from work?:** Overall, the sensitivity was 'high' (less false negatives) and the specificity was 'moderate' (more false positives). This pattern was true for all the specific causative agents considered, with the exception of isocyanates and latex. For latex there were two studies (Baur 1998 and Vandenplas 2001): one study showed high specificity but low sensitivity, while this pattern was reversed in the other study. There was no obvious reason for this difference in comparison to other agents; the GC speculated that latex exposure, unlike agents such as isocyanates, might occur outside the workplace. The numbers studied for each specific agent were small.
	• The trend was similar for the other questions (high sensitivity and a moderate specificity):
	  ○ Symptoms better at weekend
	  ○ Symptoms better during vacation.

Economic considerations	No relevant economic evaluations were identified. The two questions to detect occupational asthma are asked by the healthcare professional and in itself this has a negligible cost. However, this question may trigger further investigations which could be costly (for example, other tests or referral to specialist care). As occupational asthma is the only potentially curable type of asthma, there are considerable health benefits associated with the correct diagnosis of this type of asthma: people are able to eliminate the source of asthma, may return to a normal quality of life, and unnecessary treatment is averted. For this reason, the GC considered a high sensitivity (i.e. minimisation of false negatives) more important than a high specificity (minimisation of false positives). The clinical evidence has shown that asking these questions helps identify people with occupational asthma, i.e. this question has a high sensitivity, and therefore it is

[e] 'Holiday' here means any longer time away from work than usual breaks at weekends or between shifts.

	considered to be cost-effective.
Quality of evidence	The evidence ranged from very low to high quality. Only four studies were identified, and the sample sizes and number of people with a final diagnosis of occupational asthma were generally small. The heterogeneity for the sensitivity and specificity for the question of whether symptoms better away from work considering all causal agents was very serious. The GC also considered the evidence for different causal agents separately.
Other considerations	The GC concluded that these simple questions were of value in raising the possibility of occupational asthma. They debated when and how frequently they should be asked. There was agreement that they should be used in all working adults with a new diagnosis of asthma. In established asthma there was a view among some of the GC that the questions should be part of an annual asthma review, but their use in this fashion has not been tested, and it was considered possible that repeated use of the questions was more likely to result in falsely positive answers and consequent unnecessary detailed investigation. The GC agreed by consensus that it is appropriate to repeat the questions to adults with asthma when asthma control is found to be poor, whether this be at annual review or on other occasion; however, it is not necessary as a routine review question when control is good. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

11 Diagnosis: Spirometry

11.1 Introduction

Asthma is characterised by variation in airflow obstruction over time, greater than that seen in healthy populations. In asthma, lung function may vary between completely normal and severely obstructed in the same patient. Poorly controlled asthma is associated with greater variability in lung function than well-controlled asthma.

Spirometry is a physiological test that measures how an individual inhales or exhales volumes of air as a function of time. The primary signal measured in spirometry may be volume or flow.

The most important aspects of spirometry are the forced vital capacity (FVC), which is the volume delivered during expiration made as forcefully and completely as possible starting from full inspiration, and the forced expiratory volume (FEV) in one second, which is the volume delivered in the first second of an FVC manoeuvre. A reduced ratio of FEV1 to FVC indicates airflow obstruction.

Tests of pulmonary function should be carried out by appropriately trained staff with appropriate equipment who are able to assess the correct performance of the test by the patient and the quality of the results.

Further explanation on lung function measurement and interpretation is given in the European Respiratory Society and American Thoracic Society (ERS/ATS) guidelines[35] on standards of spirometry.

Spirometry is recommended for the diagnosis and management of asthma in national and international guidelines[22,59].

11.2 Review question: In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of spirometry/flow volume loop measures?

For full details see review protocol in Appendix C.

Table 22: Characteristics of review question

Population	People with suspected asthma (presenting with respiratory symptoms). Ages stratified into the following 2 groups: • Children/young people (5-16 years old) • Adults (>16 years old)
Index test	Spirometry measures (report separately) • FEV1/FVC ratio (<70%) • Flow volume loop (graph) • FEV1 (<80%) – if limited evidence from the above two measures Pre bronchodilator values (applies for all above measures) FEV1 and FVC should be performed using the following criteria: • Forced expiratory volume (FEV1) - patients perform manoeuvre until 3 readings are within 5% of each other (maximum 8 attempts) the measured value being the best of these 3 readings. • Forced vital capacity (FVC) - patients perform manoeuvre until 3 readings are within 5% of each other (maximum 8 attempts) the measured value being the best of these

	3 readings.
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test.
Statistical measures	Diagnostic accuracy (sensitivity and specificity).

11.3 Clinical evidence

We searched for cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of spirometry to identify whether the condition is present (as indicated by the reference standard) in people under investigation for asthma.

Six studies were included in the review[55,135,137,153,163,165] (see Table 23 and Table 24). Evidence from these studies is summarised in the summary tables and the clinical evidence profile below (Table 25). See also the study selection flow chart in Appendix D, sensitivity / specificity forest plots in Appendix J, study evidence tables in Appendix G and exclusion list in Appendix K.

All six studies were cross-sectional studies[55,135,137,153,163,165], and looked at the diagnostic accuracy of spirometry in patients with signs and symptoms. Evidence was available from one study for the ideal index test measure of FEV1/FVC ratio <70%. No evidence was available for the index test of flow volume loop measures. Due to limited evidence for FEV1/FVC ratio, evidence was included for studies reporting FEV1/FVC ratio <70% and/or FEV1 <80% (2 studies) and for studies reporting FEV1 <80% only (3 studies). The reference standard was physician's diagnosis of asthma with an objective test. A variety of objective tests and thresholds were used for the reference standard objective test (see Table 13). In children and young people, evidence was only available for the index test of FEV1<80% from one study[163]. In one study[165], evidence was available from adults, and children and young people combined (age range 8-75 years). This evidence was included in the review with the data in adults due to the mean age, and was downgraded as the combined age range is indirect to the protocol.

Summary of included studies

Table 23: Summary of studies included in the review: ADULTS Spirometry

Study	Population	Index test & cut-off	Reference standard	Comparator test
Index test vs Reference Standard				
FORTUNA 2007[55]	N=50 Adults Referred with a clinical history suggestive of asthma (dry cough, wheezing, and shortness of breath) No mention of other respiratory defects. No BMI reported.	Spirometry FEV1<80%	Methacholine challenge test (PD20 ≤16mg/ml) following guidelines of the GINA	None
PINO 1996[135]	N=84 Adults Clinically suspected of bronchial asthma No mention of other respiratory defects. No BMI reported.	Spirometry FEV1/FVC<70% and FEV1<80%	If obstructive spirometry: performed BDR (400µg salbutamol; FEV1 >15% initial) If normal spirometry: methacholine challenge test five breaths of 5mg/ml and five breaths of 25mg/ml, test positive if a 20% drop in FEV1	None
POPOVIC 2002[137]	N=195 Adults Referred by GP with suspected asthma and symptoms of breathlessness / dyspnoea. No mention of other respiratory defects. No BMI reported.	Spirometry FEV1 <80% predicted	Dx made on the basis of questionnaire, with typical medical history data of occasional asthma attacks with wheezing and nocturnal awakening because of dyspnoea, and reversible bronchial obstruction after salbutamol test (no further details stated)	None
SCHNEIDER 2009A[153]	N=219 Adults Visiting GP for the first time with complaints of suggested obstructive airway disease (OAD). Symptoms such as dyspnoea, coughing, or expectoration No mention of other respiratory defects. BMI	Spirometry at GP. FEV1/VC ≤70% and/or FEV1 <80%	Dx by respiratory physician based on whole-body plethysmography (FEV1/VC ≤70% or FEV1 <80%) followed by either BDR if obstruction is present (FEV1 ≥12% and ≥200ml) or methacholine if obstruction is not present (PC20 ≤16mg/ml or extreme increase in airway resistance accompanied by clinical symptoms in two patients)	None

Study	Population	Index test & cut-off	Reference standard	Comparator test
	reported, mean SD 25.3 (4.4)			
SMITH [165]2004	N=47 Adults and Children 8-75 years Referred to hospital pulmonary function lab by GP for possible asthma. Respiratory symptoms for a minimum of 6 weeks No mention of other respiratory defects. No BMI reported.	1. FEV1 <80% pred 2. FEV1/FVC <70%	Relevant symptom history (all patients) and a positive hypertonic saline challenge test (PD15<20ml) or BDR increase in FEV1 ≥12%	None

Table 24: Summary of studies included in the review: CHILDREN Spirometry

Study	Population	Index test & cut-off	Reference standard	Comparator test
Index test vs Reference Standard				
SIVAN 2009[163]	N=133 Children Non-specific respiratory symptoms suggestive of asthma for at least 3 months, including cough, wheezing and shortness of breath with or without trials of treatment with bronchodilators and ICS. No mention of other respiratory defects. No BMI reported.	Spirometry FEV1 <80%	Made by paediatric pulmonologist after 18 months follow-up. Based on history of 2 or more clinical exacerbations of wheezing documented by a physician; dyspnoea or cough relived by bronchodilators; documented variability in FEV1 ≥15% in response to bronchodilators at any time during the follow-up period; OR documented variability in FEV1 ≥15% over time with or without controller medications (ICS or montelukast). Results of provocation tests included when available.	None

Table 25: Clinical evidence profile: Index test vs Reference Standard (physician Dx and objective test)

Index Test (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity %	Specificity %	Area Under Curve (range)	Quality

Index Test (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity %	Specificity %	Area Under Curve (range)	Quality
Spirometry ADULTS										
FEV1/FVC <70%	1	47	No risk of bias[a]	No serious inconsistency[b]	Serious indirectness[c]	n/a[e]	35.3	100	-	MODERATE
Flow volume loop	0									
FEV1/FVC <70% and/or FEV1<80%	2	303	Serious risk of bias[a]	Serious inconsistency[b	Serious indirectness[d]	n/a[e]	Range 29 - 47	Range 41 - 59	-	VERY LOW
FEV1 <80%	3	292	Serious risk of bias[a]	Very serious inconsistency[b]	Serious indirectness[c]	n/a[e]	Median 29.4 (range 23-45)	Corresponding 100 (range 31-100)	-	VERY LOW
Spirometry CHILDREN										
FEV1/FVC <70%	0									
Flow volume loop	0									
FEV1 <80%	1	133	Very serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness[c]	n/a[e]	52.0	72.0	-	LOW

(a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.
(b) Inconsistency was assessed by inspection of the sensitivity / specificity values in each of the studies in relation to a high (≥75%), moderate (50-74%) or low (<50%) effect estimate. Inconsistency was considered high if the point estimates from individual studies were within two areas, and very high if they were in 3 areas or both high and low areas.
(c) Mixed population of adults and children/young people in one study. Reference standard was saline challenge test in one study
(d) Indirectness in the reference standard objective test cut-off
(e) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed.

11.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow diagram in Appendix E.

New cost-effectiveness analysis

An original health economic model was built for adults to assess the cost-effectiveness of several diagnostic pathways, which included spirometry. Full details of the model can be found in Appendix M. A summary of the model can be found in section 18.4.

Unit cost of performing spirometry on children

As an economic model was not feasible for children, the GC considered the unit cost of performing a spirometry test to evaluate its cost-effectiveness as part of a pathway for diagnosing asthma. This unit cost is presented below in Table 26.

Table 26: Cost of spirometry

Item	Quantity[(c)]	Unit cost	Total Cost (quantity*unit cost)	Source of unit cost
Time of GP practice nurse to conduct the test[(a)]	10-15 minutes	£0.73 per minute	£7.30 - £10.95	PSSRU[40]
Micro-lab spirometer [(b)]	1/1500	£1498.90 per spirometer	£1.00	NHS supply catalogue[47]
Bacterial filter, 3-litre syringe for calibration[(b)]	1/1500	£295.77 per syringe	£0.20	NHS supply catalogue[47]
Bacterial filter	1	£0.99 per filter	£0.99	NHS supply catalogue[47]
Total			**£9.49 - £13.14**	

(a) This range reflects the differing levels of experience of the nurse conducting the test but also the age of the child. The test is likely to be conducted quicker in older children.
(b) To calculate the marginal cost it was assumed that the equipment lasts for 5 years and is used on average 1500 times in this period.
(c) Based on GC opinion.

The GC also acknowledged the annual cost of drugs for the management of asthma in children. Preventing these costs from occurring in children without asthma would be a large benefit derived from a diagnostic strategy with a high specificity. This cost was estimated to be £201 from a study by Main et al.[102,102]

11.5 Evidence statements

Clinical

Adults

- One study with 47 adults showed that spirometry (FEV1/FVC <70%) has a sensitivity of 35.3% and a corresponding specificity of 100% for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY).
- No evidence was available for flow volume loop
- Two studies with 303 adults showed that spirometry (FEV1/FVC <70% and/or FEV1<80%) has a sensitivity range of 29-47% and a corresponding specificity range of 41-59% for diagnosing asthma in people presenting with respiratory signs and symptoms. (VERY LOW QUALITY)
- Three studies with 292 adults showed that spirometry (FEV1<80%) has a median sensitivity of 29.4% and a corresponding specificity of 100% for diagnosing asthma in people presenting with respiratory signs and symptoms. (VERY LOW QUALITY)

Children

- No evidence was available for FEV1/FVC <70% in children
- No evidence was available for flow volume loop
- One study with 133 children showed that spirometry (FEV1 <80%) has a sensitivity of 52% and a corresponding specificity of 72% for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY).

Economic

- No relevant economic evaluations were identified.
- An original health economic model found that spirometry (together with bronchodilator reversibility, FeNO, peak expiratory flow variability and histamine or methacholine challenge test) was part of the most cost-effective diagnostic pathway used to diagnose asthma in adults aged 16 and over (see diagnostic algorithm in section 4.1). This evidence is directly applicable with minor limitations.

11.6 Recommendations and link to evidence

Recommendations	**12. Offer spirometry to adults, young people and children aged 5 and over if a diagnosis of asthma is being considered. Regard a forced expiratory volume in 1 second/forced vital capacity (FEV_1/FVC) ratio of less than 70% (or below the lower limit of normal if this value is available) as a positive test for obstructive airway disease (obstructive spirometry).**
Relative values of different outcomes	The GC was interested in the utility of spirometry in the diagnosis of asthma in patients >5 years presenting with signs and symptoms. There is evidence of significant misdiagnosis of asthma, with both false positive and false negative errors (see introduction to the guideline). The GC was therefore interested in both the sensitivity and specificity of diagnostic tests and prioritised studies which reported each of these outcomes for spirometry. Sensitivity and specificity values will depend on the chosen cut-off point. The GC included studies with a FEV1/FVC ratio cut-off value of less than 70%. Cut-off values higher than this were excluded as they are likely to occur in some healthy individuals and the GC was concerned about a high number of false positives at higher cut-off values. Due to the limited evidence identified using the FEV1/FVC ratio as the spirometric measure, the GC also considered evidence from studies reporting the FEV1 alone. As the FEV1 can be influenced by conditions which cause restrictive respiratory defects, for example, obesity, details on any exclusion criteria for other conditions or BMI were extracted from the studies into the evidence tables and considered by the GC.

Trade-off between clinical benefits and harms	The studies included in the analysis demonstrated that a FEV1/FVC ratio <70% has a low sensitivity and a high specificity. A FEV1/FVC ratio <70% and/or a FEV1 <80% had a low sensitivity and a low to moderate specificity. In adults, FEV1 alone had a low sensitivity and a high specificity and, in children, FEV1 alone had a moderate sensitivity and specificity. The GC agreed spirometry should not be used in isolation for the diagnosis of asthma due to the low sensitivity of the test, and due to the fact that obstruction also occurs in other conditions such as COPD which have symptoms in common with asthma. When considering the possible placement of spirometry in a diagnostic pathway, the GC noted the importance of performing spirometry as one of the earlier investigations in all patients, to detect the presence or absence of obstruction, which then determines which other tests are appropriate (for example, the recommendation to use BDR only if obstruction is present). It was also noted that spirometry might offer an alternative explanation for a person's symptoms even if not suggestive of asthma (if restrictive spirometry is found). The contraindications for spirometry should be considered (e.g., recent MI, recent eye surgery, etc.) when testing lung function. The FEV1 alone should not be used as it will be influenced by other conditions which cause restrictive respiratory defects, for example, obesity. The FEV1/FVC ratio should be used to detect obstructive lung disease (see discussion below under 'other considerations'). The GC recognised that it is technically difficult for some young children to breathe out for long enough to achieve an accurate FVC, making it difficult to obtain the FEV1/FVC ratio cut off. However, the ability to perform spirometry will vary from child to child and the GC agreed that spirometry should be attempted in children aged 5 years and older. In young children the GC suggest using devices that show the visual trace.
Economic considerations	No economic evaluations were found which assessed the use of spirometry as part of a diagnostic pathway. Therefore, an original health economic model was built to assess the cost-effectiveness of different diagnostic pathways for asthma. The model assessed the additional costs of tests against cost-savings from unnecessary asthma medication, and the increased health outcomes from providing correct treatment. The GC agreed that spirometry should be used as one of the first diagnostic tests in all patients. The reason is that it is a widely available test that can also help with the diagnosis of other conditions such as COPD. It is also used to determine whether the patient can undergo a BDR test. However the strategy with the fewest diagnostic tests that included spirometry produced lower health outcomes at a higher cost than other strategies indicating that the cost-effectiveness of spirometry is contingent on its use as part of a pathway, and therefore if used in practice it should be followed up by the appropriate recommended tests. The most cost-effective strategy involved using spirometry, bronchodilator reversibility (BDR), FeNO, peak expiratory flow variability (PEFv) and a methacholine challenge test (MCT) in selected individuals to diagnose asthma. In this strategy everybody with symptoms of asthma would undergo a spirometry test and a FeNO test, those who had an obstructive spirometry would also receive a BDR test. Only those who had non-obstructive spirometry and conflicting FeNO and PEFv test results would receive a MCT. Adopting this strategy dominated all other strategies apart from those which performed challenge tests at more points in the pathway. The ICERs of adopting these further stratgies were above £20,000 per QALY.

	For children the GC considered the unit cost of performing a spirometry test. The GC agreed that the cost was low relative to information gained which was crucial in determining a diagnosis. Performing spirometry is necessary to determine whether it is appropriate to perform a bronchodilator reversibility test. In children the GC noted that a positive result from this test is strong evidence that the child has asthma. As with adults the GC noted that the use of spirometry on its own could lead to worse outcomes as the clinical evidence showed spirometry to have a moderate sensitivity. Therefore the cost-effectiveness of spirometry is contingent on its use as part of the recommended diagnostic pathway (see section 21).
Quality of evidence	The quality of studies ranged from very low to moderate by GRADE criteria. The main limitation of the included studies was the reference standard. Due to the limited evidence, studies were considered where the reference standard cut-off threshold used to identify a positive diagnosis did not match the one considered by the GC as the optimal cut-off threshold. The GC also considered evidence from one study[165] where the objective test was a hypertonic saline challenge test, a test not commonly used in clinical practice. • In children aged <5 years, we did not search for studies as spirometry is not able to be performed in this age-group. • In children aged 5-16 years, one study was included[163]; however, this study only reported the diagnostic accuracy of FEV1 alone and not the ratio. The evidence was of low quality. • In adults, there were five included studies; however, only one study reported the diagnostic accuracy of the FEV1/FVC ratio alone. The evidence was of moderate quality. The economic evidence was assessed as directly applicable with minor limitations.
Other considerations	Spirometry is only useful if a good quality spirogram is obtained that is both accurate and reproducible, which will require training of personnel performing the test. The GC discussed the need for patients to have access to high quality spirometry tests, but how this is achieved may depend on local factors. An accreditation scheme has recently been introduced and advocated by NHS England, with the purpose of ensuring that every patient seen in primary care has access to quality-assured spirometry testing. The GC recognised that using the FEV1/FVC ratio with a fixed cut-off valueof 70% for all ages could lead to inaccuracy at the extremes of age. Children have a higher FEV1/FVC ratio and this falls with age so that 70% is within normal in the elderly population. Using 70% in all cases may therefore lead to an underestimation of airflow obstruction in children, and over-estimation in the elderly. The use of the 'lower limit of normal' (LLN) may be more appropriate and the GC agreed that the LLN should be used, particularly in children aged 5-16 years when the calculation is available. However, calculation of the LLN may not always be readily available, and not all healthcare professionals are familiar with this. The GC had to compromise between absolute scientific accuracy and useability in clinical practice. Although there were concerns that a threshold of 70% in children may lead to some inaccuracy in estimation of airflow obstruction the GC noted that, for the diagnosis of asthma, spirometry is only the first step in an algorithm and asthma would not be discounted on the basis of a normal spirometry alone. Small differences in the threshold used to define airways obstruction may affect people with borderline spirometry results, but these people would always have further objective tests; and those with measurements markedly different from the 70% value will be correctly classified in any case. The recommendation was therefore formulated around the FEV1/FVC ratio of 70% as the cut-off for normality, but indicates that LLN should be used when available.

Details of the feasibility project are available in appendix Q. The aim of the feasibility project was to assess the impact and feasibility of adopting the diagnostic tests (e.g. FeNO and spirometry) recommended, into primary care.

Overall the GC concluded that the results from the feasibility project were positive. Five of the 7 practices said they would continue to use the algorithms if the guideline was published with no further amendments. However, the GC acknowledged the practical limitations that sites experienced in implementing the algorithm and agreed that additional changes could help to alleviate these.

General concerns

Firstly it was recognised that children may not be able to perform some of the tests. . An additional recommendation has now been made which informs the clinician what to do should this problem arise. The GC agreed that objective testing was imperative before a diagnosis of asthma could be made. However, until objective testing is possible, symptoms should be treated and monitored.

Secondly the feasibility report identified that in the current format the diagnostic algorithms could be difficult to follow in some places. Therefore the GC agreed a new format should be designed that would simplify the algorithm and make them easier to interpret.

Concerns specific to spirometry

The project highlighted concerns about conducting spirometry accurately. The GC considered that this issue was not specific to asthma as spirometry is key in the diagnostic assessment of other common respiratory conditions such as chronic obstructive pulmonary disorder (COPD). Moving forward, the GC considered that the use of diagnostic hubs could help alleviate this issue, as only a single practice, hospital or community trust in a given locality would need to be trained in providing such a service. The diagnostic hub would then provide accurate and timely spirometry for the local area. It was agreed this could reduce costs and improve the practicality of implementing the algorithm. A recommendation was developed, aimed at clinical commissioners, to consider establishing asthma diagnostic hubs to achieve economies of scale in implementing the diagnostic algorithms (see section 21.2).

Concerns were also expressed about the time needed to perform spirometry in all those with potential asthma - the test has previously been generally used in primary care for COPD diagnosis but not asthma.

The feasibility project confirmed that spirometry was not a good stand alone test for asthma as it commonly came out non-obstructive, although this was contingent on the individual's symptoms at the time of visit. Satisfactory spirometry measurements were obtained in 90.5% of the people assessed in the feasibility study, and of those 17.7% showed airflow obstruction. Therefore the GC agreed it was important to maintain the prominent role of spirometry in the diagnostic algorithm.

Finally the project highlighted that some individuals found it difficult to conduct spirometry and after this test was completed they felt out of breath and unable to conduct a FeNO test. Therefore, when a patient is undergoing both FeNO and spirometry tests, the clinican should conduct FeNO first as this test is easy for the patient to perform and would not impact on their ability to perform, or the quality of, the spirometry test.

	Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

12 Diagnosis: Bronchodilator reversibility

12.1 Introduction

The aim of this review was to assess the clinical and cost-effectiveness of bronchial reversibility testing in those with obstructive spirometry in the diagnosis of asthma in adults and children. A determination of airflow-limitation reversibility with drug administration is commonly undertaken as part of lung function testing in those in whom obstruction is observed. There is no clear consensus about what constitutes reversibility in subjects with airflow obstruction, although the ATS/ERS Task Force: Standardisation of lung function testing currently provides the clearest guidance and is most widely used.

The first step in interpreting any bronchodilator test is to determine if any change greater than random variation has occurred[126]. The patient should undergo baseline function testing when not taking any drugs prior to the test. Short-acting inhaled drugs (e.g. the β2-agonist albuterol/salbutamol or the anticholinergic agent ipratropium bromide) should not be used within four hours of testing. Long-acting beta-agonist bronchodilators (e.g. salmeterol or formoterol) and oral therapy with aminophylline or slow-release β-agonists should be stopped for at least 12 hours prior to the test, and for 24 hours for ultra-long acting agents with a long half-life (e.g. indecaterol, vilanterol). Smoking should be avoided for >1 hour prior to testing and throughout the duration of the test procedure.

The ATS/ERS Task Force-recommended[109] procedure for assessing bronchodilator response are:

- Assess lung function at baseline. If obstruction is present (FEV1/FVC ratio <70%): Administer four separate doses of 100mcg salbutamol through a spacer and re-assess lung function after 15 minutes.
- An increase in FEV1 ≥ 12% and ≥ 200ml above baseline FEV1 after short-acting β2 agonist constitutes a positive bronchodilator response.
- The lack of a spirometric bronchodilator response in the laboratory does not preclude a clinical response to bronchodilator therapy.

12.2 Review question: In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of bronchodilator response (using PEF or FEV1)?

For full details see review protocol in Appendix C.

Table 27: PICO characteristics of review question

Component	Description
Review question	In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of bronchodilator response (using PEF or FEV1)?
Objectives	To evaluate the diagnostic test value of bronchodilator response (using PEF or FEV1) in diagnosing asthma
Population / Target condition	People with suspected asthma (presenting with respiratory symptoms). Ages stratified into the following 2 different groups: • Children/ young people (5-16 years old) • Adults (>16 years old)
Index test	Bronchodilator response, measured using the following: • PEF

	• FEV1 ○ change in FEV1 % initial and change in FEV1 litres Exclusions: • Change in FEV1 % initial alone • Change in FEV1 absolute litres alone • Change in FEV1 % predicted (ΔFEV1 %pred) • Standardised residual (SR)-FEV1 • Change in FEV1 % of possible maximal response (ΔFEV1 %max)
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test
Outcomes	• Diagnostic accuracy (sensitivity and specificity)

12.3 Clinical evidence

We searched for cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of bronchodilator response to identify whether the condition is present (as indicated by the reference standard) in people under investigation for asthma.

Four studies were included in the review[21,32,84,141] (see Table 28). Evidence from these are summarised in the clinical evidence profile below (Table Table 29). See also Appendix D, sensitivity and specificity plots in Appendix J, clinical evidence tables in Appendix G and excluded studies in Appendix K.

All studies were in adults and no evidence was identified for the 5-16 year age group.

The included population of all the included studies was people with asthma or COPD, rather than suspected asthma. These studies aimed to assess the diagnostic accuracy of BDR in distinguishing between asthma and COPD. These studies were included in the review due to limited evidence in the suspected asthma population. The reference standard in the included studies was physician diagnosis. With the exception of two studies[21,141], it was unclear if the reference standard included an objective test for asthma. In the Quadrelli 1999 study, it was unclear whether all people received the objective test. Again, due to limited evidence these studies were included but the quality of the evidence was downgraded for population and reference standard indirectness.

Summary of included studies

Table 28: Summary of studies included in the review: Bronchodilator reversibility vs. reference standard (physician diagnosis): adults

Study	N	Index test/reference standard	Index test cut-off for positivity (measures in 'bold' are those specified in the protocol, and thus used in the analyses)	Population	Age
Brand 1992[21]	150	• Response to inhaled terbutaline 1000µg a) change [Δ]FEV1 % init; b) ΔFEV1[l] i.e. absolute value in litres; c) ΔFEV1 % init and ΔFEV1[l]; d) ΔFEV1 %pred; e) standardised residual [SR]-FEV1; f) FEV1 post-bronchodilator [pb] %pred Standardised history using criteria of ATS: asthma = attacks of breathlessness and wheeze (asthma attacks) without chronic (>3 months/year) cough or sputum production; COPD = Current or former smokers without a history of asthma attacks reporting either chronic cough +/- sputum production, or dyspnoea when walking quietly on level ground, or both Plus hyper-reactivity to inhaled histamine	a) ΔFEV1 % init>15%; b) ΔFEV1[l]> 0.200; **c) ΔFEV1 % init and ΔFEV1[l]:>15% and > 0.200;** d) ΔFEV1 %pred>9%; e) SR-FEV1> 0.5; f) FEV1 pb %pred>80%	Adults with chronic respiratory symptoms (asthma or COPD) in university hospital outpatients departments; baseline FEV1 >1.2 litres and 1.64-4.5 residual standard deviations below predicted value, or FEV1/inspiratory vital capacity ratio >1.64 RSD below predicted; hyperresponsive to inhaled histamine	18-60 years
Chhabra 2005[32]	354	• Response to inhaled salbutamol 200µg: a) absolute change in FEV1 (ΔFEV1); b) ΔFEV1%init; c) ΔFEV1%pred; d) ΔFEV1≥0.2l and ΔFEV1%init ≥12% Physician diagnosis based on clinical criteria suggested by the National Institute of Health Global Strategy for Asthma Management and Prevention (asthma = recurrent episodes of breathlessness and wheezing, with or without cough and phlegm, with seasonal and diurnal variations and any identifiable trigger factors) and the Global Initiative for Chronic Obstructive Lung Disease (COPD = history of smoking >10 pack-years,	a) absolute change in FEV1 (ΔFEV1) a1: 0.2l; a2: 0.3l; a3: 0.4l; b) ΔFEV1%init b1: 12%; b2: 15%; b3: 20%; c) ΔFEV1%pred c1: 9%; c2: 15%; **d) ΔFEV1 and ΔFEV1%init: ≥0.2l and ≥12%**	Clinical diagnosis of asthma (non-smokers) or COPD; stable clinical state with no history of acute exacerbation in previous 4 weeks; acceptable performance of spirometry; FEV1/FVC ratio 70% or less. Participants were already on corticosteroid treatment.	Asthma mean 35.60 (12.47); COPD mean 56.28 (9.57) years

Study	N	Index test/reference standard	Index test cut-off for positivity (measures in 'bold' are those specified in the protocol, and thus used in the analyses)	Population	Age
		cough with expectoration for at least 3 consecutive months in a year for 2 years or more and progressive dyspnoea on exertion).			
Kim 2012[84]	514	• Bronchodilator response to salbutamol 400µg Clinical decision (no definite diagnostic criteria) by specialists in allergy or pulmonary departments	**Increase in FEV1 >200mL and >12% above baseline**	Adults with chronic obstructive airways disorders included in an asthma cohort or a COPD cohort; all had at least one chronic persistent respiratory symptom (dyspnoea, cough, sputum production or wheeze) for >3 months or repetition of the symptom for >3 months	48 (16) years for asthma and 65 (8) years for COPD
Quadrelli 1999[141]	119	• Response to inhaled salbutamol 200µg a) ΔFEV1[L]; b) ΔFEV1%init; c) ΔFEV1[L] plus ΔFEV1%init; d) ΔFEV1%pred; e) ΔFEV1%max (% of maximal possible response) Clinical diagnosis: asthma = attacks of breathlessness or wheeze according to ATS criteria (smokers excluded) and at least 2 of: 1; history of symptoms since childhood or adolescence; 2. symptomatic-free periods of >3 months; 3. spontaneous variations in FEV1 during the year of >20% of baseline value; 4. histamine challenge test <8mg/mL. COPD = heavy current or ex-smokers with no history of asthma reporting chronic cough or sputum (non-smokers excluded)	a) ΔFEV1[L]: 200mL; b) ΔFEV1%init: 15%; **c) ΔFEV1[L] and ΔFEV1%init: >200mL and >15%;** d) ΔFEV1%pred: 9%; e) ΔFEV1%max (% of maximal possible response): 50%	Patients with previously diagnosed airways obstruction; present baseline spirometry: FEV1/FVC relationship1.64 SEE below predicted value or lower; people with asthma had FEV1 <55% predicted (to match with COPD patients' baseline lung function).	Overall asthma : 55.4 (19.0) years; COPD 67.3 (7.0) years

Table 29: Clinical evidence profile: Bronchodilator reversibility vs. Physician Dx of asthma

Bronchodilator reversibility (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Median Sensitivity % (range)	Median Specificity % (range)[e]	Area Under Curve (range)	Quality
ADULTS										
ΔFEV1%init ≥12% **and** ΔFEV1[L] ≥0.2L	2	868	Serious risk of bias[a]	Serious inconsistency[b]	Serious indirectness[d,e]	N/A[c]	Range 0.17-0.65	Range 0.61-0.81	-	VERY LOW
ΔFEV1%init >15% **and** ΔFEV1[L] >0.2L	2	269	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[d,e]	N/A[c]	Range 0.69-0.69	Range 0.55-0.71	-	LOW
CHILDREN 5-16 years										
No evidence identified	0									

12.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow chart in Appendix E.

New cost-effectiveness analysis

An original health economic model was built for adults to assess the cost-effectiveness of several diagnostic pathways, which included bronchodilator reversibility. Full details of the model can be found in Appendix M. A summary of the model can be found in section 18.4.

Unit cost of performing bronchodilator reversibility on children

As an economic model was not feasible for children, the GC considered the unit cost of performing bronchodilator reversibility to evaluate its cost-effectiveness as part of a pathway for diagnosing asthma. This unit cost is presented below in Table 30.

Table 30: Cost of bronchodilator reversibility

Item	Quantity[(c)]	Unit cost	Total Cost (quantity*unit cost)	Source of unit cost
Time taken to administer bronchodilator and check for reversibility[(a)]	8-17 minutes	£0.73 per minute	£5.84-£12.41	PSSRU[40]
Volumatic spacer	1	£3.81 per spacer	£3.81	NHS supply catalogue[47]
MDI	1	£5.50 per MDI	£5.50	NHS supply catalogue[47]
Spirometry equipment to check for reversibility[(b)]	1	£2.20	£2.20	NHS supply catalogue[47]
Total			**£17.35-£23.92**	

(a) This range reflects the differing levels of experience of the nurse conducting the test but also the age of the child. The test is likely to be conducted quicker in older children.
(b) When a bronchodilator reversibility test is being performed the first spirometry reading will have already been taken.
(c) Based on GC opinion.

The GC also acknowledged the annual cost of drugs for the management of asthma in children. Preventing these costs from occurring in children without asthma would be a large benefit derived from a diagnostic strategy with a high specificity. This cost was estimated to be £201 from a study by Main et al.[102,102]

12.5 Evidence statements

Clinical

- Two studies with 868 adults showed that bronchodilator reversibility (ΔFEV1%init ≥12% **and** ΔFEV1[L] ≥0.2L) has a sensitivity range of 0.17 to 0.65 and a corresponding specificity range of 0.61 to 0.81 for diagnosing asthma in people presenting with respiratory signs and symptoms and obstructive airways disease. (VERY LOW QUALITY)
- Two studies with 269 adults showed that bronchodilator reversibility (ΔFEV1%init >15% **and** ΔFEV1[L] >0.2L) has a sensitivity range of 0.69 to 0.69 and a corresponding specificity range of 0.55 to 0.71 for diagnosing asthma in people presenting with respiratory signs and symptoms and obstructive airways disease. (LOW QUALITY)
- No evidence was identified in children aged 5-16 years

Economic

- No relevant economic evaluations were identified.
- An original health economic model found that bronchodilator reversibility (together with spirometry, FeNO, peak expiratory flow variability and histamine or methacholine challenge test) was part of the most cost-effective diagnostic pathway used to diagnose asthma in adults aged 16 and over (see diagnostic algorithm in section 4.1). This evidence is directly applicable with minor limitations.

12.6 Recommendations and link to evidence

Recommendations	**13. Offer a BDR test to adults (aged 17 and over) with obstructive spirometry (FEV_1/FVC ratio less than 70%). Regard an improvement in FEV_1 of 12% or more, together with an increase in volume of 200 ml or more, as a positive test.** **14. Consider a BDR test in children and young people (aged 5 to 16) with obstructive spirometry (FEV_1/FVC ratio less than 70%). Regard an improvement in FEV_1 of 12% or more as a positive test.**
Relative values of different outcomes	The GC was interested in the utility of BDR for the diagnosis of asthma in adults and children over the age of 5. There is evidence of significant misdiagnosis of asthma, with both false positive and false negative errors (see introduction to the guideline). The GC was therefore interested in both the sensitivity and specificity of diagnostic tests and prioritised studies which reported each of these outcomes for bronchodilator reversibility. The GC considered the combination of a change in FEV1 as percentage initial and an absolute change in volume in litres in response to a standard dose of bronchodilator to be more appropriate than either measure alone. This is in accordance with the ERS/ATS taskforce 2006 guideline. The GC did not consider percentage change in FEV1 alone because there is a risk of over-diagnosis in people with small initial lung volumes. The GC did not consider absolute change in FEV1 in litres alone because the relevance of this measure will depend on a patient's starting FEV1. A gold standard test therefore would include both the percentage change in FEV1 and the absolute change in FEV1 in mls. The BDR test should always be performed following standard spirometry procedures (if there is an obstructive spirometry). The GC discussed the diagnostic cut-offs used to identify a positive test, evidence was available for a threshold of 12% and 200ml and for a threshold of 15% and 200ml. No studies were identified using PEF to measure the extent of bronchodilator

	reversibility.
Trade-off between clinical benefits and harms	The studies included in the analysis demonstrated that BDR at a change in FEV1 threhsold of 12% and 200ml has a moderate to high specificity and generally a lower sensitivity in the diagnosis of asthma. At this threshold, BDR could therefore be used reasonably well as a rule-in test. BDR at a change in FEV1 threhsold of 15% and 200ml had both a moderate sensitivity and specificity. The GC considered a change in FEV1 of ≥12% and a change in volume of ≥200mls to be a more appropriate threshold as evidence of a positive test in response to a standard dose of bronchodilator. This is in accordance with the ERS/ATS taskforce 2006 guideline. All the studies included in the analysis were conducted in a population with obstructive spirometry. Therefore, the recommendation has been made specifically for this population. The GC could not comment on the use of BDR in the context of normal spirometry. A BDR test is considered safe, quick and non-invasive, with no significant harms to the patient. A meaningful result will be useful, but will only be obtained if the test is correctly performed to standard spirometric guidance techniques by a trained professional. In particular, the GC noted that in children the technique can be variable, and often needs repeating to ensure reliability. If small numbers for FEV1 are obtained, then they should be interpreted with caution. Children <5 years of age are unlikely to be able to perform the test reliably and thus the literature was not searched for this age group.
Economic considerations	No economic evaluations were found which assessed the use of bronchodilator reversibility as part of a diagnostic pathway. Therefore, an original health economic model was built to assess the cost-effectiveness of different diagnostic pathways for asthma. The model assessed the additional costs of tests against cost-savings from unnecessary asthma medication, and the increased health outcomes from providing correct treatment. The GC agreed that bronchodilator reversibility tests should be used on all patients with an obstructive spirometry in all assessed diagnostic pathways. The reason is that bronchodilator reversibility can be performed at a low cost immediately after initial spirometry, and a positive result is recognised as strong indication that the individual has asthma. However the GC noted that the clinical evidence showed it did not have a high specificity and that there were other obstructive airway diseases, such as COPD, that could produce a positive result. Due to this the GC decided to assess diagnostic strategies where further testing is conducted after positive and negative BDR results and another strategy where no further testing is conducted after a positive BDR. The model results showed that the strategy that did no further testing after a positive BDR test was dominated. Therefore the cost-effectiveness of bronchodilator reversibility testing is contingent on the recommended diagnostic pathway being completed after the results are produced. Stopping the diagnostic pathway after the BDR is conducted will lead to higher costs and poorer health outcomes. The most cost-effective strategy involved using spirometry, bronchodilator reversibility (BDR), FeNO, peak expiratory flow variability (PEFv) and a methacholine challenge test (MCT) in selected individuals to diagnose asthma. In this strategy, everybody with symptoms of asthma would undergo a spirometry test and a FeNO test, those who had an obstructive spirometry would also receive a BDR test. Only those who had non-obstructive spirometry and conflicting FeNO and PEFv test

	results would receive a MCT. Adopting this strategy dominated all other strategies apart from those which performed challenge tests at more points in the pathway. The ICERs of adopting these further stratgies were above £20,000 per QALY gained. For children the GC considered the unit cost of conducting a bronchodilator reversibility test. The GC noted that, unlike adults, there are far fewer diseases that would lead to an obstructive spirometry that does not reverse in children. Although there was no diagnostic accuracy evidence available for children the general consensus of the GC, as well as the recommendations from other guidelines, suggests that a positive bronchodilator reversibility test is enough to confirm the diagnosis of asthma in children. Therefore in children a bronchodilator reversibility test has high value relative to its low cost. The GC noted that a negative BDR test would not rule out the diagnosis of asthma and therefore there was value in further testing beyond this point to prevent false-negative diagnoses.
Quality of evidence	The quality of the evidence ranged from very low to low by GRADE criteria. Unfortunately, none of the studies available to assess the utility of BDR in diagnosing asthma were purpose-designed with this aim. The included population of the majority of the studies was people with asthma or COPD, rather than suspected asthma. These studies aimed to assess the diagnostic accuracy of BDR in distinguishing between asthma and COPD. These studies were included in the review due to limited evidence in the suspected asthma population. This was deemed relevant by the GC as only populations with obstructive airways would be tested using BDR in the diagnostic algorithm. These people either have asthma or COPD, and the task from this point onwards is to distinguish between the two. The reference standard in the included studies was physician diagnosis. With the exception of two studies[21,141], it was unclear if the reference standard included an objective test for asthma. In one study[141], it was unclear whether all people received the objective test. Again, due to limited evidence, these studies were included, but the quality of the evidence was downgraded for population and reference standard indirectness. The GC noted that the dose and bronchodilator used for the BDR test varied between studies; however, they thought that this should not negate the results. • In children aged <5 years, we did not search for studies as BDR is not able to be performed very well in this age group. • In children aged 5-16 years, there were no included studies that addressed the use of BDR in this age group. • In adults, there were four included studies that addressed the use of BDR in this age group. Evidence was of low and very low quality. The economic evidence was assessed as directly applicable with minor limitations.
Other considerations	The GC agreed that there was sufficient evidence to make a recommendation in adults. BDR is a commonly used, simple-to-perform test, that could be carried out in primary or secondary care by a trained professional to standard techniques, and most patients should have no difficulty performing the test. However, children under 5 years old are unable to perform this test reliably. There were no studies included for children 5-16 years old; for this reason, the recommendation is based on extrapolation from the adult data and thus limited by its directness. The strength of the recommendation in children reflected the fact that the recommendation was based on GC consensus and extrapolation from adult data.

	Details of the feasibility project are available in appendix Q. The aim of the feasibility project was to assess the impact and feasibility of adopting the diagnostic tests (e.g. FeNO and spirometry) recommended, into primary care. **General concerns** Firstly it was recognised that children may not be able to perform some of the tests. An additional recommendation has now been made which informs the clinician what to do should this problem arise. The GC agreed that objective testing was imperative before a diagnosis of asthma could be made however, before objective testing is possible, symptoms should be treated and monitored. Secondly the feasibility report identified that in the current format the diagnostic algorithms could be difficult to follow in some places. Therefore the GC agreed a new format should be designed that would simplify the algorithm and make them easier to interpret. **Concerns specific to bronchodilator reversibility** The feasibility project did not identify any issues specific to bronchodilator reversibility, other than those already discussed in section 11.6 related to performing the spirometric measurements which are integral to reversibility testing. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

13 Diagnosis: Peak expiratory flow variability

13.1 Introduction

Peak expiratory flow (PEF) is an objective measure of lung function that has been widely used in the diagnosis and monitoring of asthma for many years. It is a measure of the maximum rate of expiration, generally expressed in litres/minute, and reduces as the airways become narrowed due to bronchoconstriction. Variations in PEF occurring in an individual over time, either spontaneously or in response to medication or to challenges with allergens or inhaled bronchoconstrictors, can be useful in demonstrating variable bronchoconstriction in diagnosing asthma and in assessing the degree of bronchoconstriction. It is assessed with a peak flow meter, a small, inexpensive hand-held device. A variety of peak flow meters are available, and they can be provided to individual patients for home monitoring. The level of peak flow variability (assessed typically as the best of 3 recordings measured between twice and four times a day) over a period of self-monitoring (e.g. 2 weeks) can be used as a diagnostic test for asthma, and similar monitoring during days at work and days away from work in the diagnosis of occupational asthma. Personal asthma action plans based on monitoring of the percentage of best PEF recorded are widely used. Although other measures of lung function can provide more detailed information, the ease of use and simple, inexpensive nature of the monitoring equipment have made PEF monitoring common and popular.

13.2 Review question: In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of peak expiratory flow (PEF) variability?

For full details see review protocol in Appendix C.

Table 31: PICO characteristics of review question

Component	Description
Population / Target condition	People with suspected asthma (presenting with respiratory symptoms). Ages stratified into the following 2 different groups: • Children/ young people (5-16 years old) • Adults (>16 years old)
Index test	PEF variability (diurnal variability usually expressed as amplitude (highest – lowest reading) as a percentage of the mean or the highest reading). PEFv values should be recorded as the mean over a period of at least 3 days)
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test
Outcomes	• Diagnostic accuracy (sensitivity, specificity)

13.3 Clinical evidence

We searched for cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of peak expiratory flow (PEF) variablity to identify whether the condition is present (as indicated by the reference standard) in people under investigation for asthma.

Four studies were included in the review[23,44,174,181] (see Table 32). Evidence from these studies is summarised in the summary tables and the clinical evidence profile below (Table 33). See also the

study selection flow chart in Appendix D, sensitivity / specificity forest plots in Appendix J, study evidence tables in Appendix G and exclusion list in Appendix K.

All four studies assessed the diagnostic accuracy of PEF variability versus a physician diagnosis of asthma plus objective tests. Two studies provided evidence in adults[44,174] and two studies provided evidence in children aged 5-16 years[23,181]. These age groups were analysed in the separate strata.

Summary of included studies

Table 32: Summary of studies included in the review

Study	N	Index test/reference standard	Index test cut-off for positivity	Population	Age
PEF variability vs. reference standard (physician diagnosis with objective test)					
BROUWER 2010[23]	61	PEF variation Asthma diagnosed by paediatric pulmonologist including history. physical examination and lung function tests including methacholine challenge	>95th centile for healthy children i.e. ≥12.3%	Children with non-specific respiratory symptoms such as cough and breathlessness in whom GP uncertain of diagnosis	6 to 16 years; mean 10.4 years
DENOTTER1997[44]	323	PEF variability = ($PEF_{highest}$ – PEF_{lowest})/ PEF_{mean} x 100% = amplitude % mean (average over period) Reference standard =physician diagnosis plus BHR, defined as a PC20 histamine of ≤8 mg/ml	>5% or 10% or 15%	Adults with signs or symptoms indicating asthma (persistent or recurrent respiratory symptoms or signs of reversible bronchial obstruction)	Adults 25–70 years old; mean 43 (12) years
THIADENS1998[174]	170	PEF variability (DPV) = ($PEF_{highest}$ – PEF_{lowest})/ $PEF_{highest}$ x 100% = amplitude % highest (a) MDPV = mean over 2 week period (b) DPV more than threshold on 4 days or more (c) DPV more than	Cut-off values: (a) MDPV > 10% and MDPV >15% (b) DPV >15% on 4 days or more (c) DPV >20% on 3 days or more	18–75 yrs of age, who consulted their GP with coughing that had lasted for at least 2 weeks	Mean 44 (16) years

Study	N	Index test/reference standard	Index test cut-off for positivity	Population	Age
		threshold on 3 days or more Reference standard: A patient was considered to have asthma if there had been a previous period of respiratory symptoms for >3weeks in the last year, accompanied by a provocative dose causing a 20% fall in FEV1 (PD20) ≤15.6 µmol methacholine and/or reversibility ≥9% of predicted			
ULRIK2005[181]	74 people with asthma out of sample of 609adolescents and young adults in survey	PEF variability (amp%mean) Reference standard: 1) Histamine challenge test; cut off PC20 <16.0mg/mL histamine (airways hyper-reactivity) 2) Bronchodilator reversibility: change in FEV1 (ΔFEV1%post) >10%	PEF amp%mean ≥20%	Current asthma (symptoms of asthma, episodes of wheezing and/or treatment for asthma in preceding 12 months)	Aged 13-23 years (mean 18.5 (2.8) years

Table 33: Clinical evidence profile: PEF variability vs. Physician Dx of asthma

PEF variability (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Median Sensitivity % (range)	Median Specificity % (range)[e]	Area Under Curve (range)	Quality
ADULTS >16 years										
Diurnal PEFv as amp%mean (mean over 3 weeks >5%)	1	323	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[c]	0.56	0.69	-	MODERATE
Diurnal PEFv as amp%mean (mean over 3 weeks >10%)	1	323	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[c]	0.14	0.96	-	MODERATE
Diurnal PEFv as amp%mean (mean over 3 weeks >15%)	1	323	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[c]	0.05	0.98	-	MODERATE
Diurnal PEFv as amp%highest (diurnal variation >15% on 4 or more days)	1	170	No risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[c]	0.20	0.97	-	HIGH
Diurnal PEFv as amp%highest (diurnal variation >20% on 3 or more days)	1	170	No risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[c]	0.12	0.99	-	HIGH
Diurnal PEFv as amp%highest (mean over 2 weeks >10%)	1	170	No risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[c]	0.14	0.97	-	HIGH
Diurnal PEFv as amp%highest (mean over 2 weeks >15%)	1	170	No risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[c]	0.03	0.99	-	HIGH
CHILDREN 5-16 years										
Diurnal PEFv as amp%mean (mean over 2 weeks >12.3%)	1	61	No risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[c]	0.50	0.72	-	HIGH
Amp%mean (>20% versus PC20 histamine >16mg/ml)	1	74	No risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[c]	0.46	0.80	-	HIGH
Amp%mean (>20% versus bronchodilator reversibility change in FEV1 >10%)	1	74	No risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[c]	0.71	0.58	-	HIGH

(a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.
(b) Inconsistency was assessed by inspection of the sensitivity / specificity values in each of the studies in relation to a high (≥75%), moderate (50-74%) or low (<50%) effect estimate. Inconsistency was considered high if the point estimates from individual studies were within two areas, and very high if they were in 3 areas or both high and low areas.
(c) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed

13.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow chart in Appendix E.

New cost-effectiveness analysis

An original health economic model was built for adults to assess the cost-effectiveness of several diagnostic pathways, which included peak expiratory flow variability. Full details of the model can be found in Appendix M. A summary of the model can be found in section 18.4.

Unit costs of performing peak expiratory flow variability on children

As an economic model was not feasible for children, the GC considered the unit cost of performing peak expiratory flow variability to evaluate its cost-effectiveness as part of a pathway for diagnosing asthma. This unit cost is presented below (Table 34).

Table 34: Cost of peak expiratory flow variability

Item	Quantity[(a)]	Unit cost	Total Cost (quantity*unit cost)	Source of unit cost
Time taken to instruct patient how to use test with GP practice nurse	10 minutes	£0.73 per minute	£7.30	PSSRU[40]
Time taken to interpret results by GP practice nurse	10 minutes	£0.73 per minute	£7.30	PSSRU[40]
Mini wright peak flow meter	1	£6.48 per meter	£6.48	NHS supply catalogue[47]
Total			**£21.08**	

(a) Based on GC opinion.

The GC also acknowledged the annual cost of drugs for the management of asthma in children. Preventing these costs from occurring in children without asthma would be a large benefit derived from a diagnostic strategy with a high specificity. This cost was estimated to be £201 from a study by Main et al.[102,102]

13.5 Evidence statements

Clinical

- One study with 323 adults showed that PEF variability (mean amp%mean >5%) has a sensitivity of 0.56 and a corresponding specificity of 0.69 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 323 adults showed that PEF variability (mean amp%mean >10%) has a sensitivity of 0.14 and a corresponding specificity of 0.96 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)

- One study with 323 adults showed that PEF variability (mean amp%mean >15%) has a sensitivity of 0.05and a corresponding specificity of 0.98 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 170 adults showed that PEF variability (amp%highest >15% on 4 days or more) has a sensitivity of 0.20 and a corresponding specificity of 0.97 for diagnosing asthma in people presenting with respiratory signs and symptoms. (HIGH QUALITY)
- One study with 170 adults showed that PEF variability (amp%highest >20% on 3 days or more) has a sensitivity of 0.12 and a corresponding specificity of 0.99 for diagnosing asthma in people presenting with respiratory signs and symptoms. (HIGH QUALITY)
- One study with 170 adults showed that PEF variability (mean amp%highest >10%) has a sensitivity of 0.14 and a corresponding specificity of 0.97 for diagnosing asthma in people presenting with respiratory signs and symptoms. (HIGH QUALITY)
- One study with 170 adults showed that PEF variability (mean amp%highest >15%) has a sensitivity of 0.03 and a corresponding specificity of 0.99 for diagnosing asthma in people presenting with respiratory signs and symptoms. (HIGH QUALITY)
- One study with 61 children and young people showed that PEF variability (mean amp%mean >12.3%) has a sensitivity of 0.50 and a corresponding specificity of 0.72 for diagnosing asthma in people presenting with respiratory signs and symptoms. (HIGH QUALITY)
- One study with 74 children and young people showed that PEF variability (amp%mean >20% versus PC20 histamine >16mg/mL) has a sensitivity of 0.46 and a corresponding specificity of 0.80 for diagnosing asthma in people presenting with respiratory signs and symptoms. (HIGH QUALITY)
- One study with 74 children and young people showed that PEF variability (amp%mean >20% versus bronchodilator reversibility change in FEV1 >10%) has a sensitivity of 0.71 and a corresponding specificity of 0.58 for diagnosing asthma in people presenting with respiratory signs and symptoms. (HIGH QUALITY)

Economic

- No relevant economic evaluations were identified.
- An original health economic model found that peak expiratory flow variability (together with spirometry, bronchodilator reversibility, FeNO and histamine or methacholine challenge test) was part of the most cost-effective diagnostic pathway used to diagnose asthma in adults aged 16 and over (see diagnostic algorithm in section 4.1). This evidence is directly applicable with minor limitations.

13.6 Recommendations and link to evidence

Recommendations	**15. Monitor peak flow variability for 2 to 4 weeks in adults (aged 17 and over) if there is diagnostic uncertainty after initial assessment and a FeNO test and they have either:** • **normal spirometry or** • **obstructive spirometry, reversible airways obstruction (positive BDR) but a FeNO level of 39 ppb or less.** **Regard a value of more than 20% variability as a positive test.** **16. Consider monitoring peak flow variability for 2 to 4 weeks in adults (aged 17 and over) if there is diagnostic uncertainty after initial assessment and they have:** • **obstructive spirometry and**

	• **irreversible airways obstruction (negative BDR) and** • **a FeNO level between 25 and 39 ppb.** **Regard a value of more than 20% variability as a positive test.** **17. Monitor peak flow variability for 2 to 4 weeks in children and young people (aged 5 to 16) if there is diagnostic uncertainty after initial assessment and a FeNO test and they have either:** • **normal spirometry or** • **obstructive spirometry, irreversible airways obstruction (negative BDR) and a FeNO level of 35 ppb or more.** **Regard a value of more than 20% variability as a positive test.**
Relative values of different outcomes	There is evidence of significant misdiagnosis of asthma, with both false positive and false negative errors (see introduction to the guideline). The GC was therefore interested in both the sensitivity and specificity of diagnostic tests and prioritised studies which reported each of these outcomes for bronchodilator reversibility.. Data were available from a wide range of PEF variability measures and cut-off thresholds. The GC was primarily interested in diurnal PEFv expressed as the amplitude as a percentage of the highest daily PEF. One study reported the diurnal variation as the amplitude as a percentage of the mean. It was acknowledged that this is slightly different to the amplitude as a percentage of the highest or mean, and may give different estimates of accuracy, although the two indices will correlate closely since their derivation is so similar. Ideally, the diurnal variation should then be averaged over a period of 2 or more weeks. Recordings over only a few days have less chance of capturing the day to day variation which is classical of asthma, and there is often a learning effect in the first days of measurement. The traditional cut-off for normality is a variability of >20% (as recommended by the BTS/SIGN guideline). Some studies reported cut-off thresholds of >5%, >10% or >15%. As expected, higher percentage variability cut-off thresholds increased the specificity of the test but at the expense of a decreased sensitivity. However, the GC was interested in PEFv as part of a diagnostic algorithm, and a negative result on its own would not be used to rule-out asthma.
Trade-off between clinical benefits and harms	In adults, only one study reported a PEFv at the widely used cut-off threshold of >20%. However, this used a daily variability of >20% on 3 or more days as the cut-off threshold, rather than the mean variability over 2-3 weeks. This was considered acceptable as a person with asthma would not necessarily have >20% variation every single day, and it was the only available evidence at the 20% threshold. At this threshold, PEFv had a high specificity but a low sensitivity for the diagnosis of asthma. PEFv would therefore be a better rule-in than rule- out test. The GC noted that within a research environment, PEFv is a relatively specific test and can be used to positively rule in asthma. However, in clinical practice it is likely that there will be more false positives, since inaccurate recording is less likely in a trial setting. The PEFv diary is simple, non-invasive and available in primary care, and there are no significant risks to the patient in performing PEF recording. The GC noted that it is currently a mainstay of the primary care diagnostic process for asthma although the test relies on patient technique, effort and concordance with the frequency of readings. The GC discussed the use of PEFv in children. Children under 5 are unable to reliably

	perform the technique. In older children, PEF monitoring may be difficult and the available evidence suggested the specificity was lower in children aged 5-16 years than in adults, and although the sensitivity was higher, it was still moderate to low. The GC noted that as not all children would receive a PEFv test the average unit cost of including PEFv in the proposed diagnostic algorithm would be less.
Economic considerations	No economic evaluations were found which assessed the use of PEFv as part of a diagnostic pathway. Therefore, an original health economic model was built to assess the cost-effectiveness of different diagnostic pathways for asthma. The model assessed the additional costs of tests against cost-savings from unnecessary asthma medication, and the increased health outcomes from providing correct treatment. The GC decided that PEFv should appear further down a diagnostic pathway as FeNO had a higher sensitivity and specificity and could therefore confirm an asthma diagnosis sooner when combined with spirometry and BDR results. Therefore PEFv was reserved for patients who could not perform a BDR or did not have a clear diagnosis using the results from spirometry, BDR and FeNO. Two strategies were tested which did not include PEFv, but the strategies which included PEFv dominated those without it; therefore, the model supports the use of PEFv as a cost-effective test to use following within certain diagnostic pathways. The most cost-effective strategy involved using spirometry, bronchodilator reversibility (BDR), FeNO, peak expiratory flow variability (PEFv) and a methacholine challenge test (MCT) in selected individuals to diagnose asthma. In this strategy, everybody with symptoms of asthma would undergo a spirometry test and a FeNO test, those who had an obstructive spirometry would also receive a BDR test. Only those who had non-obstructive spirometry and conflicting FeNO and PEFv test results would receive a MCT. Adopting this strategy dominated all other strategies apart from those which performed challenge tests at more points in the pathway. The ICERs of adopting these further stratgies were above £20,000 per QALY gained. For children the GC considered the unit cost of performing a PEFv test. The GC noted from the clinical review that PEFv tests had a moderate/high specificity and a moderate sensitivity. Therefore a positive result was likely to indicate asthma whereas a negative result would not rule out asthma. The GC therefore considered that due to its low cost PEFv would have value in a diagnostic pathway.
Quality of evidence	The quality of evidence ranged from moderate to high in adults. For the ideal cut-off threshold of 20%, evidence was of high quality, but the evidence was only available from one study. In children aged 5-16 years, the quality of the evidence was high; however, the evidence for each PEFv measure was only available from one small study. The included studies varied in terms of the length of diary monitoring. Most studies stipulated at least 4 readings a day, perhaps more than would typically be expected in clinical practice. • In children aged <5 years, we did not search for studies as PEFv cannot be performed in this age group. • In children aged 5-16 years, there were two included studies (Brouwer 2010, Ulrik 2005) using the best reference standard (physician diagnosis with objective test) that addressed the use of PEFv in this age group. The evidence at the preferred cut-off value of 20% was of high quality. • In adults, there were two included studies (Thiadens 1998, and Denotter 1997) using the best reference standard (physician diagnosis with objective test) that addressed the use of PEFv in this age group. The evidence at the preferred cut-off value of 20% was of high quality. The economic evidence was assessed as directly applicable with minor limitations.

Other considerations	After reviewing the available evidence, the GC proposes a variability of greater than 20% in PEF readings in accordance with consensus practice. The GC recognises that this cut-off is most appropriate when 4 readings a day are taken, but this may not always be possible in routine practice. The GC noted that healthcare professionals will need to calculate the amplitude of variation of PEF. The common practice of looking at the pattern of variation in a diary is unlikely to be as accurate. If GPs have access to computerised tools for calculating PEFv, amplitude as a percentage of mean is the best measure to use. In practice amplitude as a percent of highest value is easier to calculate and is acceptable. The evidence suggests that PEF variability has some value as a rule-in test for asthma (specificity was high) and it has the advantage of being well-established for use in primary care. However, the studies did not give a clear indication of the optimal cut-off point for clinical use. PEF recording is not possible in children less than 5 and requires concordance of the patient and good technique to obtain reproducible accurate readings. Some patient groups may struggle with this. Details of the feasibility project are available in appendix Q. The aim of the feasibility project was to assess the impact and feasibility of adopting the diagnostic tests (e.g. FeNO and spirometry) recommended, into primary care. **General concerns** Firstly it was recognised that children may not be able to perform some of the tests. An additional recommendation has now been made which informs the clinician what to do should this problem arise. The GC agreed that objective testing was imperative before a diagnosis of asthma could be made. However, before objective testing is possible, symptoms should be treated and monitored. Secondly the feasibility report identified that in the current format the diagnostic algorithms could be difficult to follow in some places. Therefore the GC agreed a new format should be designed that would simplify the algorithm and make them easier to interpret. **Concerns specific to peak flow variability** The project identified that compliance to peak flow diaries was fairly poor. This is consistent with the GC's experience, and reaffirmed their view that peak flow variability should only be assessed if diagnostic uncertainty remained after conducting FeNO, spirometry and bronchodilator reversibility. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

14 Diagnosis: Skin prick tests

14.1 Introduction

Asthma can be divided into extrinsic asthma (atopic or allergic), intrinsic asthma (non-atopic) and occupational asthma. Allergic asthma is the commonest type, is associated with atopy and usually develops in childhood or early adulthood. Atopy is defined as a genetic predisposition to produce immunoglobulin E (IgE) against common environmental aeroallergens such as house dust mites, animal dander, pollens and moulds. Approximately 80% of people with asthma are atopic compared with 30% of the general population.

People with allergic asthma are initially sensitised to allergens and subsequently develop symptoms on re-exposure. Continued exposure may lead to a chronic inflammatory response, which is characterised by persistent symptoms, airways hyper-reactivity and bronchospasm.

Skin prick tests for the most common aeroallergens can be performed to confirm the presence or absence of atopy to individual aeroallergens. However, the diagnostic test accuracy of skin prick tests to diagnose asthma is currently uncertain.

14.2 Review question: In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of skin prick tests?

For full details see review protocol in Appendix C.

Table 35: PICO characteristics of review question

Component	Description
Population / Target condition	People with suspected asthma (presenting with respiratory symptoms). All ages, stratified into the following 3 different groups: • Children (1-<5 years old) • Children/young people (5-16 years old) • Adults (>16 years old)
Index test	Skin prick tests for the most common allergens (report separately) • House dust mites • Cat • Dog • Grass pollen* (native UK grasses) • Tree pollen* (native UK trees) • Mixed pollens* (native UK species) • *Aspergillus* • *Alternaria* • *Cladosporium* Cut off values: 3mm WHEAL (skin reaction) greater than the negative control in the presence of a positive control
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test. In children 1-<5 years, objective tests cannot be performed so the reference standard will be physician diagnosis based on recurrent and persistent wheezing.

Outcomes	• Diagnostic accuracy (sensitivity, specificity)

14.3 Clinical evidence

We searched for cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of skin prick tests to identify whether the condition is present (as indicated by the reference standard) in people under investigation for asthma.

Six studies were included in the review[51,58,105,110,137,168]. Evidence from these are summarised in Table 36 and the clinical evidence profile below (Table 37). See also the study selection flow chart in Appendix D, forest plots in Appendix J, study evidence tables in Appendix G and exclusion list in Appendix K.

Three of the studies[51,137,168] fully complied with the protocol (ie. the reference standard included an objective test for asthma). Out of these three studies, two were in adults and the other study was in children, these strata were reported separately. The other three studies[58,105,110] did not mention an objective test as part of the reference standard. The reference standard was a physician diagnosis of asthma. These studies were included since there was very little evidence otherwise, but were reported separately in the analysis. Of these studies, two fell into the 5-16 years strata and one fell into the adult strata, however the mean age was not reported in two of the studies. No evidence was identified in children aged <5 years.

Summary of included studies

Table 36: Summary of studies included in the review

Study	N	Index test	Index test cut-off for positivity	Population	Age
Skin prick test vs. reference standard (physician diagnosis with objective test)					
DRKULEC 2013[51]	131 (N=71 asthma)	• SPT for Dust mite *D. pteronyssinus* • *Ambrosia artemisifoliae* • *Phleum pratense*	Not stated. Each allergen separately or positive SPT to 1 or more allergens	Children with chronic cough	Median 7.5 years
POPOVIC 2002	N=195 (N=141 asthma)	• SPT for ≥1 aeroallergen (dust mite *D. pteronyssinus,* grass pollen, tree pollen, cat fur, dog fur)	Wheal ≥3mm	Suspected asthma: had dyspnoea and treated for breathlessness.	Mean 36.5 years
SORIANO 1999A	N=1816 (N=136 asthma)	• SPT for ≥1 aeroallergen • Also Individual allergens: • Dust mite *D. pteronyssinus* • *Cladosporium* • *Alternaria* • Timothy grass • Birch pollen	Wheal ≥3mm	Suspected asthma: subsample of general population who had respiratory symptoms	Mean 32 years

Study	N	Index test	Index test cut-off for positivity	Population	Age
		• Cat			
Skin prick test vs. reference standard (physician diagnosis without objective test)					
GAIG 1999[58]	94 (N=41 asthma)	• Dust mite *D. pteronyssinus* and *D. farina*	Skin wheal diameter to at least one of the two mites 3mm larger than control	Patients attending outpatient allergy clinic who had been sharing a bunk with a sibling for >6 months, occupying always the same position (top or bottom bunk)	Mean 16 years (range not reported)
MAY 1990[105]	446 (N=190 asthma)	• *Gramineae* (grasses both wild and cultivated) • *Artemisia vulgaris* (weed: mugwort)	3+ or 4+	Consecutive unselected patients for allergological consultation for conjunctivitis, rhinitis and/or asthma which appeared or deteriorated in late spring and summer	Range 6 - 56 years, (mean not reported)
MIRAGLIA DEL GIUDICE 2002[110]	1426 (N=925 asthma)	• SPT for ≥1 aeroallergen [house dust mites, HDM (*D. pteronyssinus* and *D. farina)*, *Parietaria officinalis*, grasses (*Dactylis glomerata, Lolium perenne, Phaleum pratense*), moulds (*Alternaria, Aspergillus, Cladosporium*), dog fur, cat fur, egg albumin, and cow's milk	Wheal was at least 3 mm in diameter	Children referred to Paediatric Asthma and Allergy Center because of allergic symptoms: children in whom a diagnosis of asthma, allergic rhinoconjunctivitis, atopic dermatitis or food allergy was confirmed by a pediatric allergologist.	Range 0 - 12 years, (mean not reported)

SPT = skin prick test; BPT = bronchial provocation test; IST = intradermal skin test

Table 37: Clinical evidence profile: Skin prick test vs. Physician Dx of asthma

Skin prick test	No of studies	N	Risk of bias	Inconsistency	Indirectness	Imprecision	Median Sensitivity % (range)	Median Specificity % (range)[e]	Area Under Curve (range)	Quality
PHYSICIAN DX WITH OBJECTIVE TEST										
ADULTS										
D. pteronyssinus (Der P) +/- *D. farinae* (house dust mite).	1	1816	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[e]	N/A[d]	0.39	0.80	-	LOW
Alternaria temius (mould).	1	1816	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[e]	N/A[d]	0.07	0.99	-	LOW
≥1 positive from mixed allergens (all studies included mite and grass, plus ≥1 of weed, tree, dust, cat, dog, feathers, mould, egg, milk).	2	2011	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[e]	N/A[d]	Range 0.61 – 0.62	Range 0.63-0.69	-	LOW
Grasses mixed or timothy only.	1	1816	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[e]	N/A[d]	0.32	0.87	-	LOW
Cladosporium	1	1816	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[e]	N/A[d]	0.07	0.97	-	LOW
Cat	1	1816	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[e]	N/A[d]	0.21	0.94	-	LOW
CHILDREN 5-16 years										
D. pteronyssinus (Der P) +/- *D. farinae* (house dust mite).	1	131	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[d]	0.83	0.71	-	MODERATE
Phleum pratense (Phl P) timothy grass from Gramineae family.	1	131	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[d]	0.66	0.50	-	MODERATE
Ambrosia artemisifoliae (Amb A) common ragweed.	1	131	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[d]	0.66	0.48	-	MODERATE
≥1 positive from mixed allergens (all studies included mite and grass, plus ≥1 of weed, tree, dust, cat, dog, feathers, mould, egg, milk).	1	131	Serious risk of bias[a]	Serious inconsistency[b]	No serious indirectness	N/A[d]	0.79	0.92	-	MODERATE

Skin prick test	No of studies	N	Risk of bias	Inconsistency	Indirectness	Imprecision	Median Sensitivity % (range)	Median Specificity % (range)[e]	Area Under Curve (range)	Quality
Grasses mixed or timothy only.	1	131	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[d]	0.66	0.50	-	MODERATE
CHILDREN <5 years										
No evidence identified	0									
PHYSICIAN DX - NO OBJECTIVE TEST										
ADULTS										
Gramineae (grasses) both wild and cultivated.	1	446	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[c]	N/A[d]	0.89	0.11	-	LOW
Artemisia vulgaris (mugwort).	1	446	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[c]	N/A[d]	0.48	0.63	-	LOW
Grasses mixed or timothy only.	1	446	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[c]	N/A[d]	0.89	0.11	-	LOW
CHILDREN 5-16 years										
D. pteronyssinus (Der P) +/- *D. farinae* (house dust mite).	1	67	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[c]	N/A[d]	0.85	0.35	-	LOW
≥1 positive from mixed allergens (all studies included mite and grass, plus ≥1 of weed, tree, dust, cat, dog, feathers, mould, egg, milk).	1	1426	No risk of bias	No serious inconsistency[b]	Serious indirectness[c]	N/A[d]	0.44	0.56	-	MODERATE
CHILDREN <5 years										
No evidence identified	0									

(a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.
(b) Inconsistency was assessed by inspection of the sensitivity / specificity values in each of the studies in relation to a high (≥75%), moderate (50-74%) or low (<50%) effect estimate. Inconsistency was considered high if the point estimates from individual studies were within two areas, and very high if they were in 3 areas or both high and low areas.
c) None of the studies had objective test as part of the physician's diagnosis of asthma. Population age range spans adult and children population strata with no subgroup analysis.
d) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed.
e) Population is people with symptoms identified from a questionnaire in the general population (not visiting the GP with symptoms).

14.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow chart in Appendix E.

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs were provided to aid consideration of cost-effectiveness. The GC considered this unit cost alongside the diagnostic pathway evaluated in the economic model.

The cost of a skin prick test was identified using the NHS reference costs[45]. Skin prick tests fall under the HRG code JC11Z 'Other Diagnostic Skin Tests', the OPCS code is U27.8. The average 'Total HRG' unit cost is quoted as £195. The GC agreed that 'outpatient with procedure' was the only setting a skin prick test would be performed in for the consideration on asthma patients. The unit cost quoted for this setting is £173 and it was therefore agreed that this is the most relevant cost to use for skin prick tests.

14.5 Evidence statements

Clinical

Physician Dx with objective test:

- One study with 1816 adults showed that skin prick test (D. pteronyssinus (Der P) +/- D. farina (house dust mite)) has a sensitivity of 0.39 and a corresponding specificity of 0.80 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 1816 adults showed that skin prick test (Alternaria temius (mould)) has a sensitivity of 0.07 and a corresponding specificity of 0.99 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- Two studies with 2011 adults showed that skin prick test (≥1 positive from mixed allergens (all studies included mite and grass, plus ≥1 of weed, tree, dust, cat, dog, feathers, mould, egg, milk)) has a sensitivity range of 0.61 to 0.62 and a corresponding specificity range of 0.63 to 0.69 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 1816 adults showed that skin prick test (grasses mixed or timothy only) has a sensitivity of 0.32 and a corresponding specificity range of 0.87 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 1816 adults showed that skin prick test (Cladosporium) has a sensitivity of 0.07 and a corresponding specificity of 0.97 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 1816 adults showed that skin prick test (cat) has a sensitivity of 0.21 and a corresponding specificity of 0.94 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 131 children and young people showed that skin prick test (D. pteronyssinus (Der P) +/- D. farina (house dust mite)) has a sensitivity of 0.83 and a corresponding specificity of 0.71 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 131 children and young people showed that skin prick test (Phleum pratense (Phl P) timothy grass from Gramineae family) has a sensitivity of 0.66 and a corresponding specificity

of 0.50 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)

- One study with 131 children and young people showed that skin prick test (Ambrosia artemisifoliae (Amb A) common ragweed) has a sensitivity of 0.66 and a corresponding specificity of 0.48 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 131 children and young people showed that skin prick test (≥1 positive from mixed allergens (all studies included mite and grass, plus ≥1 of weed, tree, dust, cat, dog, feathers, mould, egg, milk)) has a sensitivity of 0.79 and a corresponding specificity of 0.92 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 131 children and young people showed that skin prick test (grasses mixed or timothy only) has a sensitivity of 0.66 and a corresponding specificity of 0.50 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- No evidence was identified in children <5 years

Physician Dx with no objective test:

- One study with 446 adults showed that skin prick test (Gramineae (grasses) both wild and cultivated) has a sensitivity of 0.89 and a corresponding specificity of 0.11 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 446 adults showed that skin prick test (Altemisia vulgaris (mugwort)) has a sensitivity of 0.48 and a corresponding specificity of 0.63 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 446 adults showed that skin prick test (grasses mixed or timothy only) has a sensitivity of 0.89 and a corresponding specificity of 0.11 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 67 children and young people showed that skin prick test (D. pteronyssinus (Der P) +/- D. farina (house dust mite)) has a sensitivity of 0.85 and a corresponding specificity of 0.35 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 1426 children and young people showed that skin prick test (≥1 positive from mixed allergens (all studies included mite and grass, plus ≥1 of weed, tree, dust, cat, dog, feathers, mould, egg, milk)) has a sensitivity of 0.44 and a corresponding specificity of 0.56 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- No evidence was identified in children <5 years

Economic

- No relevant economic evaluations were identified.

14.6 Recommendations and link to evidence

Please see section 15.6.

15 Diagnosis: Serum IgE measures

15.1 Introduction

Asthma can be divided into extrinsic asthma (atopic or allergic), intrinsic asthma (non-atopic) and occupational asthma. Allergic asthma is the commonest type, is associated with atopy and usually develops in childhood or early adulthood. Atopy is defined as a genetic predisposition to produce immunoglobulin E (IgE) against common environmental aeroallergens such as house dust mites, animal dander, pollens and moulds. Approximately 80% of people with asthma are atopic compared with 30% of the general population.

People with allergic asthma are initially sensitised to allergens and subsequently develop symptoms on re-exposure. Continued exposure may lead to a chronic inflammatory response, which is characterised by persistent symptoms, airways hyper-reactivity and bronchospasm.

Both total IgE and IgE specific for aeroallergens can be measured in the serum of individual patients. As a large proportion of people with asthma are atopic, this raises the question of whether measuring total or specific IgE is a good diagnostic test, in the diagnosis of asthma in people presenting with symptoms.

15.2 Review question: In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of total and specific serum IgE measures?

For full details see review protocol in Appendix C.

Table 38: PICO characteristics of review question

Population	People with suspected asthma (presenting with respiratory symptoms). All ages, stratified into the following 3 different groups: • Children (1-<5 years old) • Children/young people (5-16 years old) • Adults (>16 years old)
Index test	Serum IgE • Total IgE • Specific IgE* (including RAST test) *Report separately for the most common aero-allergens (dust mites, grass pollen, tree pollen, dog, cat, *Aspergillus, Alternaria, Cladosporium*). NOTE: serum IgE must have been assessed using ELISA (apart from RAST) as other techniques are not current/no longer used.
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test In children 1-<5 years, objective tests cannot be performed so the reference standard will be physician diagnosis based on recurrent and persistent wheezing.
Outcomes	• Diagnostic accuracy (sensitivity, specificity)

15.3 Clinical evidence

We searched for cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of total and specific serum IgE to identify whether the condition is present (as indicated by the reference standard) in people under investigation for asthma.

Five studies were included in the review[2,96,136,167,178]. Evidence from these are summarised in Table 39 and the clinical evidence profile below (Table 40). See also the study selection flow chart in Appendix D, forest plots in Appendix J, study evidence tables in Appendix G and exclusion list in Appendix K. All studies were conducted in adults. No evidence was identified for the strata of children aged 5-16 years or children <5 years.

All of the included studies[2,96,136,167,178] were cross-sectional studies, and looked at the diagnostic accuracy of IgE in the diagnosis of asthma in a general population, but were included as no evidence was identified in people with suspected asthma. Therefore, the evidence was downgraded for population indirectness. With the exception of one study[96], the reference standard of asthma diagnosis was established from responses to a questionnaire about previous physician diagnosis, and did not include an objective test. However, due to limited evidence, these studies were included and downgraded in quality. One study[96] included a skin prick test as an objective test, however this only confirms allergy and is not an objective test for asthma diagnosis.

Summary of included studies

Table 39: Summary of studies included in the review: IgE versus reference standard (adults)

Study	N	Index test/reference standard	Index test cut-off for positivity	Population	Age
ABRAHAM 2007[2]	702 (N=493 asthma)	Specific IgE • Dust mite • Grass (timothy) • *Alternaria* • Cat • Dog	≥0.35 kU/l	Pregnant women	Adults (21-49 yrs)
LINNEBERG 2006[96]	709 (N=51 asthma)	Specific IgE • Pollen • Dust mite	≥0.35 kU/l	General population sample	Adults (15-69 yrs)
PLASCHKE 1999A[136]	1572 (N=84 asthma)	Specific IgE • Dust mite • Grass • Birch • *Cladosporium* • Cat	≥0.70 kU/l	General population sample	Adults (20-44 yrs)
SORIANO 1999[167]	1816 (N=136 asthma)	Specific IgE or SPT • *Cladosporium* • Dust mite • Grass (timothy) • Cat	≥0.35 kU/l	General population sample	Adults (20-44 yrs)
TSCHOPP 1998[178]	8329 (N=153	Total IgE	≥100 kU/l	General population	Adults

Study	N	Index test/reference standard	Index test cut-off for positivity	Population	Age
	allergic asthma)	Specific IgE • Pollens • Dust mite • Moulds		sample	(18-60 yrs)

Table 40: Clinical evidence profile: IgE vs. Reference Standard

IgE (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Median Sensitivity % (range)	Median Specificity % (range)[e]	Area Under Curve (range)	Quality
ADULTS										
DUST MITE Specific IgE										
IgE ≥0.35 kU/L	3	3227	Serious risk of bias[a]	Serious inconsictency[b]	Serious indirectness[d]	N/A[e]	Median 0.39 (range 0.38 to 0.84)	Corresponding 0.80 (range 0.62 to 0.80)	-	VERY LOW
IgE ≥0.70 kU/L	1	1572	Serious risk of bias[a]	No serious inconsistency	Serious indirectness[d]	N/A[e]	0.19	0.94	-	LOW
BIRCH Specific IgE										
IgE ≥0.70 kU/L	1	1572	Serious risk of bias[a]	No serious inconsistency	Serious indirectness[d]	N/A[e]	0.30	0.90	-	LOW
GRASS Specific IgE										
IgE ≥0.35 kU/L	2	2518	Serious risk of bias[a]	Serious inconsistency[c]	Serious indirectness[d]	N/A[e]	Range 0.33 to 0.68	Range 0.81 to 0.87	-	VERY LOW
IgE ≥0.70 kU/L	1	1572	Serious risk of bias[a]	No serious inconsistency	Serious indirectness[d]	N/A[e]	0.36	0.87	-	LOW
***ALTERNARIA* Specific IgE**										
IgE ≥0.35 kU/L	1	702	Serious risk of bias[a]	No serious inconsistency	Serious indirectness[d]	N/A[e]	0.34	0.85	-	LOW
***CLADOSPORIUM* Specific IgE**										
IgE ≥0.35 kU/L	1	1816	Serious risk of bias[a]	No serious inconsistency	Serious indirectness[d]	N/A[e]	0.07	0.97	-	LOW
IgE ≥0.70 kU/L	1	1572	Serious risk of bias[a]	No serious inconsistency	Serious indirectness[d]	N/A[e]	0.04	0.99	-	LOW
POLLEN Specific IgE										

IgE (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Median Sensitivity % (range)	Median Specificity % (range)[e]	Area Under Curve (range)	Quality
IgE ≥0.35 kU/L	1	709	No serious risk of bias	No serious inconsistency	Serious indirectness[d]	N/A[e]	0.96	0.64	-	MODERATE
TOTAL IgE										
IgE ≥100 kU/L	1	709	No serious risk of bias	No serious inconsistency	Serious indirectness[d]	N/A[e]	0.57	0.78	-	MODERATE
CAT Specific IgE										
IgE ≥0.35 kU/L	2	2518	Serious risk of bias[a]	No serious inconsistency	Serious indirectness[d]	N/A[e]	Range 0.20 to 0.40	Range 0.88 to 0.94	-	LOW
IgE ≥0.70 kU/L	1	1572	Serious risk of bias[a]	No serious inconsistency	Serious indirectness[d]	N/A[e]	0.40	0.91	-	LOW
DOG Specific IgE										
IgE ≥0.35 kU/L	1	702	Serious risk of bias[a]	No serious inconsistency	Serious indirectness[d]	N/A[e]	0.34	0.88	-	LOW
CHILDREN 5-16 years										
No evidence identified	0									
CHILDREN <5 years										
No evidence identified	0									

a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.

b) Inconsistency was assessed by inspection of the sensitivity / specificity RevMan 5 plots. Linneberg is the outlier, the difference between this study and the others (ie. possible reasons for heterogeneity) may be that the study is conducted in a wider age range which included some children (but was mostly adults).

c) Inconsistency was assessed by inspection of the sensitivity / specificity RevMan 5 plots. There was inconsistency for sensitivity but not specificity. The differences between the two studies (ie. possible reasons for heterogeneity) could be that Abraham study was conducted specifically in pregnant women.

d) Studies were considered as indirect because they were not conducted in people with 'suspected asthma'

e) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed.

15.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow chart in Appendix E.

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs were provided below (Table 41) to aid consideration of cost-effectiveness. The GC considered these unit costs alongside the diagnostic pathway evaluated in the economic model.

Table 41: Unit costs for total and specific serum IgE tests

Item	Unit cost	Quantity	Sub total	Source
IgE	£5	1	£5	GC opinion
RAST	£12 - £20[a]	1 per allergen	£12-20 per allergen	GC opinion
Nurse time	£0.75 per minute	5 minutes	£3.75	GC estimate and PSSRU[40,40]
GP time	Average cost of GP appointment (11.7 min) = £36	2 GP appointments (1 for referral and 1 to discuss the results)	£72	PSSRU[40,40]
TOTAL			£92.75 - 112.75[(b)]	

(a) The cost for a RAST is dependent on whether it is identifying common allergen, a recombinant allergen, or a mix.
(b) This is the cost for one allergen, for additional allergens the cost of additional RASTS would need to be added assuming they are all done at the same time.

15.5 Evidence statements

Clinical

DUST MITE-specific IgE:

- Three studies with 3227 adults showed that DUST MITE-specific IgE (IgE cut-off ≥0.35 Ku/L) has a median sensitivity of 0.39 and a corresponding specificity of 0.80 for diagnosing asthma in the general population. (VERY LOW QUALITY)
- One study with 1572 adults showed that DUST MITE-specific IgE (IgE cut-off ≥0.70 Ku/L) has a sensitivity of 0.19 and a corresponding specificity of 0.94 for diagnosing asthma in the general population. (LOW QUALITY)

BIRCH-specific IgE:

- One study with 1572 adults showed that BIRCH-specific IgE (IgE cut-off ≥0.70 Ku/L) has a sensitivity of 0.30 and a corresponding specificity of 0.90 for diagnosing asthma in the general population. (LOW QUALITY)

GRASS-specific IgE:

- Two studies with 2518 adults showed that GRASS-specific IgE (IgE cut-off ≥0.35 Ku/L) has a sensitivity range of 0.33 to 0.68 and a corresponding specificity range of 0.81 to 0.87 for diagnosing asthma in the general population. (VERY LOW QUALITY)

- One study with 1572 adults showed that GRASS-specific IgE (IgE cut-off ≥0.70 Ku/L) has a sensitivity of 0.36 and a corresponding specificity of 0.87 for diagnosing asthma in the general population. (LOW QUALITY)

ALTERNARIA-specific IgE:

- One study with 702 adults showed that ALTERNARIA-specific IgE (IgE cut-off ≥0.70 Ku/L) has a sensitivity of 0.34 and a corresponding specificity of 0.85 for diagnosing asthma in the general population. (LOW QUALITY)

CLADOSPORIUM-specific IgE:

- One study with 1816 adults showed that CLADOSPORIUM-specific IgE (IgE cut-off ≥0.35 Ku/L) has a sensitivity of 0.07 and a corresponding specificity of 0.97 for diagnosing asthma in the general population. (LOW QUALITY)
- One study with 1572 adults showed that CLADOSPORIUM-specific IgE (IgE cut-off ≥0.70 Ku/L) has a sensitivity of 0.04 and a corresponding specificity of 0.99 for diagnosing asthma in the general population. (LOW QUALITY)

POLLEN-specific IgE:

- One study with 709 adults showed that POLLEN-specific IgE (IgE cut-off ≥0.35 Ku/L) has a sensitivity of 0.96 and a corresponding specificity of 0.64 for diagnosing asthma in the general population. (MODERATE QUALITY)

TOTAL-specific IgE:

- One study with 709 adults showed that TOTAL-specific IgE (IgE cut-off ≥100 Ku/L) has a sensitivity of 0.57 and a corresponding specificity of 0.78 for diagnosing asthma in the general population. (MODERATE QUALITY)

CAT-specific IgE:

- Two studies with 2518 adults showed that CAT-specific IgE (IgE cut-off ≥0.35 Ku/L) has a sensitivity range of 0.20 to 0.40 and a corresponding specificity of 0.88 to 0.94 for diagnosing asthma in the general population. (LOW QUALITY)
- One study with 1572 adults showed that CAT-specific IgE (IgE cut-off ≥0.70 Ku/L) has a sensitivity of 0.40 and a corresponding specificity of 0.91 for diagnosing asthma in the general population. (LOW QUALITY)

DOG-specific IgE:

- One study with 702 adults showed that DOG-specific IgE (IgE cut-off ≥0.35 Ku/L) has a sensitivity of 0.34 and a corresponding specificity of 0.88 for diagnosing asthma in the general population. (LOW QUALITY)
- No evidence was identified in children aged 5-16 years.
- No evidence was identified in children <5 years.

Economic

- No relevant economic evaluations were identified.

15.6 Recommendations and link to evidence

Recommendations	**18.Do not offer the following as diagnostic tests for asthma:** • **skin prick tests to aeroallergens** • **serum total and specific IgE.**

	19.Use skin prick tests to aeroallergens or specific IgE tests to identify triggers after a formal diagnosis of asthma has been made.
Relative values of different outcomes	The GC was interested in the diagnostic test accuracy of skin prick tests and of total and specific serum IgE tests in the diagnosis of asthma. There is evidence of significant misdiagnosis of asthma, with both false positive and false negative errors (see introduction to the guideline). The GC was therefore interested in both the sensitivity and specificity of diagnostic tests and prioritised studies which reported each of these outcomes for skin prick tests and IgE measurements. It is important to note that: • In the case of serum IgE, the data are extrapolated from population-based data rather than from people with respiratory symptoms. • For skin prick testing, the data reviewed were based on people with suspected asthma, or used people with allergic symptoms as a reference population. The GC agreed that skin prick test, and specific IgE test plus total IgE, are different modes of testing for a similar phenomenon.
Trade-off between clinical benefits and harms	Total IgE is potentially a useful marker of an allergic state and can be elevated in the absence of positive skin prick or specific IgE titres. A high total IgE has a number of causes and may require further investigation. In adults, the evidence from the studies identified suggests that, for the majority of allergens, total IgE, as well as specific IgE or skin prick tests has poor sensitivity for diagnosing asthma. For the majority of allergens, this poor sensitivity is coupled with a moderate or high specificity. Using these tests in the diagnosis of asthma could help to prevent over-diagnosis of asthma, but would result in a large number of people with asthma being missed. In children aged 5-16 years, evidence suggests that skin prick tests have a moderate sensitivity and specificity for the diagnosis of asthma. Evidence was only available from one study, and no evidence was available for total or specific IgE. In children aged <5 years there was no evidence identified. The GC debated the risks and benefits and considered that in adults the risks and limitations from conducting blood tests (for serum total IgE and specific IgE) were potentially negligible. In relation to paediatrics, it was noted that children may find blood tests distressing, so they are not often used as a diagnostic test in the paediatric population. However, this short-term distress can be minimised by a variety of mechanisms. The risks of skin prick testing were thought to be negligible in adults and children. The GC noted that skin prick tests have the additional benefit that a result will be available within 15 minutes, i.e. within one visit. Whether total or specific IgE or skin prick testing is used is likely to vary depending on the setting. The GC noted that there would be no additional value in performing both tests (skin prick and specific IgE). Despite the low risk associated with performing these tests, the GC were concerned about the low sensitivity of the tests and chose not to recommend either test.
Economic considerations	**Serum IgE:** No relevant economic evaluations were identified for serum total or specific IgE tests. The unit cost for specific and total IgE tests was estimated to be £93 - £113 depending on the allergen identified. An additional £12- £20 would be incurred per extra allergen. **Skin prick tests:** No relevant economic evaluations were identified for skin prick

	tests. Skin prick tests are estimated to cost £173. The GC considered the unit costs of these tests, as well as the downstream implications of correct and incorrect diagnoses. The GC did not think that serum IgE or skin prick tests would be cost-effective as first-line diagnostic tests, as the clinical evidence does not show that they offer sufficient diagnostic accuracy. An original health economic model was built to assess the cost-effectiveness of different diagnostic pathways for asthma in adults. To be of benefit the test would need to lead to change in the diagnostic decision. The GC considered that the results from SPT/IgE tests would not be enough to overturn any of the decisions made in the diagnostic pathway and therefore they were not considered a cost-effective use of resources.
Quality of evidence	**SKIN PRICK** The studies of skin prick testing included populations of patients suspected of having asthma, or having symptoms of asthma and allergic symptoms. The GC referred back to one of the original papers, Soriano1999[168]. This was a large study that dominated the evidence; however, patients were identified from a screening questionnaire as a subset of people with respiratory symptoms (not because they were presenting to the GP). Therefore, the evidence was downgraded for indirectness. The studies included to review the utility of skin prick testing used a cut-off of > 3mm wheal for a positive result, with the exception of one study[51] which did not mention the cut-off that was used. Three included studies used the ideal reference standard of a physician diagnosis with an objective test, and were reported separately. The GC focused on the evidence from these studies. • In children aged <5 years, there were no included studies using the best reference standard (physician diagnosis with objective test) that addressed the use of skin prick tests in this age group. • In children aged 5-16 years, there was one included study (Drkulec) using the best reference standard (physician diagnosis with objective test) that addressed the use of skin prick tests in this age group. The evidence was of moderate quality. • In adults, there were two included studies (Popovic and Soriano) using the best reference standard (physician diagnosis with objective test) that addressed the use of skin prick tests in this age group. The evidence was of low quality. **IgE** Due to a lack of evidence in the review population, the studies included in the analysis of serum total IgE and specific IgE testing were in the general population and the GC had to extrapolate the findings from these. Several studies were found, but were excluded because the reference standard was very similar to the index test (was allergen-specific, e.g. studies comparing specific IgE test to skin prick test or inhaled allergen challenges), and thus do not provide useful information on the utility of the index test in the diagnosis of asthma. The studies relied on questionnaire responses of previous asthma diagnosis, as the reference standard and none of the studies identified had the ideal reference standard of physician diagnosis with an objective test. • In children aged <5 years, there were no included studies that addressed the use of total or specific IgE in this age group. • In children aged 5-16 years, there were no included studies that addressed the use of total or specific IgE in this age group. • All five included studies were conducted in adults. For the majority of the tests,

	evidence was only available from 1 study. Evidence for all except two of the tests was of low and very low quality.
Other considerations	The GC agreed that there is variation in the access and use of specific IgE testing in primary care for aero- and other allergens. Access to skin prick testing in primary care is generally even more limited. It was noted that there are circumstances in which it is extremely useful to know which allergens a person with asthma is sensitised to. This can be useful therapeutically, for example in terms of avoiding exposure and therefore triggering an attack. However, this benefit applies when the diagnosis of asthma has already been established. The GC concluded that the evidence was not strong enough to recommend measurement of IgE or skin tests, but made a recommendation which expressed this in terms of using these as diagnostic tests since they did not wish to discourage appropriate use further along the management pathway. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

16 Diagnosis: Fractional exhaled nitric oxide (FeNO)

16.1 Introduction

Asthma can be divided into extrinsic asthma (atopic or allergic), intrinsic asthma (non-atopic) and occupational asthma. Atopy is defined as a genetic predisposition to produce immunoglobulin E (IgE) against common environmental aeroallergens such as house dust mites, animal dander, pollens and moulds. Approximately 80% of people with asthma are atopic compared with 30% of the general population. Atopic asthma is characterised by Th2 lymphocyte driven inflammation within the airways.

Exhaled nitric oxide (NO) mainly originates from the respiratory epithelium and is produced by inducible NO synthase (iNOS). In people with asthma, iNOS expression is upregulated by interleukin-4 and -13, both archetypal Th2 cytokines. Thus exhaled NO primarily signals Th2-lymphocyte-driven inflammation in the bronchial mucosa and consequently has potential utility in the diagnosis of asthma. The aim of this review is to assess the diagnostic test accuracy of FeNO as a relatively new diagnostic tool.

16.2 Review question: In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of fractional exhaled nitric oxide (FeNO) measures?

For full details see review protocol in Appendix C.

Table 42: Characteristics of review question

Component	Description
Population / Target condition	People with suspected asthma (presenting with respiratory symptoms). All ages, stratified into the following 3 different groups: • Children (1-<5 years old) • Children/young people (5-16 years old) • Adults (>16 years old)
Index test	Fractional exhaled nitric oxide (FeNO) with a cut-off threshold between 20-50ppb and a flow rate of 50ml/s or equivalent
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test. In children 1-<5 years, objective tests cannot be performed so the reference standard will be physician diagnosis based on recurrent and persistent wheezing.
Outcomes	• Diagnostic accuracy (Sensitivity and specificity) • FeNO levels

16.3 Clinical evidence

We searched for cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of FeNO to identify whether the condition is present (as indicated by the reference standard) in people under investigation for asthma.

Eighteen studies were included in the review[19,27,30,36,37,49,70,88,90,101,149,158,159,163,191,197,201] (see Table 43, Table 44 and Table 45). Evidence from these are summarised in the clinical evidence profile below

(Table 46). See also the study selection flow chart in Appendix D, and sensitivity / specificity forest plots in Appendix J, study evidence tables in Appendix G and exclusion list in Appendix K.

Eight of the total included studies were cross-sectional and compared the diagnostic accuracy of FeNO with a physician diagnosis plus objective test for asthma in people with suspected asthma[37,57,70,90,149,163,191,197]. Studies using a cut-off threshold for the reference standard objective test that differed from the protocol were not included in this review as there was enough evidence available from studies with the ideal reference standard. A variety of index test cut-off thresholds were used in the included studies, these are summarised in Table 43.

- Three of these studies were in adults only[57,90,149].
- Three of these studies were in a mixed population of adults and children/young people (data not separated), and were analysed in the adult strata due to the average age of the population (>16 years): CORDEIRO[37](age 7 and above), HEFFLER[70] (age 11 and above), VOUTILAINEN[191] (age 14 and above).
- Two studies were in children/young people alone: WOO[197] (aged 8 to 16 years) and Sivan[163].

One study was a cross-sectional study and compared the diagnostic accuracy of FeNO with a methacholine challenge test reference standard in adults only[30].

Nine studies were case-control studies and assessed FeNO levels in people with asthma or asthma vs. other respiratory diseases or healthy controls[19,27,36,49,88,101,158,159,201]. FeNO levels were also included from the cross-sectional studies, comparing those with a final diagnosis of asthma with those with symptoms but without a final diagnosis of asthma. In total eighteen studies were included for FeNO levels and median values are summaried for all, adults alone and children/young people alone in Table 43**.**

Summary of included studies

Table 43: Summary of studies included in the review: diagnostic accuracy of FeNO test versus physician diagnosis with objective test (could use >1 test)

Study	Presentation	Target condition	Index test	Reference standard	Comments
Cordeiro 2011[37]	New referrals to outpatient allergy clinic N=114 mixed population	Asthma	• FeNO: 27ppb • Flow rate 50ml/s • Niox-Flex device	History of typical respiratory symptoms and FEV1 improvement >12% and >200mL with salbutamol 400µg or PC20 histamine ≤8mg/mL	See below for FeNO levels from same study
Fukuhara 2011[57]	Outpatients referred to pulmonary medicine department. At least 1 of the subjective symptoms: recurrent cough, wheezing or dyspnoea (including chest tightness) N=61 Adults	Asthma	• FeNO: ≥40ppb • Flow rate 50ml/s • NA623N , Chest MI Japan device	At least 2 of the following: induced sputum eosinophilia, AHR, reversible airway obstruction. Airway reversibility defined as a change in FEV1 of 200ml or ≥12% after SABA or after 2-4 weeks treatment with ICS or bronchodilator. AHR defined as dose of MCh at which airway resistance began to rise (cut-off <12.5U). And other diseases ruled out using chest radiography, computed tomography and other lab tests.	
Heffler 2006[70]	Patients referred to allergy department for diagnostic evaluation of persistent rhinitis and asthma-like lower airways symptoms (cough, dyspnoea, chest tightness and wheezing) for 2 months N=48 mixed population	Asthma	• FeNO: 36ppb • Flow rate 50ml/s • Niox device	Typical symptoms and significant response to bronchodilator (≥12% improvement in FEV1 with salbutamol) or airway hyper-reactivity to methacholine (PD20 FEV1 ≤800µg)	See below for FeNO levels from same study
Kowal 2009[90]	Patients with chronic cough (at least 8 weeks) referred to asthma clinic for	Asthma	• FeNO: 40ppb flow	Significant diurnal changes in PEF or significant improvement of FEV1 with 200µg salbutamol	See below for FeNO levels from same

Study	Presentation	Target condition	Index test	Reference standard	Comments
	evaluation N=540 Adults		rate 50ml/s • NOA 280 Sievers device		study
Sato 2008[149]	Prolonged cough or wheezing >3 weeks attending Department of Pulmonary Medicine N=71 Adults	Asthma	• FeNO: 38.8ppb • Flow rate 50ml/s • Device from Kimoto, Japan (no further details given).	• Bronchial asthma: cough + wheeze >3 weeks, sputum eosinophilia + airway hyper-reactivity to methacholine or reversible airflow limitation (improvement in FEV1 of 200mL and ≥12%) with salbutamol or long-acting β2-agonist • Cough variant asthma: cough without wheezing >3 weeks, sputum eosinophilia + airway hyper- reactivity to methacholine or reversible airflow limitation	Asthma group = bronchial asthma + cough variant asthma; compared with non-asthma group = eosinophilic bronchitis without asthma, post-infectious cough, post-nasal drip, COPD, chronic bronchitis, cough with GERD or sino-bronchial syndrome. See below for FeNO levels from same study
SIVAN 2009[163]	Non-specific respiratory symptoms suggestive of asthma for at least 3 months, including cough, wheezing and shortness of breath with or without trials of treatment with bronchodilators and ICS. No mention of other respiratory defects. No BMI reported.	Asthma	• FeNO: 25ppb • Online single exhalation technique recommended by ERS/ATS guidelines	• Made by paediatric pulmonologist after 18 months follow-up. Based on history of 2 or more clinical exacerbations of wheezing documented by a physician; dyspnoea or cough relived by bronchodilators; documented variability in FEV1 ≥15% in response to bronchodilators at any time during the follow-up period; OR documented variability in FEV1 ≥15% over time with or without controller medications (ICS or montelukast). Results of provocation tests	

Study	Presentation	Target condition	Index test	Reference standard	Comments
	N=133 children			included when available.	
Voutilainen 2013[191]	Sedentary patients remitted to an allergy and asthma clinic because of respiratory symptoms (cough, dyspnoea or wheeze) N=87 (study also included a group of elite athletes N=87, not included in this review)	Asthma	• FeNO: 30ppb • Online single exhalation technique recommended by ATS • NIOX	• Based on general guidelines including typical symptoms and the objective confirmation of variable airway obstruction documented in hospital records. Such evidence was based either on BDR ≥12%, PEFv ≥20%, BDR of PEF ≥15%, exercise challenge test ≥15% or BHR MCh PD20 or hist PD15 ≤0.4mg	See below for FeNO levels from same study
Woo 2012[197]	Children presenting with non-specific respiratory symptoms e.g. cough, wheezing, shortness of breath, referred to paediatric outpatients for evaluation of asthma N=245 Children/young people	Asthma	• FeNO: 22ppb • flow rate 50ml/s • NIOX MINO device	History + reversible airflow obstruction (≥12% improvement in FEV1 with inhaled β-agonist) and/or airway hyper-reactivity (methacholine PC20 ≤8mg/mL)	Unclear if treatment naïve. See below for FeNO levels from same study

Table 44: Summary of studies included in the review : diagnostic accuracy of FeNO test versus individual comparator test

Study	Presentation	Target condition	Index test	Comparator tests	Comments
Chatkin 1999[30]	Chronic cough (>3 weeks) of unknown cause referred for diagnosis N=38 Adults	Asthma	• FeNO: 30ppb • Flow rate 45ml/s • Sievers 280	Positive to methacholine challenge (PC20 ≤8mg/mL)	See below for FeNO levels from same study

Study	Presentation	Target condition	Index test	Comparator tests	Comments
			device		

Table 45: Summary of studies included in the review : comparison of FeNO levels between people with asthma and non-asthma

Study	Presentation	Target condition	Index test	Non-asthma conditions	Comments
BERLYNE 2000[19]	Asthma (steroid naiive). Symptoms of wheeze, breathlessness or cough in past year plus MCT PC20 <8 mg/ml if the FEV1/VC >70%; or a post-BD FEV1 >15% if the FEV1/VC was <70%. Not received ICS in previous month.	Asthma	• FeNO; flow rate 45ml/s • Sievers 240 device	• eosinophilic bronchitis • healthy controls - atopic • healthy controls – nonatopic	
CARDINALE 2005[27]	mild intermittent asthma. History of symptoms, pulmonary function tests and response to inhaled beta-adrenergic agents according to international guidelines. History of at least 1 episode of asthma in past year and stable at time of study	Asthma	• FeNO; flow rate 50ml/s • NOA Tm280 Sievers device	• allergic rhinitis • healthy controls	
Chatkin 1999*[30]	Chronic cough (>3 weeks) of unknown cause referred for diagnosis	Asthma	• FeNO; flow rate 45ml/s • Sievers 280 device	Chronic cough (but methacholine negative)	
CIPRANDI 2013[36]	Allergic asthma. Paediatrician using validated criteria (GINA). Consistent symptoms and signs, lung function impairment and BDR. BDR FEV1>12%. Allergy by SPT for common aeroallergens	Asthma	• FeNO; flow rate 50ml/s • Sievers 280 device	• allergic rhinitis	
Cordeiro 2011*[37]	New referrals to outpatient allergy clinic	Asthma	• FeNO;	• Allergic rhinitis	Unclear if treatment

Study	Presentation	Target condition	Index test	Non-asthma conditions	Comments
			flow rate 50ml/s • Niox-Flex	• Allergic rhinitis, nonallergic rhinitis, eczema, urticarial, other (all together)	naive
Deykin 2002[49]	History of asthma, with either a 12% improvement in FEV1 after inhalation of a beta-agonist or a methacholine PC20 of 8 mg/ml or less	Asthma	• FeNO; flow rate 50ml/s. • NOA 280 Sievers device	• Healthy controls	Gives FeNO levels at other flow rates as well as 50ml/s
Fukuhara 2011*[57]	Outpatients referred to pulmonary medicine department. At least 1 of the subjective symptoms: recurrent cough, wheezing or dyspnoea (including chest tightness)	Asthma	• FeNO: flow rate 50ml/s • NA623N , Chest MI Japan device	Did not meet criteria for diagnosis of asthma but final diagnoses not reported	
Heffler 2006*[70]	Patients referred to allergy department for diagnostic evaluation of persistent rhinitis and asthma-like lower airways symptoms (cough, dyspnoea, chest tightness and wheezing) for 2 months	Asthma	• FeNO; flow rate 50ml/s. • Niox	Did not meet criteria for diagnosis of asthma but final diagnoses not reported	
Kostikas 2008*	Subjects with at least one asthma symptom on a screening questionnaire among students	Asthma	• FeNO; flow rate 50ml/s. • NIOX MINO device	• Allergic rhinitis • non-specific respiratory symptoms • healthy controls	Subjects had not presented to healthcare professionals. Only data from non-smokers included in FeNO levels analysis.
Kowal 2009*[90]	Patients with chronic cough (at least 8 weeks) referred to asthma clinic for evaluation	Asthma	• FeNO; flow rate 50ml/s.	• Rhinitis/sinusitis • gastroesophageal reflux • healthy controls	

Study	Presentation	Target condition	Index test	Non-asthma conditions	Comments
			• NOA 280 Sievers device		
Louhelainen 2008A[101]	Patients with newly-diagnosed asthma. (wheezing, prolonged cough and shortness of breath plus significant bronchial reversibility i.e. reduction in post-exercise PEF and/or FEV1 ≥15% or improvement in FEV1 ≥12% after bronchodilator or PD15 of histamine <0.4mg or ≥20% diurnal variation in PEF values and/or ≥15% improvement in PEF after bronchodilator at home)	Asthma	• FeNO; flow rate 50ml/s. • Niox device	• Healthy controls	Patients with asthma and healthy controls grouped by age (adult = 16-72 yrs; child = 7-14 yrs); COPD all adult
Sato 2008*[149]	Prolonged cough or wheezing >3 weeks attending Department of Pulmonary Medicine	Asthma	• FeNO; flow rate 50ml/s. • Device from Kimoto, Japan (no further details given).	• Eosinophilic bronchitis without asthma • Post-infectious cough, post-nasal drip, COPD, chronic bronchitis, cough with GERD or sino-bronchial syndrome	
Shimoda 2013[158]	Patients referred to asthma clinic with cough variant asthma or bronchial asthma	Asthma	• FeNO; flow rate 50ml/s • NOA 280 Sievers device	Healthy controls.	Each type of asthma compared separately with healthy controls.
Shome 2006[159]	Patients with newly-diagnosed asthma	Asthma	• FeNO; flow rate 50ml/s • CLD 88sp,	Healthy controls	Patients with asthma grouped by mild versus moderate/ severe disease

Study	Presentation	Target condition	Index test	Non-asthma conditions	Comments
			EcoPhysics device		
Voutilainen 2013*[191]	Sedentary patients remitted to an allergy and asthma clinic because of respiratory symptoms (cough, dyspnoea or wheeze)	Asthma	• FeNO: 30ppb • Online single exhalation technique recommended by ATS • NIOX	Non-asthma (not BDR ≥12%, PEFv ≥20%, BDR of PEF ≥15%, exercise challenge test ≥15% or BHR MCh PD20 or hist PD15 ≤0.4mg); final diagnosis not stated	
Woo 2012*[197]	Children presenting with non-specific respiratory symptoms e.g. cough, wheezing, shortness of breath, referred to paediatric outpatients for evaluation of asthma	Asthma	• FeNO; flow rate 50ml/s • NIOX MINO device	Non-asthma (not airway hyper-reactivity (cut off for methacholine PC20 of 8mg/mL) or reversible airflow obstruction (12% improvement in FEV1 with inhaled β-agonist); final diagnoses not stated	Unclear if treatment naïve. Asthma and non-asthma groups also sub-divided by atopic vs. non-atopic
ZIETKOWSKI 2006A[201]	Steroid-naïve patients with mild to moderate asthma (56 allergic and 45 nonallergic). Asthma Dx according to GINA	Asthma	FeNO; flow rate 50ml/s	Healthy controls	

Table 46: Clinical evidence profile: Diagnostic Test Accuracy for FeNO

Index Test (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity % (range/median 95% CI)	Specificity % (range/median 95% CI)	Area Under Curve (range)	Quality

Index Test (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity % (range/median 95% CI)	Specificity % (range/median 95% CI)	Area Under Curve (range)	Quality
FeNO vs. Physician Dx with objective test: Adults										
FeNO >27ppb	1	114	No risk of bias[a]	No serious inconsistency	Serious indirectness[c][d][e]	n/a[f]	78.6	91.7	-	MODERATE
FeNO >30ppb	1	87	Serious risk of bias[a]	No serious inconsistency	Serious indirectness[c]	n/a[f]	43.0	89.0	0.738	LOW
FeNO >36ppb	1	48	No risk of bias[a]	No serious inconsistency	Serious indirectness[c]	n/a[f]	77.8	60.0	-	MODERATE
FeNO >38.8ppb	1	71	No risk of bias[a]	No serious inconsistency	Serious indirectness[e]	n/a[f]	79.2	91.3	-	MODERATE
FeNO >40ppb	2	601	Very serious risk of bias[a]	No serious inconsistency	Serious indirectness[d][e]	n/a[f]	Range 78.6 - 88.3	Range 82.6 - 89.5	-	VERY LOW
FeNO vs. Physician Dx with objective test: Children/young people										
FeNO >22ppb	1	245	No risk of bias[a]	No serious inconsistency	No serious indirectness	n/a[f]	56.9	87.2	-	HIGH
FeNO 25ppb	1	113	Very serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness[c]	n/a[f]	75	89		LOW
FeNO vs. methacholine ≤8mg/mL										
12 FeNO >30ppb	1	38	No risk of bias[a]	No serious inconsistency	No serious indirectness	n/a[f]	75.0	86.7	-	HIGH

(a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.
(b) Inconsistency was assessed by inspection of the sensitivity / specificity values in each of the studies in relation to a high (≥75%), moderate (50-74%) or low (<50%) effect estimate. Inconsistency was considered high if the point estimates from individual studies were within two areas, and very high if they were in 3 areas or both high and low areas.
c) Mixed population of adults and children/young people
d) Unclear if treated at baseline in one study
e) Smokers included in the study
f) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed.

16.4 Economic evidence

Published literature

Three economic evaluations relating to this review question were identified but were excluded due to a combination of limited applicability and methodological limitations.[18,66,140] These are summarised in Appendix L, with reasons for exclusion given. All these three studies assessed individual tests in isolation, while the GC was interested in knowing the cost-effectiveness of diagnostic strategies including more than one test.

New cost-effectiveness analysis

An original health economic model was built for adults to assess the cost-effectiveness of several diagnostic pathways, which included FeNO. Full details of the model can be found in Appendix M. A summary of the model can be found in section 18.4 .

Unit costs of performing FeNO on children

As an economic model was not feasible for children the GC considred the unit cost of performing FeNO to evaluate its cost-effectiveness as part of a pathway for diagnosing asthma. This unit cost is presented below (Table 47).

Table 47: Cost of FeNO

Item	Quantity[(a)]	Unit cost	Total Cost (quantity*unit cost)	Source of unit cost
Time taken to conduct test with GP practice nurse	5-10 minutes	£0.73 per minute	£3.65-£7.30	PSSRU[40]
Cost of FeNO equipment per use	1	£6.36 per use	£6.36	Harnan et al[66]
Total			**£10.01-£13.66**	

(a) Based on GC opinion.

The GC also acknowledged the annual cost of drugs for the management of asthma in children. Preventing these costs from occurring for children without asthma would be a large benefit derived from a diagnostic strategy with a high specificity. This cost was estimated to be £201 from a study by Main et al.[102,102]

16.5 Evidence statements

Clinical

- One study with 114 adults showed that FeNO (cut-off 27ppb) has a sensitivity of78.6% and a corresponding specificity of 91.7% for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY).
- One study with 87 adults showed that FeNO (cut-off 30ppb) has a sensitivity of43.0% and a corresponding specificity of 89.0% for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY).

- One study with 48 adults showed that FeNO (cut-off 36ppb) has a sensitivity of77.8% and a corresponding specificity of 60.0% for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY).
- One study with 71 adults showed that FeNO (cut-off 38.8ppb) has a sensitivity of79.2% and a corresponding specificity of 91.3% for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY).
- Two studies with 601 adults showed that FeNO (cut-off 40ppb) has a sensitivity range of 78.6-88.3% and a corresponding specificity range of 82.6-89.5% for diagnosing asthma in people presenting with respiratory signs and symptoms. (VERY LOW QUALITY).
- One study with 245 children showed that FeNO (cut-off 22ppb) has a sensitivity of 56.9% and a corresponding specificity of 87.2% for diagnosing asthma in people presenting with respiratory signs and symptoms. (HIGH QUALITY)
- One study with 113 children showed that FeNO (cut-off 25ppb) has a sensitivity of 75% and a corresponding specificity of 89% for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 38 adults showed that FeNO (cut-off30ppb) has a sensitivity of 75.0% and a corresponding specificity of 86.7% for a positive methacholine challenge test in people presenting with respiratory signs and symptoms. (HIGH QUALITY)

Economic

- No relevant economic evaluations were identified
- An original health economic model found that FeNO (together with spirometry, BDR and when there are conflicting results PEFv and MCT) was part of the most cost-effective diagnostic pathway used to diagnose asthma in adults aged 16 and over (see diagnostic algorithms in section 4.1). This evidence is directly applicable with minor limitations.

16.6 Recommendations and link to evidence

Recommendations	**20. Offer a FeNO test to adults (aged 17 and over) if a diagnosis of asthma is being considered. Regard a FeNO level of 40 parts per billion (ppb) or more as a positive test.** **21. Consider a FeNO test in children and young people (aged 5 to 16)[f] if there is diagnostic uncertainty after initial assessment and they have either:** • **normal spirometry or** • **obstructive spirometry with a negative bronchodilator reversibility (BDR) test.** **Regard a FeNO level of 35 ppb or more as a positive test.** **22. Be aware that a person's current smoking status can lower FeNO levels both acutely and cumulatively. However, a high level remains useful in supporting a diagnosis of asthma.**
Relative values of different diagnostic measures and	The GC was interested in the diagnostic test accuracy of fractional exhaled nitric oxide (FeNO) measures using a flow rate of 50ml/s for the FeNO test. There is evidence of significant misdiagnosis of asthma, with both false positive and false

[f] Children at the lower end of the age range may not be able to do the FeNO test adequately. In these cases, apply the principles in recommendation 13.

outcomes	negative errors (see introduction to the guideline). The GC was therefore interested in both the sensitivity and specificity of diagnostic tests and prioritised studies which reported each of these outcomes for FeNO measurements. Sensitivity and specificity values will depend on the chosen cut-off point. It was noted that ATS guidelines[52] have classified low FeNO levels as <25ppb in adults and <20ppb in children, these levels indicating normal airways, and high FeNO levels as >50ppb in adults and >35ppb in children indicating eosinophilic inflammation. The ATS guidelines specify that FeNO levels between these ranges should be interpreted with caution. For the diagnosis of asthma with FeNO as the index test, the GC included studies with a cut-off threshold for diagnosis between 20-50ppb. Studies using a cut-off threshold for the reference standard objective test that differed from the protocol were not included in this review, as there was enough evidence available from studies with the ideal reference standard. The GC noted that the number of people diagnosed with asthma will vary with the cut-off threshold for the reference standard objective test, and therefore only the optimal threshold should be used if this can be determined. The GC considered evidence from a summary ROC curve when assessing the heterogeneity in the results and any threshold effect. The summary ROC curve can be found in Appendix J (sub-section 10.1.1; page 457). They also viewed evidence from case-control and cross-sectional studies comparing FeNO levels in patients with asthma vs. other respiratory conditions or healthy controls when choosing an appropriate cut-off value for FeNO as a diagnostic test.
Trade-off between clinical benefits and harms	The FeNO test can be performed in around 10 minutes and can be performed within primary care. The sensitivity and specificity of FeNO in adults was high, with the exception of one study with a moderate specificity (cut-off >36ppb[70]) and one study with a low sensitivity (cut-off >30ppb[191]). In children/young people, FeNO had a moderate sensitivity in one study[197] and high sensitivity in the other[163] with high specificity in both.
Economic considerations	No economic evaluations were included which assessed the use of FeNO as part of a diagnostic pathway. Therefore, an original health economic model was built to assess the cost-effectiveness of different diagnostic pathways for asthma. The model assessed the additional costs of tests against cost-savings from unnecessary asthma medication, and the increased health outcomes from providing correct treatment. FeNO had the best diagnostic accuracy out of all the tests that can be conducted in primary care. Due to this the GC agreed that FeNO should appear in every diagnostic pathway as it would be pivotal in making a diagnosis. In the model, strategies that gave FeNO to all patients in the pathway dominated the strategies that did not. Therefore FeNO is a highly cost-effective component of the diagnostic algorithm. The most cost-effective strategy involved using spirometry, bronchodilator reversibility (BDR), FeNO, peak expiratory flow variability (PEFv) and a methacholine challenge test (MCT) in selected individuals to diagnose asthma. In this strategy, everybody with symptoms of asthma would undergo a spirometry test and a FeNO test, those who had an obstructive spirometry would also receive a BDR test. Only those who had non-obstructive spirometry and conflicting FeNO and PEFv test results would receive a MCT. Adopting this strategy dominated all other strategies apart from those which performed challenge tests at more points in the pathway. The ICERs of adopting these further stratgies were above £20,000 per QALY gained. In a sensitivity analysis when the cost of FeNO increased above £93 none of the diagnostic strategies were cost-effective at a £20,000 per QALY threshold and therefore current practice became the most cost-effective strategy. If the cost of

	FeNO was £93 then the cost-effective ranking of strategies remained unchanged. For the marginal cost of FeNO to rise to £93 the machine would only be used approximately 28 times in a 5 year time span. The GC noted that even for small GP practices under the most conservative assumptions of the number of new diagnoses made each year, this level of use would still be attainable. In children the GC recognised FeNO as having considerable value in a diagnostic pathway as it can be performed at a low cost and the clinical evidence showed it had a high specificity. Apart from children who produce a positive bronchodilator reversibility result, the GC considered that FeNO measurements would add value to all other points in the pathway. The GC noted that as not all children would receive a FeNO test the average unit cost of including FeNO in the proposed diagnostic algorithm would be less.
Quality of evidence	The evidence ranged from very low to high quality by GRADE criteria. The limitations of the studies included the variable inclusion criteria, making directness to the clinical question variable, and the risk of bias as assessed using the QUADAS II checklist. • In children aged <5 years, no studies were identified. • In children and young people aged 5-16 years, there were two included studies[163, 197] with cut-offs of 22ppb and 25ppb, using the reference standard (physician diagnosis with objective test). FeNO showed moderate sensitivity in one study[197], high sensitivity in the other[163] and high specificity in both. The evidence quality ranged from low to high. • In adults, there were five included studies. At cut-off thresholds ranging from 27-40ppb, FeNO had a high sensitivity and a high specificity, with the exception of two studies with a low sensitivity at a cut-off of 30ppb[191] and a moderate specificity at a cut-off of 36ppb[70]. The quality of the evidence was downgraded if there was a mixed population of adults and children/young people with no breakdown of the results, if there was uncertainty about medications at baseline or if the population included smokers. One study (Cordeiro 2011) had an indirect population with general allergic symptoms rather than respiratory symptoms. The quality of the evidence ranged from very low to high quality. Additionally, in adults, one study comparing FeNO with methacholine challenge test as a proxy for asthma showed high sensitivity and high specificity at a cut-off value of 30ppb. The evidence for FeNO levels in asthma, other respiratory conditions and healthy controls is limited by the lack of numbers for some populations. The economic evidence was assessed as directly applicable with minor limitations.
Other considerations	The GC discussed variables which may affect the FeNO test. A standard flow-rate of 50ml/s should be used and this flow-rate is independent of body size. FeNO levels can be altered by corticosteroids, smoking or previous smoking history and diet. The GC excluded studies in which more than 50% of the population were taking corticosteroids, or if the smoking history of the population was unclear. FeNO measures are independent of peripheral blood eosinophil levels. There is some uncertainty about the best FeNO cut-off threshold to distinguish between asthma and non-asthma. When considering a cut-off threshold, the GC discussed the sensitivity and specificity data, and evidence for the FeNO levels in patients with asthma, other respiratory conditions and healthy controls. A cut-off of greater than or equal to 40ppb was chosen in adults as this cut-off value has a high sensitivity and specificity and is above the range of FeNO levels observed in a population of mixed respiratory symptoms without a diagnosis of asthma. FeNO levels were generally lower in children with asthma and healthy controls. The GC

discussed the high prevalence of rhinitis in children and that the cut-off threshold for the diagnosis of asthma should ideally lie above the range of FeNO values in children with rhinitis. The available studies suggested a cut-off point in the range 20-25ppb but specificity at this level appears higher than sensitivity. A slightly higher cut-off would reduce the number of false positive tests. The GC discussed making a recommendation with a grey area, specifying clear-cut levels for rejecting the diagnosis of asthma set at 20ppb or below and for supporting the diagnosis of asthma set at 35ppb or higher. However, this was rejected as being too complicated and potentially confusing, particularly bearing in mind that FeNO is an unfamiliar test to many current practitioners in the UK. A cut-off of greater than or equal to 35ppb was therefore chosen in children after considering the issue of allergic rhinitis and the limitations of cut-off values lower than these, as more false positives occur.

The GC discussed whether it was possible to recommend different cut-off levels for smokers. Most of the studies included in the review specifically excluded smokers. The GC also noted that there is not clear advice within the manufacturer's instructions on the magnitude of the effect smoking has on FeNO levels. The GC agreed that there was not enough evidence to recommend a cut-off that should be used for smokers. However, there was agreement that smoking lowers FeNO levels and that the smoking status of an individual should be considered when interpreting FeNO results[106,111,188].

The GC considered the placement of FeNO in a diagnostic pathway, taking into account the diagnostic accuracy of the test and the practicality of performing the test. The GC noted that FeNO can be performed fairly easily in primary care and is a simple test for the patient to complete. The GC also noted the high diagnostic accuracy of FeNO relative to all other tests that can be performed in primary care. For these reasons, the GC agreed it would be appropriate to give all patients who had undergone a spirometry test and (where relevant) bronchodilator reversibility (BDR) testing, a FeNO test.

The GC also considered the additional benefit that FeNO offers of identifying patients likely to be steroid-responsive.

For adults aged 17 years and over with obstructive spirometry, reversible airways (positive BDR), no variability in peak flow readings and a very low FeNO count of less than 25 ppb, the decision to either consider an alternative diagnosis or refer for a specialist opinion will differ on a case-by-case basis.

For instance, a 55-year-old with a 30 pack year smoking history, who demonstrates obstructive spirometry with reversibility, but a FeNO of 10ppb and no peak flow variability, is likely to have COPD. On the other hand, a 20-year-old who has never smoked, with obstructive spirometry with reversibility, a FeNO of 24ppb and no peak flow variability, may well need a specialist opinion for further investigation, rather than considering an alternative diagnosis.

In children aged 5-16 years, evidence was only available from fewer studies than in adults. The GC decided that the strength of the recommendation in children should refect the fact that the recommendation was based on limited evidence. The GC also discussed that some children at the lower end of this age range may find it difficult to perform the FeNO test adequately. In these cases, the principles of the recommendation 27 should apply until the child is old enough to perform the FeNO test adequately.

Details of the feasibility project are available in appendix Q. The aim of the feasibility project was to assess the impact and feasibility of adopting the diagnostic tests (e.g.

	FeNO and spirometry) recommended, into primary care. **General concerns** Firstly it was recognised that children may not be able to perform some of the tests. An additional recommendation has now been made which informs the clinician what to do should this problem arise. The GC agreed that objective testing was imperative before a diagnosis of asthma could be made however, before objective testing is possible, symptoms should be treated and monitored. Secondly the feasibility report identified that in the current format the diagnostic algorithms could be difficult to follow in some places. Therefore the GC agreed a new format should be designed that would simplify the algorithm and make them easier to interpret. **Concerns specific to FeNO** From the feasibility project results, the main barrier to implementation was cited as the cost of the device and consumables rather than the practicaility or accuracy of the test. The project cited positive feedback for the FeNO machine with very good patient compliance. All sites agreed that the device was easy to use and training was not lengthy (less than for spirometry). Moreover, fewer patients were unable to complete FeNO measurement than spirometry (5 vs 9). Moving forward, the GC considered that the use of diagnostic hubs could help alleviate the issue of cost as the cost of the machine would only need to be incurred once. Likewise there are economies of scale that arise through bulk purchasing of the consumables. It was considered that this would reduce costs and improve the practicality of implementing the algorithm. A recommendation was developed, aimed at clinical commissioners, to consider establishing asthma diagnostic hubs to achieve economies of scale in implementing the diagnostic algorithms (see section 21.2). Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. One study was identified which the GC agreed did not suggest a change in the recommendations was warranted.

17 Diagnosis: Peripheral blood eosinophil count

17.1 Introduction

Eosinophils are a form of white blood cells produced by the bone marrow. Their exact role in health has yet to be determined, but it is believed that they play a role in fighting parasitic infections and primarily reside within the lining of the gut.

Biopsies taken from the lungs of people with asthma have frequently demonstrated increased numbers of eosinophils and the number of eosinophils is also often increased in sputum samples taken from people with asthma. Measurement of sputum eosinophil numbers have been used to aid the diagnosis and management of asthma. However, this is a time consuming procedure, which is only performed in a specialist setting. Eosinophils travel from the bone marrow to the lung, it is therefore logical to investigate whether measurement of blood eosinophils is a useful tool for asthma diagnosis.

17.2 Review question: In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of eosinophil blood count measures?

For full details see review protocol in Appendix C.

Table 48: PICO characteristics of review question

Population	People with suspected asthma (presenting with respiratory symptoms). All ages, stratified into the following 3 different groups: • Children (1-<5 years old) • Children/young people (5-16 years old) • Adults (>16 years old)
Index test	Peripheral blood eosinophil count (may be part of FBC)
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test In children 1-<5 years, objective tests cannot be performed so the reference standard will be physician diagnosis based on recurrent and persistent wheezing.
Outcomes	• Diagnostic accuracy (sensitivity, specificity) • Eosinophil levels

17.3 Clinical evidence

We searched for cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of peripheral blood eosinophil counts to identify whether the condition is present (as indicated by the reference standard) in people under investigation for asthma.

Twenty studies were included in the review[10,65,73,82,89,91,92,108,117,129,137,147,157,161,162,175,176,179,189,201]. Evidence from these are summarised in Table 49 and the clinical evidence profile below (Table 50). See also the study selection flow chart in Appendix D, forest plots in Appendix J, study evidence tables in Appendix G and exclusion list in Appendix K.

Four of the studies[89,137,157,175] were cross-sectional studies, and looked at the diagnostic accuracy of peripheral blood eosinophil counts in people with suspected asthma. Three of these studies had a reference standard of physician diagnosis plus and objective test[89,137,175]. The remaining study[157] had

a physician diagnosis of atopic asthma without an objective test, but was included due to limited evidence with the ideal reference standard. This study was downgraded for indirectness. Two studies were in adults[137,175] and the other two studies were in children[89,157] and evidence was analysed in the 5-16 year strata. No evidence was identified in children <5 years.

The remaining sixteen studies[10,65,73,82,91,92,108,117,129,147,161,162,176,179,189,201] were either case-series (of asthma only patients, no comparison), or were case-control studies, which compared the levels of peripheral blood eosinophils counts in people who had already been diagnosed with asthma vs. healthy controls and/or other respiratory symptoms or conditions. Although one of these studies (Backer 2002[11]) was a cross-sectional study but only reported blood eosinophil counts (and not sensitivity and specificity). The studies are summarised in Table 49. Evidence from these studies is summarised in Appendix G. Studies measuring eosinophil count simply stating 'cells/mm^3' have not been included in the pooled summary results.

Table 49: Summary of included studies

Study	N	Index test*/reference standard	Index test cut-off for positivity	Population	Age
Adults: PBE vs. reference standard					
POPOVIC 2002[137]	195 (N=141 asthma)	PBE vs. Physician Dx + objective test (BDR)	Not reported	Suspected asthma (dyspnoea)	Adults (mean 39 yrs)
TILEMANN 2011[175]	210 (N=86 asthma)	PBE vs. BDR	≥4.15%	Suspected obstructive airways disease.	Adults (mean 49 yrs)
Children 5-16 years: PBE vs. reference standard					
KOTANIEMI 2002[89]	82 (N=33 asthma)	PBE vs. Physician Dx + objective test (exercise challenge test)	≥0.45 x 10^9/l.	Suspected asthma (wheeze)	Children (mean 7.2 yrs)
SHIELDS 1999[157]	137 (N=60 asthma)	PBE vs. Physician Dx	>4% and >8%	History of wheezing	Children (range 1-15 yrs)
Studies reporting PBE counts					
BACKER 2002[10,11]	624 (N=103 asthma)	PBE counts only	N/A	General population sample	Adults (range 19-29 yrs)
HALVANI 2012[65]	98 (N=61 asthma)	PBE counts only	N/A	Asthma and healthy controls	Adults (mean 38 yrs)
HUNTER 2002[73]	110 (N=89 asthma)	PBE counts only	N/A	Asthma and healthy controls	Mainly adults (mean 39 yrs)
KHAKZAD 2009[82]	62 (N=50 asthma)	PBE counts only	N/A	Asthma and healthy controls	Adults (mean 40 yrs)
KROEGEL 1998[91]	56 (N=14	PBE counts only	N/A	Asthma,	Adults

Study	N	Index test*/reference standard	Index test cut-off for positivity	Population	Age
	asthma)			suspected asthma, COPD and healthy controls	(mean 55 yrs)
LABBE 2001[92]	143 (N=88 asthma)	PBE counts only	N/A	Asthma and healthy controls	Children (mean 7 yrs)
METSO 2000[108]	190 (N=160 asthma)	PBE counts only	N/A	Asthma and healthy controls	Mainly adults (range 16-60)
NORDLUND 2012[117]	39	PBE counts only	N/A	Asthma	Children (mean 14 yrs)
PIIPPOSAVOLAINEN 2007[129]	83	PBE counts only	N/A	Wheezing/ bronchiolitis	Children (<2 yrs, mean not reported)
RYTILA 2000[147]	68 (N=25 asthma)	PBE counts only	N/A	Asthma and healthy controls	Mainly adults (mean 38 yrs)
SILVESTRI 2001A[160,161]	112	PBE counts only	N/A	Asthma	Children (mean 10.6 yrs)
SILVESTRI 2003[160,162]	92	PBE counts only	N/A	Asthma	Children (mean 10.7 yrs)
TOMASIAKLOZOWSKA 2012[176]	110 (N=91 asthma)	PBE counts only	N/A	Asthma and healthy controls	Adults (mean 38 yrs)
TUCHINDA 1987[179]	1000	PBE counts only	N/A	Asthma	Children <13 years (mean not reported)
VILA-INDURAIN 1999[189]	57 (N=36 asthma)	PBE counts only	N/A	Asthma and healthy controls	Children (8-18 yrs, mean not reported)
ZIETKOWSKI 2006A[201]	140 (N=101 asthma)	PBE counts only	N/A	Asthma and healthy controls	Adults (mean 35 yrs)

PBE = peripheral blood eosinophil count

Table 50: Clinical evidence profile: PBE count vs. Physician Dx of asthma:

Blood eosinophil count	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Median Sensitivity, or PBE counts (range)	Median Specificity or PBE counts (range)[e]	Area Under Curve (range)	Quality
ADULTS										
≥4.15%	1	210	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[c]	0.36	0.83	-	MODERATE
cut-off not reported	1	195	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness [e]	N/A[c]	0.15	0.39	-	LOW
CHILDREN 5-16 years old										
>4%	1	137	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness [d]	N/A[c]	0.62	0.67	-	LOW
>8%	1	137	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness [d]	N/A[c]	0.38	0.93	-	LOW
≥0.45 x 10^9/l	1	82	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	N/A[c]	0.55	0.84	-	MODERATE
CHILDREN 1-<5 years old										
No evidence identified	0									

a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.

b) Inconsistency was assessed by inspection of the sensitivity / specificity values in each of the studies in relation to a high (≥75%), moderate (50-74%) or low (<50%) effect estimate. Inconsistency was considered high if the point estimates from individual studies were within two areas, and very high if they were in 3 areas or both high and low areas.

c) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed.

d) Reference standard did not include an objective test.

e) Index test cut-off not reported.

17.4 Economic evidence

Published literature

No relevant economic evaluations were identified. See also the economic article selection flow chart in Appendix E.

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs were provided to aid consideration of cost-effectiveness. The GC considered this unit cost alongside the diagnostic pathway evaluated in the economic model. If this test was part of a diagnostic pathway then the incremental cost of performing this test may not include GP referral time as this could be incurred at the beginning of the pathway for all relevant tests.

Table 51: Unit costs for eosinophil blood count

Item	Unit cost	Quantity	Sub total	Source
Lab costs associated with eosinophils	£10.33	1	£10.33	GC estimate
Nurse time	£0.75 per minute	10 minutes	£7.50	PSSRU[40,40]
GP time	Average cost of GP appointment (11.7 min) = £36	2 GP appointments (1 for referral and 1 to interpret and discuss the results)	£72	PSSRU[40,40]
TOTAL			£89.83	

17.5 Evidence statements

Clinical

- One study with 210 adults showed that blood eosinophil count (cut-off ≥4.15%) has a sensitivity of 0.36 and a corresponding specificity of 0.83 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 195 adults showed that blood eosinophil count (no cut-off reported) has a sensitivity of 0.15 and a corresponding specificity of 0.39 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 137 children and young people showed that blood eosinophil count (cut-off >4%) has a sensitivity of 0.62 and a corresponding specificity of 0.67 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 137 children and young people showed that blood eosinophil count (cut-off >8%) has a sensitivity of 0.38 and a corresponding specificity of 0.93 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 82 children and young people showed that blood eosinophil count (cut-off ≥0.45 x 10^9/l) has a sensitivity of 0.55 and a corresponding specificity of 0.84 for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- No evidence was identified in children <5 years.

Economic

- No relevant economic evaluations were identified.

17.6 Recommendations and link to evidence

Recommendations	**23. Do not offer a peripheral blood eosinophil count as a diagnostic test for asthma.**
Research recommendations	**1. What is the acceptability and diagnostic accuracy of objective tests that could be used to comprise a diagnostic pathway for asthma in children aged 5-16 years old (for example, exercise challenge, direct bronchial challenge with histamine or methacholine, indirect bronchial challenge with mannitol and peripheral blood eosinophil count)?**
Relative values of different outcomes	The GC was interested in the diagnostic test accuracy of blood eosinophil counts in the diagnosis of asthma. There is evidence of significant misdiagnosis of asthma, with both false positive and false negative errors (see introduction to the guideline). The GC was therefore interested in both the sensitivity and specificity of diagnostic tests and prioritised studies which reported each of these outcomes. Sensitivity and specificity values will depend on the chosen cut-off point, but in most studies (despite often using different cut-offs), the cut-off chosen appeared to give better specificity than sensitivity. Eosinophil levels from population studies including people with asthma and people without asthma were also considered if the studies reported levels separately in those with asthma. These studies show that an asthma diagnosis is more likely with increasing blood eosinophil level, but the overlap between asthma and non-asthma is considerable.
Trade off between clinical benefits and harms	The benefit of measuring eosinophil levels was considered purely in terms of the value as a diagnostic test, and therefore it is dependent on the sensitivity and specificity. In adults, a cut-off threshold of ≥4.15% resulted in a low sensitivity but a high specificity. In children, the sensitivity at all thresholds was low or moderate, and the specificity was moderate or high. Over-reliance on an eosinophil level as a test of asthma would be a potential harm, but, apart from this, the only disadvantage is that a blood sample has to be provided. In most adults this does not present a problem, although a few people have needle phobia. Obtaining a blood sample poses a greater problem in children, but although the procedure can be distressing the use of anaesthetic patches should ameliorate this. There is a delay between taking blood and obtaining the eosinophil result which constitutes a slight disadvantage in comparison to some other potential tests for asthma, which can provide immediate answers. The GC also noted that some GP practices do not have a phlebotomy service and their patients would have to make a separate visit to a hospital for the blood sample to be taken.
Economic considerations	No relevant economic evaluations were identified for blood eosinophil count. The GC considered the unit costs of these tests, as well as the downstream implications of correct and incorrect diagnoses. The GC did not think that blood eosinophil count would be cost-effective as first-line diagnostic tests, as the clinical evidence does not show that they offer sufficient diagnostic accuracy. An original health economic model was built to assess the cost-effectiveness of different diagnostic pathways for asthma. The GC agreed there were no endpoints of the diagnostic pathway where an eosinophil test would be of benefit. To be of

	benefit the test would need to lead to change in the diagnostic decision. The GC concurred that the results from eosinophil tests would not be enough to overturn any of the decisions made in the diagnostic pathway and therefore they were not considered a cost-effective use of resources.
Quality of evidence	In children aged <5 years there was little evidence. There were no studies giving sensitivity or specificity values in this age group, and only one study (PIIPPOSAVOLAINEN 2007) which provided eosinophil levels in people with asthma. In children aged 5-16 years, two studies assessed the diagnostic accuracy of PBE counts for asthma diagnosis. One study had an ideal reference standard of physician diagnosis plus an objective test. The other study had a physician diagnosis of atopic asthma without an objective test, but was included due to limited evidence with the ideal reference standard. The evidence was of low to moderate quality. Other studies of eosinophil levels in different population groups of similar age ranges (groups with asthma vs. groups without asthma), showed that peripheral blood eosinophil (PBE) counts were generally higher in: asthma vs. non-asthma or healthy controls; allergic asthma vs. non-allergic asthma; asthma with normal FEV_1 vs. those with low FEV_1. In adults, two studies assessed the diagnostic accuracy of PBE counts for asthma diagnosis using the ideal reference standard of physician diagnosis plus an objective test. The evidence was of low to moderate quality. Other studies of eosinophil levels in different population groups (groups with asthma vs. groups without asthma), showed that PBE counts were generally higher in: asthma vs. non-asthma or healthy controls; allergic asthma vs. non-allergic asthma.
Other considerations	The GC noted that eosinophil counts are known to be elevated in conditions other than asthma and therefore would be unreliable as a diagnostic test for asthma in isolation. The GC considered the use of eosinophil blood tests in a diagnostic algorithm taking into account practicality of performing the test and diagnostic accuracy. They noted that there were no areas of diagnostic uncertainty in the proposed algorithm where a blood eosinophil count could be of significant benefit, since the results from the test would not be conclusive enough to reverse any diagnostic decisions made on the basis of other tests. The GC made a high-priority research recommendation to investigate the acceptability and performance characteristics of objective tests, including peripheral blood eosinophil tests, which could be used to comprise a diagnostic pathway for asthma in children aged 5-16 years old. Further details on the high-priority research recommendation made can be found in appendix N, along with the full list of research recommendations made. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

18 Diagnosis: Direct bronchial challenge test with histamine and methacholine

18.1 Introduction

Hyper-reactivity of the airways to non-specific stimuli is a key feature of asthma. Bronchial hyper-reactivity (BHR) can be measured in a number of different ways. Inhalation of the bronchoconstrictors histamine and methacholine can be used to measure BHR. In both histamine and methacholine challenge tests, incremental doses of one or the other agent are administered by inhalation until there is a fall in the person's FEV1 of at least 20% from the baseline value. The result is expressed as the PC20 (provocation concentration) or PD20 (provocation dose) of bronchoconstrictor required to produce a 20% fall in FEV1. However, the diagnostic test accuracy of bronchial challenge tests with histamine or methacholine to diagnose asthma is currently uncertain.

18.2 Review question: In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of airway hyper-reactivity (non-specific bronchial challenge) with histamine and methacholine?

For full details see review protocol in Appendix C.

Table 52: Characteristics of review question

Population	People with suspected asthma (presenting with respiratory symptoms). Ages stratified into the following 2 different groups: • Children/young people (5-16 years old) • Adults (>16 years old)
Index test	• Histamine PC20 or PD20 • Methacholine PC20 or PD20 Cut-off threshold of 8mg/ml or a cut-off threshold identified from a ROC curve
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test
Statistical measures	• Diagnostic accuracy (sensitivity and specificity)
Study design	Cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses)

18.3 Clinical evidence

We searched for cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of airway hyper-reactivity (non-specific bronchial challenge) with histamine and methacholine to identify whether the condition is present (as indicated by the reference standard) in people under investigation for asthma.

The GC considered evidence from studies reporting a cut-off value of ≤8 mg/ml. The GC excluded studies with a cut-off value of a fall in FEV1 greater than or equal to 20% (PC20) at a concentration of ≤16mg/ml from the review (with one exception noted below). Values at the higher end of this range are likely to occur in some healthy individuals and the GC was concerned about a high number of

false positives at this cut-off value. Six studies were included in the review[7,69,87,90,116,137] (see Table 53 and Table 54). Evidence from these are summarised in the clinical evidence profile below (Table 55 and Table 56). See also clinical article selection in Appendix D, sensitivity and specificity forest plots in Appendix J, clinical evidence tables in Appendix G and excluded clinical studies in Appendix K.

Five studies[7,69,90,116,137] were cross-sectional studies, and looked at the diagnostic accuracy of the methacholine or histamine challenge test in patients with suspected asthma or asthma symptoms. The reference standard was physician's diagnosis of asthma with an objective test. A variety of objective tests were used for the reference standard (see Table 53). Four studies were in adults and one study in children aged 5-16 years. In children aged 5-16 years, no evidence was available using the ideal cut-off of 8mg/ml. Therefore, evidence was included from one study using a cut-off value for a positive test of 16mg/ml[7]. This study also reported data in adults, children and young people combined, however it was not included in the review in adults here, as other studies with the preferred threshold of 8mg/ml were available in the adult population.

The remaining study[87]was a case-series, and looked at the diagnostic accuracy of methacholine or histamine challenge tests vs. other diagnostic tests, in people who had already been diagnosed with asthma. In this case, the index test was taken to be histamine challenge and the other test as the reference standard.

Summary of included studies

Table 53: Summary of studies included in the review: Adults

Study	Population	Index test & cut-off	Reference standard	Comparator test & cut-off
HEDMAN 1998[69]	N=230 Adults Referred with symptoms of cough, dyspnoea or wheezing of unknown cause	MCT PD20 6900µg	Physician Dx with objective test - according to guidelines of the American Thoracic Society (documented variation in FEV1 or PEF of ≥15% after medication, or repeatedly a ≥20% spontaneous PEFv during a period of 2 weeks AND a ≥15% decrease in FEV1 after a specific allergen provocation or during an exercise test)	None
KOSKELA 2003[87]	N=42 Adults All people with asthma - recent Dx of asthma based on Physician Dx and objective test	HCT PD15 1mg and 0.4mg	n/a – all people with asthma so use comparator test	MANNITOL PD15 <635mg
KOWAL 2009[90]	N=540 Adults (18-45 yrs) Referred to asthma clinic with chronic cough	HCT PC20 8mg/ml	Clinical follow-up 6 months (with objective test) (diurnal PEFV or significant improvement of FEV1 on administration of 200µg of salbutamol according to the GINA guidelines)	None
NIEMINEN 1992[116]	N=791 Adults Referred to pulmonary clinic with dyspnoea, wheezing, prolonged cough or history of asthma	MCT PD20 2,600 µg	Physician Dx with objective test - according to guidelines of the American Thoracic Society (documented variation in FEV1 or PEF of ≥15% after medication, or repeatedly a ≥20% spontaneous PEFv during a period of 2 weeks AND a ≥15% decrease in FEV1 after a specific allergen provocation or during an exercise test)	None

Study	Population	Index test & cut-off	Reference standard	Comparator test & cut-off
POPOVIK 2002[137]	N=195 Referred by GP with suspected asthma and symptoms of breathlessness / dyspnoea.	MCT PC20 8 mg/ml	Dx made on the basis of questionnaire, with typical medical history data of occasional asthma attacks with wheezing and nocturnal awakening because of dyspnoea, and reversible bronchial obstruction after salbutamol test (no further details stated)	None

Table 54: Summary of studies included in the review: children 5-16 years

Study	Population	Index test & cut-off	Reference standard	Comparator test & cut-off
ANDERSON 2009[7]	N=375 Adults and children/young people (6-50 yrs). Sn/sp given for (a) all ages; (b) <18 yrs only Signs and symptoms suggestive of asthma according to NIH questionnaire with an equivocal Dx of asthma or referred for further investigation	MCT PC20 16mg/ml	Physician Dx with objective test (access to exercise challenge test result, history, examination, skin tests and BDR but not methacholine and mannitol challenge tests)	None (no comparator used as population is suspected asthma)

Table 55: Clinical evidence profile: Methacholine Challenge Test/Histamine Challenge Test vs Reference Standard (physician Dx and objective test)

Index Test (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity %median (range)	Specificity %median (range)	Area Under Curve (range)	Quality
Methacholine/Histamine Challenge Test:Adults										
PC20 ≤8mg/ml	2	735	Very serious risk of bias[(a)]	No serious inconsistency[(b)]	No serious indirectness	n/a[(d)]	93 - 97	83 - 100	n/a	LOW
PD20 ≤6900µg	1	230	No risk of bias[(a)]	No serious inconsistency[(b)]	No serious indirectness	n/a[(d)]	77	82	n/a	HIGH

Index Test (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity %median (range)	Specificity %median (range)	Area Under Curve (range)	Quality
PD20 ≤2600µg	1	791	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness	n/a[d]	89	76	n/a	MODERATE
Methacholine Challenge Test: aged<18 years										
PC20 ≤16mg/ml	1	115	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness [c]	n/a[d]	66	63	n/a	LOW

(a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.
(b) Inconsistency was assessed by inspection of the sensitivity / specificity values in each of the studies in relation to a high (≥75%), moderate (50-74%) or low (<50%) effect estimate. Inconsistency was considered high if the point estimates from individual studies were within two areas, and very high if they were in 3 areas or both high and low areas.
(c) Study age range of 6-18 years does not match protocol of 6-15 years. Screening criteria included FEV1 >70% and non-atopic: selected group at screening may be a group with mild disease
(d) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed.

Table 56: Clinical evidence profile: Methacholine Challenge Test/Histamine Challenge Test vs Other Tests

Index Test (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity %median (range)	Specificity %median (range)	Area Under Curve (range)	Quality
Histamine Challenge Test vs Mannitol Challenge Test: Adults										
PD15 ≤1mg	1	37	No risk of bias[a]	No serious inconsistency[b]	Serious indirectness[c]	n/a[d]	100	39	n/a	MODERATE
PD15 ≤0.4mg	1	37	No risk of bias[a]	No serious inconsistency[b]	Serious indirectness[c]	n/a[d]	84	89	n/a	MODERATE

(a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.
(b) Inconsistency was assessed by inspection of the sensitivity / specificity values in each of the studies in relation to a high (≥75%), moderate (50-74%) or low (<50%) effect estimate. Inconsistency was considered high if the point estimates from individual studies were within two areas, and very high if they were in 3 areas or both high and low areas.
(c) Comparator test used as reference standard in people with confirmed asthma.
(d) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed.

18.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow chart in Appendix E.

Health economic modelling

Model overview/methods

Six diagnostic strategies were created using combinations of the following tests:

- spirometry
- bronchodilator reversibility
- FeNO
- peak expiratory flow variability
- challenge tests.

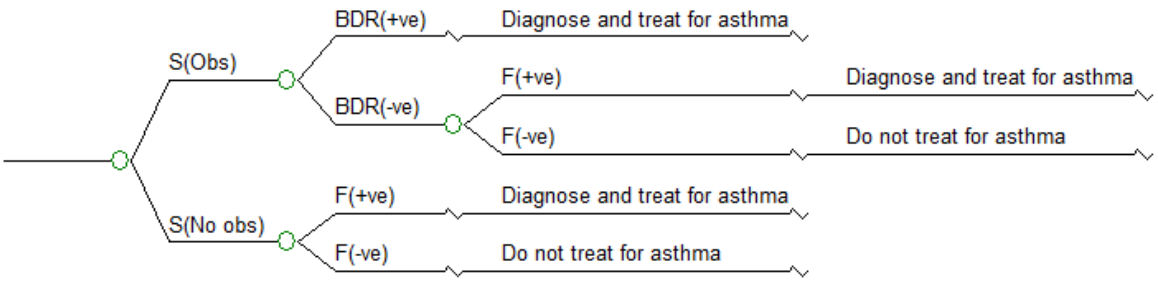
The GC agreed that only one challenge test would ever be conducted per patient meaning that challenge testing would only appear once in a diagnostic strategy. Therefore once the diagnostic strategies were developed it was proposed to duplicate each strategy which used challenge testing using the diagnostic accuracies and costs of histamine/methacholine, mannitol or exercise challenge test. However once the costs of an exercise challenge test and a methacholine challenge test had been established it was apparent that the exercise challenge test was the more expensive test. The clinical review also found that exercise challenge tests had a lower sensitivity and specificity when compared to a methacholine challenge test. Therefore exercise challenge tests were not modelled as they would always be dominated (more costly and provide lower health outcomes) when compared to methacholine challenge tests. Mannitol was also not modelled as the clinical review found it had a low sensitivity and specificity. Adding mannitol to the diagnostic pathway would in fact decrease the overall diagnostic accuracy of the pathway making it dominated by strategies that did not use challenge tests.

All the pathways were constructed using clinical judgement and taking into account the evidence produced in the clinical review.

Strategy 1

Strategy 1 involves the fewest number of tests. The exact point that each test appears in the diagnostic pathway and at which point patients are diagnosed with asthma is shown in Figure 4. For example in Figure 4 spirometry (S) is used as the initial test, followed by bronchodilator reversibility (BDR) if S detects obstruction (Obs) or FeNO (F) if S does not detect obstruction (No obs). If BDR is negative this is followed by F. A diagnosis of asthma is made with either a positive BDR or F, while asthma is excluded only with a negative F.

Figure 4: Strategy 1

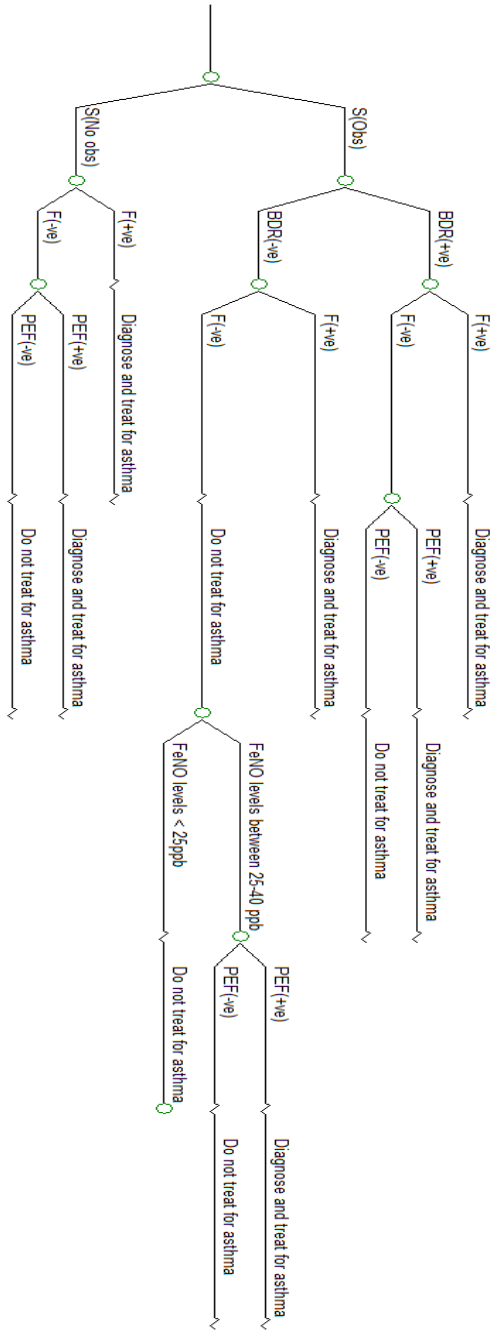


(-ve): negative; (+ve): positive; BDR: bronchodilator reversibility test; F: FeNO; S: spirometry; (Obs): obstruction

Strategy 2

The second strategy involves spirometry, bronchodilator reversibility, FeNO and PEF variability (PEF). The diagnostic pathway is shown in Figure 5. As more tests can be conducted after a FeNO test, if a patient receives a negative FeNO test, the FeNO level that was measured in the patient is also taken into account when deciding what to do next. This test is considered negative when the FeNO level is below 40 parts per billion (ppb), however the confidence in excluding a diagnosis of asthma depends on how close to this cut off the result is. If the FeNO level is below 25 parts per billion (ppb), along with an obstructive spirometry and a negative BDR, asthma is ruled out. If the FeNO level is between 25 – 40ppb then the diagnosis of asthma still cannot be ruled out and further tests are conducted. In strategy 2 below the patient goes on to have a PEFv test.

Figure 5: Strategy 2

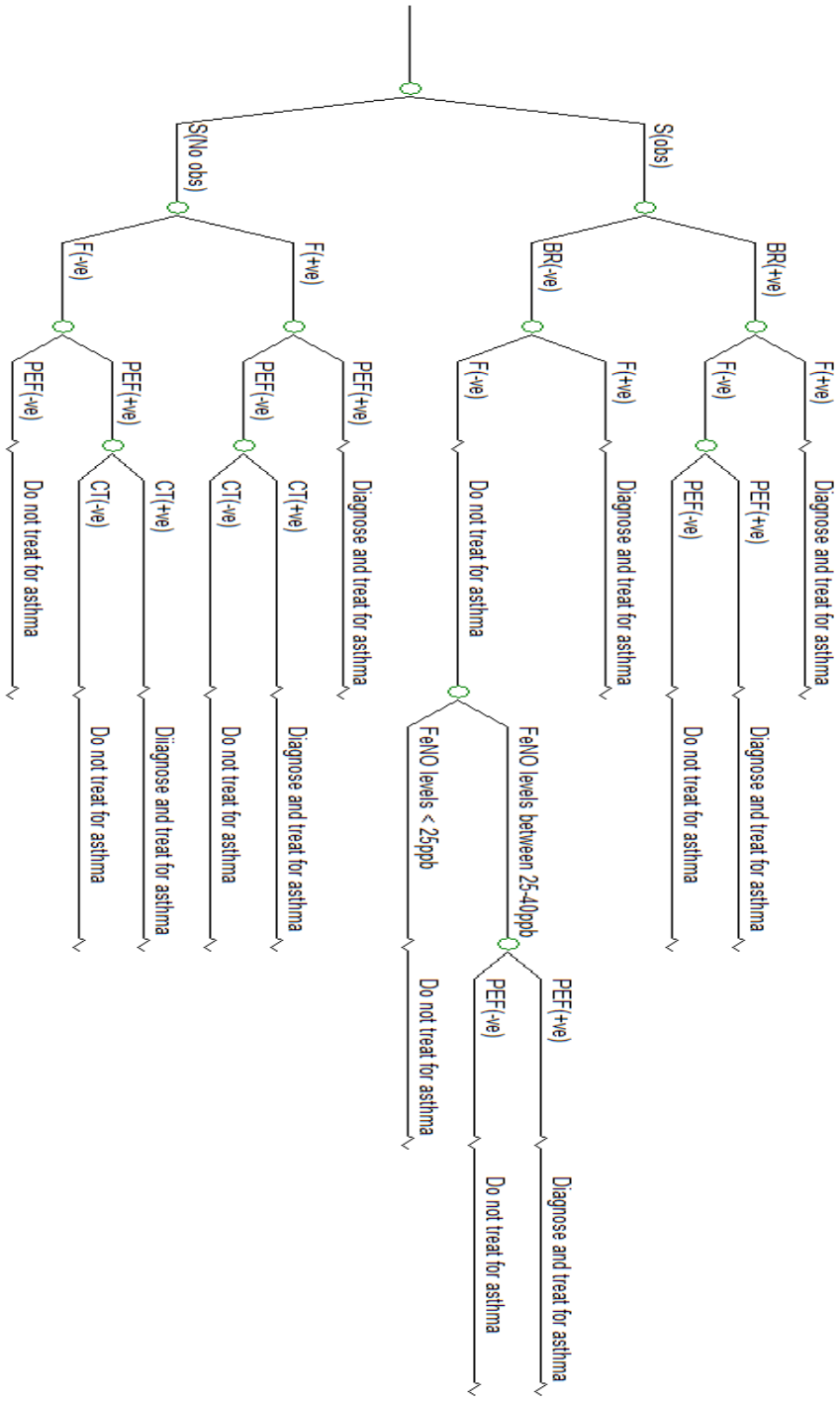


(-ve): negative; (+ve): positive; BDR: bronchodilator reversibility test; F: FeNO; S: spirometry; (Obs): obstruction; PEF: peak expiratory flow variability

Strategy 3

The third strategy uses spirometry, bronchodilator reversibility, FeNO, PEF variability and a methacholine challenge test (CT). The diagnostic pathway is shown in Figure 6. Note in this pathway challenge tests are only used on patients who have a non-obstructive spirometry.

Figure 6: Strategy 3

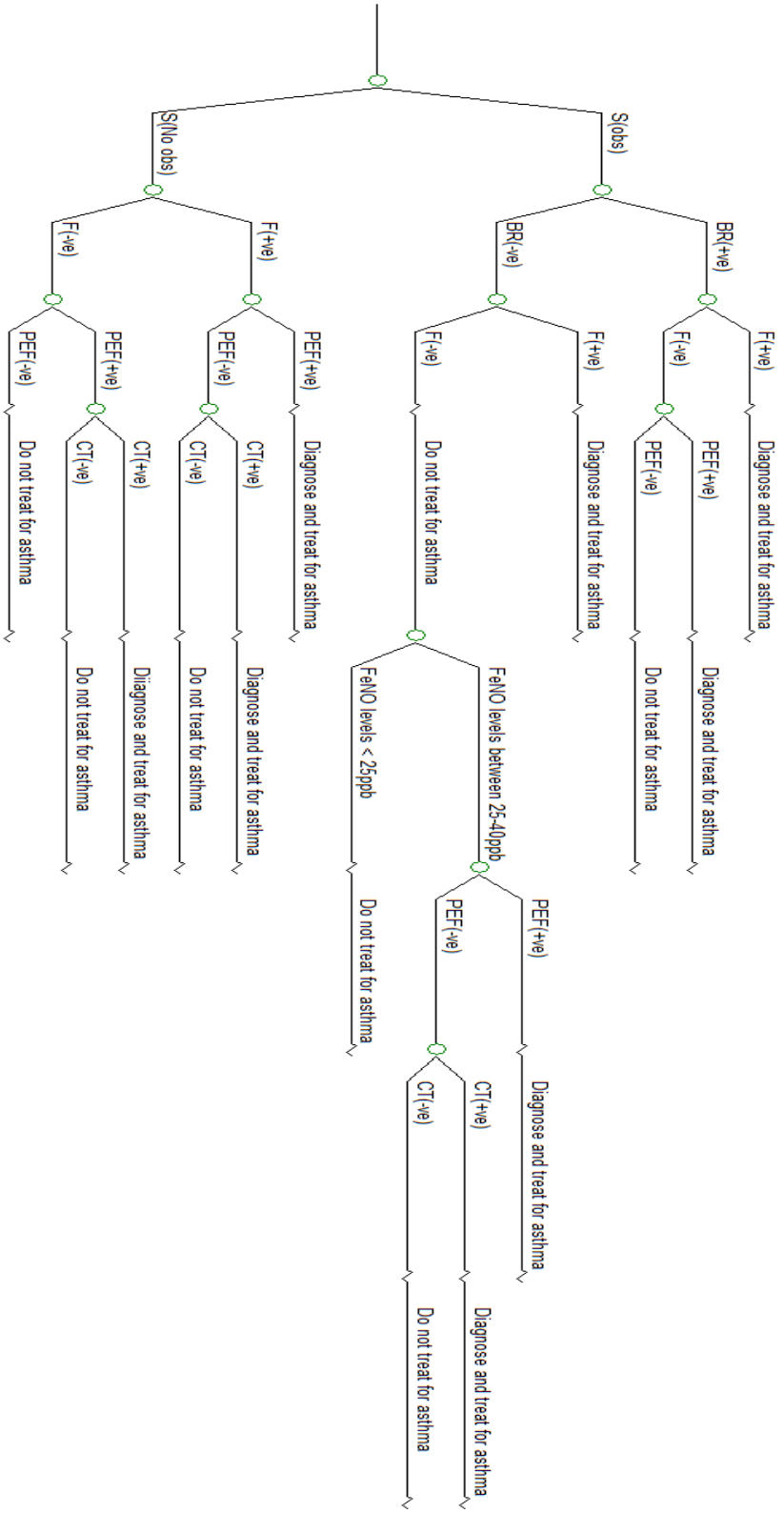


(-ve): negative; (+ve): positive; BDR: bronchodilator reversibility test; CT: challenge test; F: FeNO; S: spirometry; (Obs): obstruction; PEF: peak expiratory flow variability

Strategy 4

The forth strategy shown in Figure 7 expands the use of challenge tests as seen in strategy 3. Now a CT is also conducted on patients with a positive BDR, negative FeNO and a negative PEFv result. The use of FeNO levels is also taken into account, whereby a CT is only conducted in this arm when FeNO levels are between 25-40ppb.

Figure 7: Strategy 4

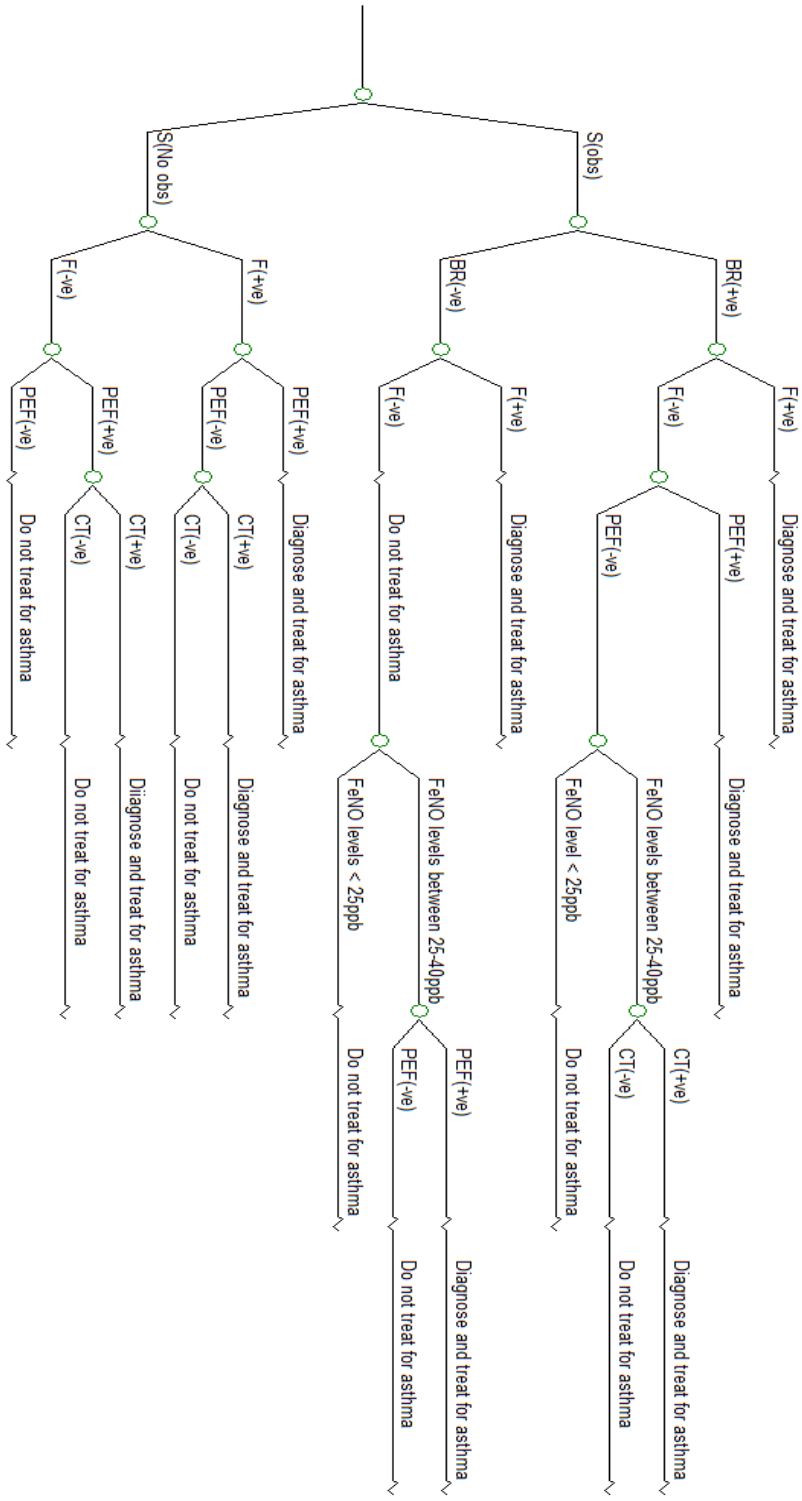


(-ve): negative; (+ve): positive; BDR: bronchodilator reversibility test; CT: challenge test; F: FeNO; S: spirometry; (Obs): obstruction; PEF: peak expiratory flow variability

Strategy 5

The fifth strategy, shown below in Figure 8, also expands the use of challenge tests, as seen in strategy 3, however places the additional CT at a different point in the pathway. Now a CT is also conducted on patients with a negative BDR, negative FeNO (between 25-40ppb) and a negative PEFv test result.

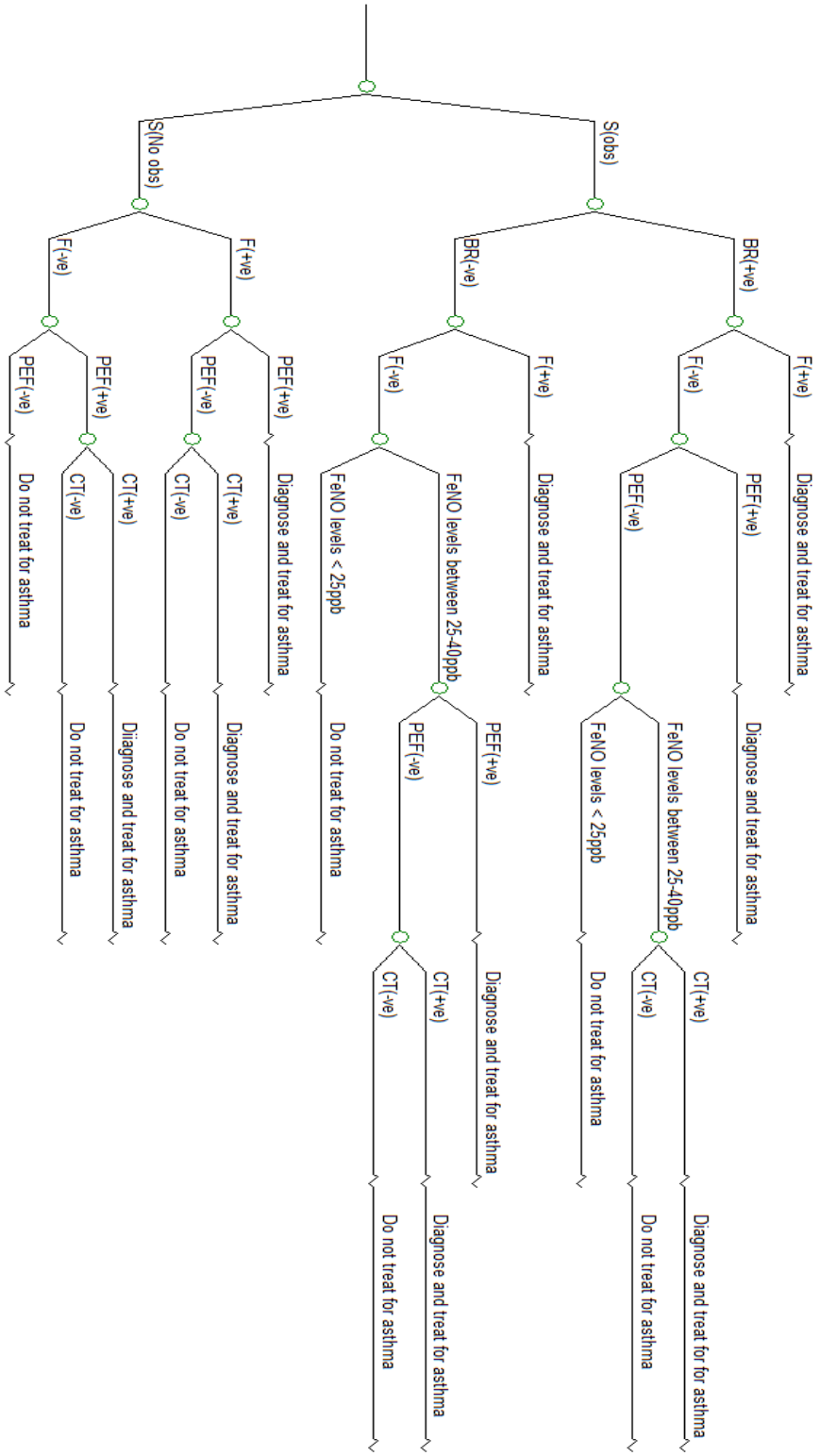
Figure 8: **Strategy 5**



Strategy 6

The sixth strategy, shown below in Figure 9, is the most comprehensive and uses the maximum number of challenge tests.

Figure 9: Strategy 6



Current practice

A final strategy considered involves not giving the patient any tests and diagnosing without the use of objective tests. To make this strategy more reflective of current practice it is assumed that some of the non-asthmatics will be correctly diagnosed as not having asthma. One prevailing thought is that one third of people currently diagnosed with asthma are misdiagnosed, ie they do not have asthma (False positive) according to a study by Aaron et al.[1,1] Therefore, the proportion of false positives calculated in this strategy will be a third of the total number of positive diagnoses made:

$$\frac{False\ positives}{False\ positives + True\ positives} = \frac{1}{3}$$

As no tests are conducted the only costs that are incurred in this strategy are those that occur after the diagnosis is made (e.g. the cost of asthma treatment). An assumption was made that all people with asthma are correctly diagnosed giving this strategy a sensitivity of 100%.

The economic evaluation was a cost-utility analysis, where lifetime costs and quality-adjusted life-years (QALYs) were considered from a UK NHS and personal social services perspective. The model was based on two parts:

- **Decision tree** - Using the sensitivity and specificity, combined with data on the prevalence of asthma in the defined population, the model identifies the proportion of patients that receive a true positive (TP), true negative (TN), false positive (FP) or false negative (FN) diagnosis.
- **Markov model** - Once the diagnosis is made the patient moves on to the second part of the model which involves a Markov model to fully evaluate the patients' health and cost outcomes. This incorporates the time spent misdiagnosed and the associated decrease in quality of life, higher mortality risks and wasted NHS resources.

The model makes some assumptions concerning:

- Conditional dependence
- The underlying condition the individual will have if they present asthma symptoms but do not have asthma
- The length of misdiagnosis
- Quality of life for incorrectly treated individuals (false positives and false negatives)

The accuracy of diagnostic tests was taken from the clinical reviews presented in this guideline and altered, where appropriate, in light of conditional dependence. Full details on how these assumptions were implemented can be found in the full model write up in appendix M.

Results

The results below in Table 57 show that diagnostic strategy 3 has the highest net monetary benefit and is therefore the most cost-effective way of diagnosing asthma. Strategy 6 produces the highest number of QALYs however is not deemed cost-effective at a £20,000 per QALY threshold. Strategy 1 produces the least QALYs and the highest cost.

Table 57: Base case results (probabilistic)

Strategy	Mean per patient		NMB at £20,000 threshold	Rank at £20,000 threshold	Probability of being CE at £20,000 threshold
	QALYs	Cost			
Current practice	16.7766	£3,730	£331,802	6	6%
Strategy 1	16.7760	£3,753	£331,768	7	0%
Strategy 2	16.7776	£3,686	£331,866	5	19%
Strategy 3	16.7783	£3,683	£331,882	1	44%

Strategy	Mean per patient		NMB at £20,000 threshold	Rank at £20,000 threshold	Probability of being CE at £20,000 threshold
	QALYs	Cost			
Strategy 4	16.7785	£3,691	£331,878	4	0%
Strategy 5	16.7784	£3,686	£331,881	2	23%
Strategy 6	16.7787	£3,695	£331,879	3	8%

Table 58 below shows the overall sensitivity and specificity of each diagnostic pathway, that is the percentage of patients with asthma that receive a true positive diagnosis and the percentage of patients without asthma that receive a true negative diagnosis.

Table 58: Diagnostic accuracies of each strategy

	Current practice	Strategy 1	Strategy 2	Strategy 3	Strategy 4	Strategy 5	Strategy 6
Sensitivity	100%	90.3%	89.3%	86.3%	88.7%	87.7%	90.3%
Specificity	65.8%	69.1%	82.4%	89.5%	89.4%	89.4%	89.4%

Table 58 shows a dramatic increase in specificity once FeNO is routinely performed on all individuals (strategy 2). Once challenge tests are added to the diagnostic pathway specificity also increases considerably (strategies 3, 4, 5 and 6). No strategy has a single highest value for sensitivity and specificity though strategy 6 has the highest diagnostic odds ratio. Finally Table 59 below shows the costs associated with objective tests for each strategy.

Table 59: Cost of testing in each strategy

	Current practice	Strategy 1	Strategy 2	Strategy 3	Strategy 4	Strategy 5	Strategy 6
Cost associated with diagnostic tests	£0	£42	£52	£92	£100	£95	£103

Table 59 shows that although the strategies that include challenge tests cost more the increase in cost is far less than the additional cost of a single challenge tests as the majority of individuals will not go on to receive one.

Overall this analysis showed that strategy 3 is the most cost-effective strategy at a £20,000 per QALY threshold. Further challenge testing on patients with an obstructive spirometry who had either (negative BDR and FeNO results) or (positive BDR, negative FeNO and PEFv results) provided higher health outcomes however were not cost-effective at a £20,000 per QALY threshold.

A series of sensitivity analyses show that, with regards to the routine use of challenge tests in asthma diagnosis, the model results are highly robust to health and cost outcomes attached to false positives, speed of re-diagnosis and the effects of conditional dependence. These are three key uncertain areas of the model. The sensitivity analyses did show however that there is scope for additional challenge tests (as detailed in strategies 5 and 6) to be cost-effective at a £20,000 per QALY threshold. In the base case the ICER for providing an additional challenge test as detailed in strategy 5 was £20,276 per QALY. The ICER of providing challenge tests at all appropriate points in the pathway, as detailed by strategy 6, was £32,565. However sensitivity analyses showed there were some scenarios where it was cost-effective to do these additional challenge tests such as when the diagnostic accuracy of FeNO changed. The sensitivity analyses also showed that the cost-

effectiveness of performing these additional challenge tests was contingent on a very high specificity of methacholine challenge tests. As the model does not fully capture conditional dependence concerning the accuracy of this test the GC were cautious to routinely recommend these additional challenge tests.

The GC believe these additional challenge tests would be cost-effective in some situations where other diagnoses could not easily be ruled in. For example if another diagnosis, such as COPD, is considered likely then further challenge testing should not be considered. Therefore these additional challenge tests should not be routinely carried out, unlike those placed in strategy 3 but should still be considered.

The main limitations of the model concerned the lack of clinical data informing parameters associated with misdiagnosis as, due to ethical reasons; this evidence will likely never be available. However the model results were robust to all the assumptions imposed arounds these parameters which therefore limits their impact on the model.

Full details of the model and results can be found in appendix M.

Unit cost of performing a direct bronchial challenge test on children

As an economic model was not feasible for children, the GC considered the unit cost of performing a direct bronchial challenge test to evaluate its cost-effectiveness as part of a pathway for diagnosing asthma. The NHS reference cost associated with 'Bronchial reactivity studies' (HRG code: DZ36Z) is £177.[46] A paediatric respiratory outpatient visit would also need to be considered to interpret the result; this is cost as £197 in the NHS reference costs.

18.5 Evidence statements

Clinical

- Two studies with 735 adults showed that methacholine/histamine challenge test (PC20 cut-off 8mg/ml) has a sensitivity range of 93-97% and a corresponding specificity range of 83-100% for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 230 adults showed that methacholine/histamine challenge test (PD20 cut-off 6900μg) has a sensitivity of 77% and a corresponding specificity of 82% for diagnosing asthma in people presenting with respiratory signs and symptoms. (HIGH QUALITY)
- One study with 791 adults showed that methacholine/histamine challenge test (PD20 cut-off 2600μg) has a sensitivity of 89% and a corresponding specificity of 76% for diagnosing asthma in people presenting with respiratory signs and symptoms. (MODERATE QUALITY)
- One study with 115 children and young people showed that methacholine/histamine challenge test (PC20 cut-off 16mg/ml) has a sensitivity of 66% and a corresponding specificity of 63% for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 37 adults showed that methacholine/histamine challenge test (PD15 cut-off 1mg) has a sensitivity of 100% and a corresponding specificity of 39% for a positive mannitol challenge test in people with asthma. (MODERATE QUALITY)
- One study with 37 adults showed that methacholine/histamine challenge test (PD15 cut-off 4mg) has a sensitivity of 84% and a corresponding specificity of 89% for a positive mannitol challenge test in people with asthma. (MODERATE QUALITY)

Economic

- An original health economic model found that histamine and methacholine challenge test (together with spirometry, BDR, FeNO and PEFv) was part of the most cost-effective diagnostic

pathway used to diagnose asthma in adults aged 16 and over (see diagnostic algorithms in section 4.1). This evidence is directly applicable with minor limitations.

18.6 Recommendations and link to evidence

Recommendations	**24. Offer a direct bronchial challenge test with histamine or methacholine[g] to adults (aged 17 and over) if there is diagnostic uncertainty after a normal spirometry and either a:** • **FeNO level of 40 ppb or more and no variability in peak flow readings, or** • **FeNO level of 39 ppb or less with variability in peak flow readings.** **Regard a PC20 value of 8 mg/ml or less as a positive test.** **25. Consider a direct bronchial challenge test with histamine or methacholine[g] in adults (aged 17 and over) with:** • **obstructive spirometry without bronchodilator reversibility and** • **a FeNO level between 25 and 39 ppb and** • **no variability in peak flow readings (less than 20% variability over 2 to 4 weeks).** **Regard a PC20 value of 8 mg/ml or less as a positive test.** **26. If a direct bronchial challenge test with histamine or methacholine is unavailable, suspect asthma and review the diagnosis after treatment, or refer to a centre with access to a histamine or methacholine challenge test.**
Research recommendations	**1. What is the acceptability and diagnostic accuracy of objective tests that could be used to comprise a diagnostic pathway for asthma in children aged 5-16 years old (for example, exercise challenge, direct bronchial challenge with histamine or methacholine, indirect bronchial challenge with mannitol and peripheral blood eosinophil count)?**
Relative values of different outcomes	The GC was interested in the diagnostic test accuracy of airway hyper-reactivity (non-specific bronchial challenge) with histamine or methacholine. There is evidence of significant misdiagnosis of asthma, with both false positive and false negative errors (see introduction to the guideline). The GC was therefore interested in both the sensitivity and specificity of diagnostic tests and prioritised studies which reported each of these outcomes. The GC excluded studies with a cut-off value of a fall in FEV1 greater than or equal to 20% (PC20) at a concentration of ≤16mg/ml from the review. Values at the higher end of this range are likely to occur in some healthy individuals, and the GC was concerned about a high number of false positives at this cut-off value. Evidence from studies reporting a PD20 cut-off of ≤6900µg or ≤2600µg were

[g] At the time of publication (November 2017), histamine and methacholine did not have UK marketing authorisation for this use. The healthcare professional should follow relevant professional guidance, taking full responsibility for the decision to use this test. Informed consent should be obtained and documented. See the General Medical Council's Prescribing guidance: prescribing unlicensed medicines for further information.

	considered, particularly when a ROC curve had been used in the study to identify the diagnostic cut-off. In children, evidence from studies with a cut-off value of ≤16mg/ml was considered due to lack of other evidence.
Trade-off between clinical benefits and harms	The sensitivity and specificity of methacholine and histamine challenge tests was high in adults at all three cut-off thresholds (including the preferred threshold of 8mg/ml), suggesting it is both a good rule-in and rule-out test at this threshold. In children aged 5-16 years, there was no evidence at the preferred cut-off threshold of 8mg/ml, so evidence was included at a cut-off of 16mg/ml. The sensitivity and specificity of the test in children aged 5-16 years was moderate. As with all functional tests, it relies on the ability to perform spirometry according to standard spirometry techniques. The methacholine or histamine challenge test is time-consuming and for safety reasons needs to be performed in secondary care. Methacholine and histamine challenge tests are very well tolerated in the vast majority of patients, but there is a rare risk of severe bronchospasm and therefore resuscitation facilities are required wherever the test is to be performed. Patients who have been started on empirical anti-asthma treatment will have to stop treatment prior to the test, which may put them at risk, although this risk is low in people in whom there is genuine diagnostic doubt. Histamine can cause throat irritation and methacholine is likely to be better tolerated.
Economic considerations	No economic evaluations were found which assessed the use of direct bronchial challenge tests as part of a diagnostic pathway. Therefore, an original health economic model was built to assess the cost-effectiveness of different diagnostic pathways for asthma. The model assessed the additional costs of tests against cost-savings from unnecessary asthma medication, and the increased health outcomes from providing correct treatment. The most cost-effective strategy involved using spirometry, bronchodilator reversibility (BDR), FeNO, peak expiratory flow variability (PEFv) and a methacholine challenge test (MCT) in selected individuals to diagnose asthma. In this strategy everybody with symptoms of asthma would undergo a spirometry test and a FeNO measurement, those who had an obstructive spirometry would also receive a BDR test. Only those who had non-obstructive spirometry and conflicting FeNO and PEFv test results would receive a MCT. Adopting this strategy dominated all other strategies apart from those which performed challenge tests at additional points in the pathway. The ICERs of adopting these further stratgies were above £20,000 per QALY gained. The model result of using challenge tests in a diagnostic pathway as opposed to no challenge tests was robust to changes in all key assumptions apart from significant changes in health benefits derived from correctly diagnosing people who do not have asthma. The GC noted that the model did not consider mortality impacts; severe misdiagnoses such as tuberculosis or lung cancer; or the adverse effects of asthma treatment. Including these aspcets would increase the cost-effectiveness of strategies that include challenge testing as they have a higher specificity. This shows there is considerable value in using highly accurate challenge tests outside of primary care to diagnose asthma. The sensitivity analysis did pick up on an element of uncertainty in the model with regards to how many MCTs should be conducted. This uncertainty has been captured in the strength of recommendations around the use of challenge tests in

	certain points of the pathway. Where challenge tests are offered in the algorithm of recommendation there is strong evidence to suggest that doing so is cost-effective at a £20,000 per QALY threshold, therefore they should be routinely performed on all patients at these points. In the base case performing challenge tests at these points was a dominant strategy, producing better health outcomes at a lower cost, than not performing challenge tests. This result was robust to a variety of sensitivity analyses which supports their stong recommendation. Where challenge tests are 'considered' in the algorithm of recommendations there is evidence to suggest they could be cost-effective at a £20,000 per QALY threshold, however this evidence is much more uncertain. At these points in the pathway MCTs should be considered on a case by case basis where the clinician will decide if they have enough evidence to exclude asthma. If the clinican believes the individual has strong signs of asthma or their symptoms are unlikely to be caused by something else then referring them for a challenge test and confirming this result will be a cost-effective use of resources. For example COPD can be ruled-out for an individual who is 30 years old and never smoked. By reserving challenge tests for these patients it can be ensured that only those with the highest diagnostic uncertainty receive the tests at this point making them cost-effective for these patients. For children, as the clinical evidence informing the diagnostic accuracy of these tests was poor and the cost of using these tests is very high, the GC agreed they could not make a recommendation concerning their use in a diagnostic pathway from consensus alone. Therefore an appropriate research recommendation was devised.
Quality of evidence	The quality of the evidence ranged from low to high. The GC noted that in included studies there were very few false positives, which increases the specificity of the test. It is possible that the test will perform less well outside trial conditions. • In children aged <5 years, we did not search for studies as spirometry is not routinely performed in this age group. • In children aged 5-16 years, there was one included study[7] using the reference standard (physician diagnosis with objective test) that addressed the use of methacholine in this age group. Methacholine challenge tests had a moderate sensitivity and moderate specificity. Evidence with a cut-off value of ≤16mg/ml was considered due to lack of evidence in children and young people. The GC was concerned that the inclusion criteria for the study (FEV1 >70% and non-atopic) may pre-select a population with mild disease, and therefore the evidence was downgraded for indirectness. Overall, this evidence was of low quality. • In adults, there were four included studies[69,90,116,137] using the best reference standard (physician diagnosis with objective test) that addressed the use of methacholine or histamine challenge test. Two studies used a cut-off threshold for a positive test of 8mg/ml. The evidence at this threshold was of low quality. Additionally, in adults there was one study comparing methacholine or histamine challenge testing with mannitol challenge testing as a proxy for asthma; this study showed high sensitivity but low specificity at a cut-off value of PC15≤1mg and a high sensitivity and high specificity at a cut-off value of PC15≤0.4mg. The economic evidence was assessed as directly applicable with minor limitations.
Other considerations	The GC recommendation is based on review of the evidence in adults and consensus opinion of the GC for a research recommendation in children. The GC made a high-priority research recommendation to investigate the acceptability and performance characteristics of objective tests, including histamine and methacholine challenge tests, which could be used to comprise a diagnostic pathway for asthma in children aged 5-16 years old. Further details on the high-priority research recommendation

made can be found in appendix N, along with the full list of research recommendations.

Based on the available evidence, the bronchial challenge is currently the best test for adults; however its use is limited by familiarity and ability to perform the test. It is limited to secondary care and, particularly in a paediatric population, there is a lack of experience using the test. As with all functional tests, it relies on the ability to perform spirometry according to standard spirometry techniques. Bronchial hyper-reactivity may vary over time so a negative test does not exclude asthma, and the test cannot be used to rule-out asthma on its own.

The available evidence suggested that the best combination of sensitivity/specificity is at the cut-off point of 8mg/ml for PC20. This was included in the recommendation. The GC accepted that calculating a PD20 rather than a PC20 is entirely valid, but PD20 will vary more with the dosimeter used in the test, and therefore a recommended cut-off value for PD20 is not given.

When considering the use of a histamine/methacholine challenge test (MCT) in a diagnostic pathway the GC considered the diagnostic accuracy of the test and the practicality of implementing the test. Unlike most tests that are used to diagnose asthma the MCT, currently, can only be performed in secondary care. The GC agreed that this test should only be performed on individuals who have considerable diagnostic uncertainty after performing other tests. Regarding the MCTs exact placement in a pathway the GC noted certain combination of test results they considered would result in a definitive diagnosis. Therefore the MCT is used after conflicting test results where diagnostic uncertainty is highest.

The GC agreed this may be an important and useful test in the diagnosis of asthma in children but there is currently inadequate evidence available to assess either its efficacy or acceptability and tolerability for children and their care givers.

The GC noted that histamine and methacholine did not have UK marketing authorisation for this use at the time of publication of the guideline and included a footnote to the recommendation to highlight this.

Details of the feasibility project are available in appendix Q. The aim of the feasibility project was to assess the impact and feasibility of adopting the diagnostic tests (e.g. FeNO and spirometry) recommended, into primary care.

Although the GC concluded that bronchial challenge should be part of a first-class diagnostic pathway, they also noted the current lack of availability and this was emphasised by the feasibility study results. Of the 143 people evaluated, 14 reached the point in the diagnostic pathway at which bronchial challenge was appropriate, but no patient had had the test at the end of the period of analysis (some were waiting for secondary care appointments, for others the test was not available locally). Therefore, the GC made a recommendation that if a histamine/methacholine challenge test is unavailable, asthma should be suspected and this diagnosis should be reviewed after treatment, or the patient should be referred to a centre with access to a histamine or methacholine challenge test.

Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

19 Diagnosis: Indirect bronchial challenge test with mannitol

19.1 Introduction

Hyper-reactivity of the airways is a key feature of asthma. Inhaled dry powder mannitol increases the osmolarity of the mucosal lining of the respiratory epithelium and leads to bronchoconstriction by inducing the release of inflammatory mediators. This is an indirect bronchial provocation test similar to exercise testing and to eucapnoiec voluntary hyperpnea. The test is performed by inhaling increasing doses of mannitol until lung function testing demonstates a 15% reduction in FEV1 from baseline. It is said to mimic the mechanism of bronchoconstriction in exercise induced asthma. As this test is easy to perform, easy to standardise and carries a low risk of severe bronchoconstriction, its utility and accuracy in the diagnosis of asthma in people with asthma symptoms is of interest.

19.2 Review question: In people under investigation for asthma, what is the diagnostic test accuracy and cost-effectiveness of airway hyper-reactivity (non-specific bronchial challenge) with mannitol?

For full details see review protocol in Appendix C.

Table 60: Characteristics of review question

Population	People with suspected asthma (presenting with respiratory symptoms). Ages stratified into the following 2 different groups: • Children/young people (5-16 years old) • Adults (>16 years old)
Index test	Mannitol
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test
Statistical measures	Diagnostic accuracy (sensitivity, specificity)

19.3 Clinical evidence

We searched for cross sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of airway hyper-reactivity (non-specific bronchial challenge) with mannitol to identify whether the condition is present (as indicated by the reference standard) in people under investigation for asthma.

One study was included in the review[7]. Evidence from this study is summarised in Table 61 and in the clinical evidence profile below (Table 62).

Evidence was available from adults and children/young people pooled together (6-50 years) and from children and young people alone (6-18 years). Data for forest plots have been separated into these two strata; however these groups are indirect to the protocol. See also Appendix D: Clinical article selection, Appendix J: Coupled sensitivity / specificity forest plots, Appendix G: Clinical evidence tables and Appendix K: Excluded clinical studies.

The included study[7] was a cross-sectional study, and looked at the diagnostic accuracy of the mannitol challenge test in patients with suspected asthma or asthma symptoms. The reference standard was physician's diagnosis of asthma with an objective test.

Summary of included studies

Table 61: Summary of studies included in the review

Study	Population	Index test & cut-off	Reference standard	Comparator test & cut-off
ANDERSON 2009[7]	N=375 Adults and children/young people (6-50 yrs). Sn/sp given for: • all ages • <18 yrs only Signs and symptoms suggestive of asthma according to NIH questionnaire with an equivocal Dx of asthma or referred for further investigation	Mannitol ≥15% fall in FEV1 ≤635mg or 10% fall between consecutive doses.	Physician Dx with objective test (access to exercise challenge test result, history, examination, skin tests and BDR but not methacholine and mannitol challenge tests)	No comparator as population is suspected asthma not confirmed asthma

Table 62: Clinical evidence profile: Mannitol Challenge Test vs Reference Standard (physician Dx and objective test)

Index Test (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity % (range)	Specificity % (range)	Area Under Curve (range)	Quality
Mannitol Challenge Test – all age groups										
≥15% fall in FEV1 ≤635mg or 10% fall between consecutive doses	1	375	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[c]	n/a[e]	56	75	n/a	LOW
Mannitol Challenge Test <18 yrs										
≥15% fall in FEV1 ≤635mg or 10% fall between consecutive doses	1	115	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[d]	n/a[e]	63	81	n/a	LOW

(a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.

(b) Inconsistency was assessed by inspection of the sensitivity / specificity values in each of the studies in relation to a high (≥75%), moderate (50-74%) or low (<50%) effect estimate. Inconsistency was considered high if the point estimates from individual studies were within two areas, and very high if they were in 3 areas or both high and low areas.

(c) Inclusion of one study with mixed population of adults and children/young people. Screening criteria included FEV1 >70% and non-atopic: selected group at screening may be a group with mild disease.

(d) Study age range of 6-18 years does not match protocol of 5-16 years. Screening criteria included FEV1 >70% and non-atopic: selected group at screening may be a group with mild disease.

(e) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed.

19.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow chart in Appendix E.

New cost-effectiveness analysis

An original health economic model was built for adults to assess the cost-effectiveness of several diagnostic pathways. Non-specific bronchial challenge test with mannitol was considered as part of the diagnostic strategy, however due to its low diagnostic accuracy, found in the clinical review, adding mannitol to the diagnostic pathway reduced the overall sensitivity and specificity. This meant the strategy would cost more and produce poorer health outcomes making it a dominated strategy. Full details of the model can be found in Appendix M. A summary of the model can be found in section 18.4.

Unit cost of performing a bronchial challenge test with mannitol on children

As an economic model was not feasible for children, the GC considered the unit cost of performing a direct bronchial challenge test to evaluate its cost-effectiveness as part of a pathway for diagnosing asthma. The NHS reference cost associated with 'Bronchial reactivity studies' (HRG code: DZ36Z) is £177.[46] A paediatric respiratory outpatient visit would also need to be considered to interpret the result; this is cost as £197 in the NHS reference costs.

19.5 Evidence statements

Clinical

- One study with 375 adults, children and young people showed that mannitol challenge test has a sensitivity of 56% and a corresponding specificity of 75% for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 115 children and young people showed that mannitol challenge test has a sensitivity of 63% and a corresponding specificity of 81% for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)

Economic

- No relevant economic evaluations were identified.

19.6 Recommendations and link to evidence

Recommendations	**No clinical recommendation.**
Research recommendations	**1. What is the acceptability and diagnostic accuracy of objective tests that could be used to comprise a diagnostic pathway for asthma in children aged 5-16 years old (for example, exercise challenge, direct bronchial challenge with histamine or methacholine, indirect bronchial challenge with mannitol and peripheral blood eosinophil count)?** **2. What is the clinical and cost effectiveness of using an indirect bronchial**

	challenge test with mannitol to diagnose asthma in adults and young people older than 16?
Relative values of different outcomes	The GC was interested in the diagnostic test accuracy of airway hyper-reactivity (non-specific bronchial challenge) with mannitol. There is evidence of significant misdiagnosis of asthma, with both false positive and false negative errors (see introduction to the guideline). The GC was therefore interested in both the sensitivity and specificity of diagnostic tests and prioritised studies which reported each of these outcomes for mannitol challenge. Sensitivity and specificity values will depend on the chosen cut-off point. The cut-off point of a fall in FEV1 greater than or equal to 15% at a cumulative dose ≤635mg of mannitol was taken as the standard cut-off currently used in clinical practice.
Trade-off between clinical benefits and harms	Data were considered from adults, children and young people pooled together, and from children <18 years alone. In both groups, the mannitol challenge test had a moderate sensitivity and a high specificity. This suggests mannitol challenge tests have more utility as a rule-in test, meaning that a positive test would identify asthma, whereas a negative test may not rule-out the condition. The mannitol challenge test is relatively less time-consuming compared to other BHR tests (but more time-consuming than some objective tests). It is currently only licensed to be performed in secondary care. There is a risk of a severe bronchospasm response to mannitol and therefore resuscitation facilities are required wherever the test is to be performed.
Economic considerations	No economic evaluations were found which assessed the use of bronchial challenge test with mannitol as part of a diagnostic pathway. An original health-economic model was built to assess the cost-effectiveness of different diagnostic pathways for asthma. The model assessed the additional costs of tests against cost savings from unnecessary asthma medication and increased health outcomes from providing correct treatment. A strategy was considered that gave patients mannitol rather than a histamine/methacholine challenge test. Only one clinical study was found for mannitol which showed it had moderate sensitivity and a moderate/high specificity. However due to its low diagnostic accuracy relative to tests that preceed it, such as FeNO, using it in a diagnostic pathway would actually reduce the overall diagnostic accuracy of the strategy thus making mannitol a dominated option. The GC concurred that the body of evidence found for mannitol was not of high quality and not strong enough to make a recommendation. The GC also noted that in the future there is scope to perform mannitol in primary care which would significantly lower its cost, as there would be no requirement for referral to secondary care. For this to occur, further research will be required to demonstrate that the mannitol challenge test is safe in primary care allowing for a change to its licensed indication. For children, as the clinical evidence informing the diagnostic accuracy of these tests was poor and the cost of using these tests is very high, the GC agreed they could not make a recommendation concerning their use in a diagnostic pathway from consensus alone. Therefore an appropriate research recommendation was devised.
Quality of evidence	Evidence from one study was included in the review. • In children aged <5 years, we did not search for studies, as spirometry is not routinely performed in this age-group and the mannitol test requires a spirometry test to be performed as it is based on induced change in FEV1. • In children aged 5-16 years, there was one included study[7] using the reference standard (physician diagnosis with objective test) that addressed the use of

	mannitol in this age group. Mannitol challenge test had a moderate sensitivity and higher specificity. The GC was concerned that the inclusion criteria for the study (FEV1 >70% and non-atopic) may pre-select a population with mild disease, and therefore the evidence was downgraded for indirectness. Overall, the evidence was of low quality. • In adults, no studies were identified using the best reference standard (physician diagnosis with objective test) that addressed the use of mannitol in adults-alone group. The one included study[7] looked at a mixed population of adults and children aged 5-16 years using the reference standard (physician diagnosis with objective test). The GC considered this evidence for the adult population. Mannitol challenge test had a moderate sensitivity and higher specificity. The quality of the evidence was downgraded, as the included study had a mixed population of adults, children and young people, and was indirect to the protocol. The GC was concerned that the inclusion criteria for the study (FEV1 >70% and non-atopic) may pre-select a population with mild disease, and therefore the evidence was downgraded for indirectness. Overall, the evidence was of low quality.
Other considerations	The GC was interested in the position of indirect challenge tests within an algorithm of diagnostic tests. However, the GC agreed there was not enough evidence of sufficient quality to make a recommendation regarding the use of mannitol to diagnose asthma. The GC made a high-priority research recommendation to investigate the clinical and cost-effectiveness of mannitol challenge tests in adults aged 17 years and over. Further details on the research recommendation made can be found in appendix N. Mannitol testing requires spirometry to be performed as part of the test. The GC noted that the spirometry should be performed in accordance with standard technical guidelines. Indirect bronchial challenge testing with mannitol is done in secondary care only, as this test is currently not licensed in primary care. The GC made a high-priority research recommendation to investigate the acceptability and performance characteristics of objective tests, including mannitol challenge tests, which could be used to comprise a diagnostic pathway for asthma in children aged 5-16 years old. Further details on the high-priority research recommendation made can be found in appendix N. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. One study was identified which the GC agreed did not suggest a change was warranted in the decision to make no clinical recommendation for this topic.

20 Diagnosis: Indirect bronchial challenge test with exercise

20.1 Introduction

In cases where a diagnosis of asthma remains in doubt and objective evidence is lacking, a bronchial challenge test can be used, and this could be either a direct or indirect challenge test. Direct bronchial challenge testing with increasing doses of inhaled methacholine is covered in chapter 18 of this guideline; currently it is rarely performed in the UK, and particularly not in children. Instead, indirect bronchial provocation testing is done in children in the form of an exercise test. Exercise testing is an 'indirect' measure of BHR whereby, in susceptible individuals, it triggers the endogenous release of inflammatory mediators, primarily eicosanoids, prostaglandins and histamine, which cause bronchoconstriction[6,124]. This is due to thermal and osmotic changes at the airway surface as a consequence of increased aerobic demand.

The European Respiratory Society (ERS) and American Thoracic Society (ATS) recommend its use in patients with an intermediate probability of asthma and no evidence of reversible airway obstruction as well as for patients with severe symptoms or a poor response to pharmacotherapy[123,124]. However, currently there is uncertainty around the diagnostic accuracy of exercise challenge tests.

20.2 Review question: In people under investigation for asthma, what is the diagnostic accuracy of bronchoconstriction in response to an exercise challenge?

For full details see review protocol in Appendix C.

Table 63: PICO characteristics of review question

Component	Description
Population / Target condition	People with suspected asthma (presenting with respiratory symptoms). Ages stratified into the following 2 different groups: • Children/young people (5-16 years old) • Adults (>16 years old)
Index test	Exercise challenge test (>10% FEV1 bronchoconstriction in response to exercise – within 15 mins) 1. Change in FEV1 ≥10% post-exercise 2. If the study has used a cut-off based on performing a ROC NOTE: usually this is a 6 minute exercise challenge test.
Reference standard	Physician diagnosis of asthma based on symptoms plus an objective test
Outcomes	Diagnostic accuracy (sensitivity and specificity)

20.3 Clinical evidence

We searched for cross-sectional studies, cohort studies, case series (including both retrospective and prospective analyses) assessing the diagnostic test accuracy of exercise challenge test to identify whether the condition is present (as indicated by the reference standard) in people under investigation for asthma.

Five studies were included in the review[9,53,81,85,94] (see Table 64). Evidence from these studies is summarised in the clinical evidence profile below (Table 65). See also the study selection flow chart in Appendix D, sensitivity/specificity forest plots in Appendix J, study evidence tables in Appendix G and exclusion list in Appendix K.

One of the studies[85] was a cross-sectional study, and looked at the diagnostic accuracy of exercise test versus the reference standard in adults. This study included a population with asthma and allergic rhinitis, but was included due to the lack of evidence in people with suspected asthma. The remaining four studies[9,53,81,94] were case-series, and looked at the diagnostic accuracy of exercise test response vs. other diagnostic tests, in people who had already been diagnosed with asthma. In this case, the index test was taken to be the exercise test and the other test as the reference standard. These studies were included due to the lack of available evidence and were downgraded for indirectness. Two studies were in adults and two studies in children. These have been reported separately in the different strata.

Summary of included studies

Table 64: Summary of studies included in the review

Study	N	Index test/reference standard	Index test cut-off for positivity	Population	Age
Exercise test vs. reference standard (physician diagnosis)					
Klepac 2004[85]	35	Exercise test (6 minute treadmill) GINA definition of asthma; symptoms of asthma and/or taking asthma medication; all positive to histamine; all met EAACI definition of allergic asthma and had positive skin prick tests to at least 1 inhaled allergen. Allergic rhinitis patients met EAACI definition; negative to histamine; positive skin prick tests to at least 1 inhaled allergen	ΔFEV1 ≥10%	Asthma or allergic rhinitis	Asthma: range 15 to 48 years; allergic rhinitis: range 15 to 45 years
Exercise test vs. other tests					
Avital 2000[9]	135	Exercise test 6 minutes treadmill Methacholine PC20 <8mg	ΔFEV1%init >8.2%	Children and young adults with asthma	Mean 12.4 (3.9) years
Egglest on 1979[53]	45	Exercise test 5 minutes treadmill Methacholine	ΔFEV1 ≥18% (cut off for 2SD from mean normal response)	Young adults with asthma	Range 16 to 30 years
Kersten 2009[81]	25	Exercise challenge running with nose clip on treadmill in cold air at ice rink (temperature 1°C) for 6 minutes Mannitol challenge up to cumulative dose 6.35mg	ΔFEV1%init >15% for both tests	Children with a history of allergic asthma and exercise induced bronchoconstriction recruited from outpatient clinic; clinically stable, otherwise	Mean 12.4 (2.0) years

Study	N	Index test/reference standard	Index test cut-off for positivity	Population	Age
				healthy; FEV1 at least 70% predicted normal value; able to run on treadmill and perform reproducible spirometry	
Lin 1991[94]	22	Exercise test (10 minute treadmill) Methacholine challenge	ΔFEV1%init >20%	Stable unmedicated asthma; FEV1 >75% normal	Range 20 to 40 years

Table 65: Clinical evidence profile: Exercise test vs. Physician Dx of asthma

PEF variability (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Median Sensitivity % (range)	Median Specificity % (range)[e]	Area Under Curve (range)	Quality
ADULTS: Exercise test versus physician diagnosis										
ΔFEV1 ≥10%	1	35	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[d]	N/A[c]	26	100	-	LOW
ADULTS: Exercise test versus other tests										
Exercise test ΔFEV1 ≥18% vs. methacholine	1	45	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[e]	N/A	80	Not estimable	-	LOW
Exercise test ΔFEV1 ≥20% vs. methacholine	1	22	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[e]	N/A	43	100	-	LOW
CHILDREN 5-16 years: Exercise test versus other tests										
Cold air exercise test ΔFEV1 % init >15% vs. mannitol ΔFEV1 % init >15%	1	25	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[e]	N/A	69	92	-	LOW
Exercise ΔFEV1 ≥8.2% vs. methacholine PC20 ≤8mg/mL	1	135	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[e]	N/A	72	67	-	LOW

(a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.
(b) Inconsistency was assessed by inspection of the sensitivity / specificity values in each of the studies in relation to a high (≥75%), moderate (50-74%) or low (<50%) effect estimate. Inconsistency was considered high if the point estimates from individual studies were within two areas, and very high if they were in 3 areas or both high and low areas.
c) The judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis, where performed.
d) Population is asthma or allergic rhinitis, not suspected asthma.
e) Population with asthma and accuracy of exercise challenge test for other tests.

20.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow chart in Appendix E.

New cost-effectiveness analysis

An original health economic model was built to assess the cost-effectiveness several diagnostic pathways. Bronchial challenge test with exercise was considered as part of the diagnostic strategy, however this tests has lower accuracy and higher costs compared to direct bronchial challenge test with histamine or methacholine, therefore it was excluded from the formal analysis. Full details of the model can be found in Appendix M. A summary of the model can be found in section 18.4.

Unit cost of performing a direct bronchial challenge test with exercise on children

As an economic model was not feasible for children, the GC considered the unit cost of performing a direct bronchial challenge test to evaluate its cost-effectiveness as part of a pathway for diagnosing asthma. The NHS reference cost associated with 'Complex lung function exercise testing' (HRG code: DZ31Z) is £180.[46] A paediatric respiratory outpatient visit would also need to be considered to interpret the result; this is cost as £197 in the NHS reference costs.

20.5 Evidence statements

Clinical

- One study with 35 adults showed that exercise test (ΔFEV1 ≥10%) versus physician diagnosis has a sensitivity of 0.26 and a corresponding specificity of 1.00 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 45 adults showed that exercise test (ΔFEV1 ≥18%) has a sensitivity of 0.80 [it was not possible to calculate corresponding specificity] for predicting a positive methacholine test in people with asthma. (LOW QUALITY)
- One study with 22 adults showed that exercise test (ΔFEV1 ≥20%) has a sensitivity of 0.43 and a corresponding specificity of 1.00 for predicting a positive methacholine test in people with asthma. (LOW QUALITY)
- One study with 25 children and young people showed that exercise test versus physician diagnosis (cold air exercise test ΔFEV1 % init >15% vs. mannitol ΔFEV1 % init >15%) has a sensitivity of 0.69 and a corresponding specificity of 0.92 for diagnosing asthma in people presenting with respiratory signs and symptoms. (LOW QUALITY)
- One study with 135 children and young people showed that exercise test (ΔFEV1 ≥8.2%) has a sensitivity of 0.72 and a corresponding specificity of 0.67 for predicting a positive methacholine challenge test (PC20 ≤8mg/mL) in people with asthma. (LOW QUALITY)

Economic

- No relevant economic evaluations were identified.

20.6 Recommendations and link to evidence

Recommendations	**27. Do not offer adults (aged 17 and over) an exercise challenge test as a diagnostic test for asthma.**
Research recommendations	**1. What is the acceptability and diagnostic accuracy of objective tests that could be used to comprise a diagnostic pathway for asthma in children aged 5-16 years old (for example, exercise challenge, direct bronchial challenge with histamine or methacholine, indirect bronchial challenge with mannitol and peripheral blood eosinophil count)?**
Relative values of different outcomes	The GC was interested in the diagnostic test accuracy of airway hyper-reactivity (non-specific bronchial challenge) with exercise. There is evidence of significant misdiagnosis of asthma, with both false positive and false negative errors (see introduction to the guideline). The GC was therefore interested in both the sensitivity and specificity of diagnostic tests and prioritised studies which reported each of these outcomes for exercise challenge.
Trade-off between clinical benefits and harms	In adults, evidence was only available from one study for the accuracy of exercise challenge test in the diagnosis of asthma using the ideal reference standard. Exercise challenge test had a low sensitivity and a high specificity. In adults there are more effective indirect challenge tests that can be performed more easily and are more readily available than exercise; therefore, the GC agreed that exercise challenges do not have a routine place in the diagnosis of adult asthma. For an adult exercise test, 2 members of staff, a treadmill and monitoring facilities are required. In children, no evidence was available using the ideal population (suspected asthma, presenting with symptoms) or the ideal reference standard (physician diagnosis with an objective test). The available data suggest an exercise test has a moderate sensitivity and a moderate-to-high specificity in predicting a positive response to a mannitol or methacholine challenge test. An exercise test in a child who is able to run poses no significant harm to the child, but the acceptability of the test to a child or care giver has not been reviewed here.
Economic considerations	No economic evaluations were found which assessed the use of bronchial challenge test with exercise as part of a diagnostic pathway. An original health economic model was built to assess the cost-effectiveness of different diagnostic pathways for asthma. The model assessed the additional costs of tests against cost savings from unnecessary asthma medication and increased health outcomes from providing correct treatment. A strategy was considered that gave patients an exercise challenge test rather than a histamine/methacholine challenge test. The cost of using exercise challenge tests as opposed to other challenge tests is higher as detailed in the NHS reference costs. As the sensitivity and specificity of exercise challenge tests were lower than direct bronchial challenge tests, exercise challenge tests were a dominated option. The GC noted that the costs presented would probably be lower for children and currently they are the only challenge tests available for children. However as the costs are still high the clinical evidence was not strong enough to allow a recommendation to be made. Therefore a future research recommendation was made for children as the body of available evidence was not strong enough to dismiss them. No model was built for children and the results from the adult model could not be extrapolated.
Quality of evidence	The studies in both adults and children where exercise has been tested to make a diagnosis of asthma are very small. Larger population based studies exist in children, but these were excluded by the GC as they represented the use of exercise as a

	screening tool for the general population and thus the results could not be extrapolated to give a sensitivity or specificity in a population suspected of asthma. The studies in children in particular were also limited by the lack of comparator test (with appropriate diagnostic criteria) other than physician diagnosis and studies were excluded in this regard. • In children aged <5 years, we did not search for studies, as an exercise test is not able to be performed very well in this age-group. • In children aged 5-16 years, there were no included studies using the best reference standard (physician diagnosis with objective test) that addressed the use of exercise testing in this age group. • In adults, there was one included study (Klepac) using the best reference standard (physician diagnosis with objective test) that addressed the use of exercise testing in this age group. This study included a population with asthma and allergic rhinitis, but was included due to the lack of evidence in people with suspected asthma. It was downgraded for indirectness. The evidence was of low quality. • All other included studies assessed exercise testing vs. other tests (rather than vs. the reference standard of physician diagnosis with objective test). The evidence was of low quality.
Other considerations	There is very limited evidence concerning the use of exercise tests in those in whom there is clinical suspicion of asthma. Exercise challenge testing requires spirometry to be performed as part of the test. The GC noted that the spirometry should be performed in accordance with standard technical guidelines. Exercise testing is currently the most commonly used test in children, more commonly performed in secondary care. The GC was interested in the position of indirect challenge tests within an algorithm of diagnostic tests. However, the true efficacy, acceptability and tolerability of the test for patients under 16 years of age is not known. The GC concurred strongly that further research is urgently needed to identify the efficacy, accuracy and acceptability in the paediatric population. The GC made a high-priority research recommendation to investigate the acceptability and performance characteristics of objective tests, including exercise challenge tests, which could be used to comprise a diagnostic pathway for asthma in children aged 5-16 years old. Further details on the high-priority research recommendation made can be found in appendix N. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

21 Diagnostic summaries

This chapter discusses the diagnostic pathway for asthma. Table 66 below summarises the cut-off thresholds for the objective tests used in the diagnostic pathway.

Table 66: Thresholds of diagnostic tests

Objective test	Diagnostic cut-off for a positive test
Spirometry	FEV1/FVC ratio <70%
Bronchodilator response	Improvement in FEV1 ≥12% and increase in volume ≥200 mls
Peak expiratory flow variability	>20% variability over 2-4 weeks
FeNO	Adults 17 years and over: ≥40 ppb Children 5-16 years: ≥35 ppb
Methacholine	≤8 mg/ml (PC20)
Histamine	≤8 mg/ml (PC20)

21.1 Diagnostic algorithms

Please see section 4.1 on page 41 to 42.

21.2 Recommendations and link to evidence

Recommendations	**Diagnostic hubs** **28. Those responsible for planning diagnostic service support to primary care (for example, clinical commissioning groups) should consider establishing asthma diagnostic hubs to achieve economies of scale and improve the practicality of implementing the recommendations in this guideline.** **Diagnosing asthma and initial treatment for young children** **29. For children under 5 with suspected asthma, treat symptoms based on observation and clinical judgement, and review the child on a regular basis (see section 1.8 of the NICE guideline on asthma: diagnosis, monitoring and chronic asthma management). If they still have symptoms when they reach 5 years, carry out objective tests (see recommendations 14 to 35 and algorithm B).** **30. If a child is unable to perform objective tests when they are aged 5:** • **continue to treat based on observation and clinical judgement** • **try doing the tests again every 6 to 12 months until satisfactory results are obtained** • **consider referral for specialist assessment if the child repeatedly cannot perform objective tests and is not responding to treatment.** **Children and young people aged 5 to 16 (algorithm B)** **31. Diagnose asthma in children and young people (aged 5 to 16) if they**

	have symptoms suggestive of asthma and: • **a FeNO level of 35 ppb or more and positive peak flow variability or** • **obstructive spirometry and positive bronchodilator reversibility.** **32. Suspect asthma in children and young people (aged 5 to 16) if they have symptoms suggestive of asthma and:** • **a FeNO level of 35 ppb or more with normal spirometry and negative peak flow variability, or** • **a FeNO level of 35 ppb or more with obstructive spirometry but negative bronchodilator reversibility and no variability in peak flow readings, or** • **normal spirometry, a FeNO level of 34 ppb or less and positive peak flow variability.** **Do not rule out other diagnoses if symptom control continues to remain poor after treatment. Review the diagnosis after 6 weeks by repeating any abnormal tests and reviewing symptoms.** **33. Refer children and young people (aged 5 to 16) for specialist assessment if they have obstructive spirometry, negative bronchodilator reversibility and a FeNO level of 34 ppb or less.** **34. Consider alternative diagnoses and referral for specialist assessment in children and young people (aged 5 to 16) if they have symptoms suggestive of asthma but normal spirometry, a FeNO level of 34 ppb or less and negative peak flow variability.** **Adults aged 17 and over (algorithm C)** **35. Diagnose asthma in adults (aged 17 and over) if they have symptoms suggestive of asthma and:** • **a FeNO level of 40 ppb or more with either positive bronchodilator reversibility or positive peak flow variability, or bronchial hyperreactivity,or** • **a FeNO level between 25 and 39 ppb and a positive bronchial challenge test, or** • **positive bronchodilator reversibility and positive peak flow variability irrespective of FeNO level.** **36. Suspect asthma in adults (aged 17 and over) with symptoms suggestive of asthma, obstructive spirometry, and:** • **negative bronchodilator reversibility, and either a FeNO level of 40 ppb or more, or** • **a FeNO level between 25 and 39 ppb and positive peak flow variability, or** • **positive bronchodilator reversibility, a FeNO level between 25 and 39 ppb and negative peak flow variability.** **Do not rule out other diagnoses if symptom control continues to remain poor after treatment. Review the diagnosis after 6 to 10 weeks**

by repeating spirometry and objective measures of asthma control and reviewing symptoms.

37. Consider alternative diagnoses, or referral for a second opinion, in adults (aged 17 and over) with symptoms suggestive of asthma, and:

- **a FeNO level below 40 ppb, normal spirometry and positive peak flow variability, or**
- **a FeNO level of 40 ppb or more but normal spirometry, negative peak flow variability, and negative bronchial challenge test, or**
- **obstructive spirometry with bronchodilator reversibility, but a FeNO level below 25 ppb, and negative peak flow variability, or**
- **positive peak flow variability but normal spirometry, a FeNO level below 40 ppb, and a negative bronchial challenge test, or**
- **obstructive spirometry with negative bronchodilator reversibility, a FeNO level below 25 ppb, and a negative peak flow variability (if measured).**

Good clinical practice in asthma diagnosis

38. Record the basis for a diagnosis of asthma in a single entry in the person's medical records, alongside the coded diagnostic entry.

Summary of objective test results for adults, young people and children (over 5)

Algorithms have been produced that summarise objective testing for asthma in adults, young people and children (over 5). Table 69 summarises the objective test threshold levels.

Table 67: Positive test thresholds for objective tests for adults, young people and children (aged 5 and over)

Test	Population	Positive result
FeNO	Adults	40 ppb or more
	Children and young people	35 ppb or more
Obstructive spirometry	Adults, young people and children	FEV1/FVC ratio less than 70% (or below the lower limit of normal if this value is available)
Bronchodilator reversibility (BDR) test	Adults	Improvement in FEV1 of 12% or more and increase in volume of 200 ml or more
	Children and young people	Improvement in FEV1 of 12% or

			more
	Peak flow variability	Adults, young people and children	Variability over 20%
	Direct bronchial challenge test with histamine or methacholine	Adults	PC20 of 8 mg/ml or less
		Children and young people	n/a
	Abbreviations: FeNO, fractional exhaled nitric oxide; FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity; PC20, provocative concentration of methacholine causing a 20% fall in FEV1.		

Relative values of different outcomes	See sections 11.6, 12.6, 13.6, 16.6 and 18.6.
Trade-off between clinical benefits and harms	See sections 11.6, 12.6, 13.6, 16.6 and 18.6.
Economic considerations	See sections 11.6, 12.6, 13.6, 16.6 and 18.6.
Quality of evidence	See sections 11.6, 12.6, 13.6, 16.6 and 18.6.
Other considerations	The diagnostic algorithms assume that the objective tests have been performed correctly. **Diagnostic hubs** A recommendation was developed, aimed at clinical commissioners, to consider establishing asthma diagnostic hubs to achieve economies of scale in implementing the diagnostic algorithms. **Children younger than 5 years old:** A certain diagnosis of asthma cannot be made in this age group as no objective tests can be conducted. The GC agreed that the only viable option would be to treat the child's symptoms accordingly; however, a diagnosis of asthma could not be confirmed until the child was old enough to perform objective tests. From age 5, objective tests should be performed as per the recommendations in children aged 5-16 years. The GC discussed whether the child should remain on treatment from age 5 years and if this would affect the results of the objective tests. It was agreed that treatment should not be stopped if a child is still symptomatic on treatment. In someone who is still symptomatic on treatment, the tests would probably reveal some abnormality anyway, so objective tests should still be performed. In someone who is asymptomatic on treatment, it would be normal practice to step down treatment and, when appropriate, withdraw treatment. This should be done before performing the objective tests and reviewing the diagnosis. **Children 5-16 years old and adults and young people older than 16 years old:** For adults and children aged 5 years and older a diagnostic algorithm was built whereby the individual could arrive at different endpoints based on their test results. The following points detail what happens at each endpoint. • **Diagnose with asthma** At these points in the pathway the evidence captured from the test results was conclusive enough to make a firm diagnosis of asthma. Although at these points in the pathway there is a very small probability that the individual may not have asthma the likelihood is so low that delaying treatment any longer for the majority of patients would lead to much worse health outcomes. • **Suspect asthma but do not rule out other diagnoses if symptom control**

continues to remain poor after treatment and review the diagnosis after 6 weeks by repeating any abnormal tests and reviewing symptoms

At these points in the pathway the evidence gathered from the objective tests suggests asthma as the most likely diagnosis but is not unequivocal; there is a possibility that the symptoms may be derived from another cause. Therefore fixing a diagnosis of asthma could lead to a few cases of long-term misdiagnosis. The GC agreed that it would be best to commence anti-asthma treatment but to not regard the diagnosis as fixed, to monitor the response carefully, and to have a low threshold for investigating for other conditions. Doing so will help further minimise the number of false positives produced by the algorithm.

The GC agreed that inhaled corticosteroids was the appropriate anti-asthma treatment to commence in these circumstances, and that review after 6 weeks was an appropriate interval to allow this treatment to take effect before re-assessing.

- **Consider alternative diagnoses and referral for specialist assessment**

At these points in the pathway there is significant evidence to rule-out the diagnosis of asthma. Ruling out the diagnosis of asthma will allow investigation to establish an alternative cause of the presenting symptoms. Although at these points in the pathway there is a very small probability that the individual may have asthma the likelihood is low and delaying investigative procedures for other conditions for the majority of patients would lead to much worse health outcomes.

- **Refer for specialist opinion (children 5-16 only)**

At this point in the algorithm the child would have an obstructive spirometry but a negative bronchodilator reversibility test and a negative FeNO. This could indicate that the child has a very unique form of asthma or another, potentially serious, condition. The GC therefore considered in this circumstance that a specialist opinion would be needed to ensure the child's condition was effectively diagnosed and managed.

- **Consider alternative diagnosis or refer to specialist (adults and young people older than 16 years old only)**

At this point in the pathway the individual is unlikely to have asthma, but they have reversible airways. A possible diagnosis for this individual could be COPD as people with this condition can produce positive results from a bronchodilator reversibility test. However if the clinician cannot rule-in an alternative diagnosis then specialist referral may be the best option to ensure the individual receives the best management.

People diagnosed with asthma

The GC discussed that currently, the basis on which a diagnosis of asthma is made is not well documented in the patient's medical records. It was agreed that it was important to record the evidence on which the diagnosis was based in the patient's medical records. A recommendation was made to reflect this.

22 Monitoring asthma control

Chapters 23 to 30 review the clinical and cost-effectiveness of interventions used to monitor asthma control.

23 Monitoring: Symptom scores and questionnaires

23.1 Introduction

Published evidence suggests that both patients and clinicians tend to underestimate asthma severity and overestimate asthma control when simply asking a patient 'How is your asthma?' As a result of this, multiple different asthma questionnaires have been developed both to assess asthma-related quality of life and asthma control. These questionnaires have primarily been validated in patients with mild to moderate asthma. Questionnaires have the potential to be used to aid the monitoring of asthma as minimally clinically important differences have been established for the majority of the questionnaires.

23.2 Review question: In people with asthma, what is the clinical and cost-effectiveness of using symptom scores / diaries or validated questionnaires measuring symptom control (eg ACT, ACQ, CACT, RCP 3 questions) and / or health related quality of life (eg AQLQ, PAQLQ) to monitor asthma?

For full details see review protocol in Appendix C.

Table 68: PICO characteristics of review question

Population	People with asthma (defined as physician Dx with objective test or recurrent persistent wheeze in <5 years). If insufficient evidence is found we will consider evidence from studies where asthma is defined as physician Dx only or on asthma treatment. All ages, stratified into the following 3 different groups: • Children (1-<5 years old) • Children/young people (5-16 years old) • Adults (>16 years old)
Intervention(s)	Monitoring the following, and using the outcomes of scores/questionnaires to adjust management/therapy according to physician decision or personalised treatment plan: • Symptom scores or diaries • Symptom/control questionnaires • Quality of life questionnaires (asthma specific)
Comparison(s)	Comparison of adjustment of asthma therapy based on symptom scores or questionnaires to: • Usual care: e.g. clinical symptoms (with/without spirometry/PEF) according to guidelines (including BTS/SIGN, GINA) Comparison of adjustment of asthma therapy based on: • Symptom scores or diaries vs. questionnaires • Control questionnaire vs. other control questionnaire • QOL questionnaire vs. control questionnaire
Outcomes	Critical outcomes • Mortality • Unscheduled healthcare utilisation (ED/A&E visit; hospital admissions; GP out of hours or walk-in centre) • Exacerbations (defined as need for course of oral steroids)

	• Asthma control questionnaires (ACT; CACT; ACQ; PACQ; RCP-3) • QoL (AQLQ; pAQLQ; St George's respiratory questionnaire) Important outcomes • Lung function (FEV1, PEF) • Symptoms (annual symptom free days) • Dose of regular asthma therapy / preventer medication (ICS dose) • Rescue medication (SABA use) • Time off school or work
Study design	RCTs

23.3 Clinical evidence

We searched for randomised trials comparing the effectiveness of monitoring using asthma control questionnaires, QOL questionnaires or symptom diaries vs monitoring according to usual care to guide asthma treatment and management. The asthma control questionnaires and QOL questionnaires considered in the review are summarised in Table 69.

Three studies (four papers) were included in the review[107,143,184] these are summarised in Table 70 below. See also the study selection flow chart in Appendix D, study evidence tables in Appendix G, forest plots in Appendix J, GRADE tables in Appendix I and excluded studies list in Appendix K. All relevant clinical studies used monitoring with asthma control questionnaires to guide management or treatment. No relevant clinical studies were identified using monitoring with asthma-specific QOL questionnaires to guide management or treatment. No relevant clinical studies were identified using monitoring with symptom scores or diaries vs. usual care. Studies comparing monitoring with symptoms scores or diaries with monitoring using PEF or FeNO are reported in the respective reviews.

In children age 5-16 years, evidence comparing monitoring with questionnaires vs. usual care to guide ongoing management was available from one study[143] summarised in the clinical evidence summary (Table 71). This study was in children with uncontrolled asthma and used the Asthma Control Questionnaire (ACQ) to guide treatment according to a treatment algorithm. Outcomes are reported at both <6months and ≥6 months.

In adults age >16 years, evidence comparing monitoring with questionnaires vs. usual care to guide ongoing management was available from two studies[107,184] summarised in the clinical evidence summary (Table 72). One study used the ACQ to guide treatment according to a treatment algorithm, the other study used the Asthma Control Test (ACT) to guide pharmacist care.

No relevant clinical studies comparing monitoring with questionnaires vs. usual care to guide ongoing management were available in children age 1-<5 years old.

Table 70 also summarises additional education interventions received by the intervention or comparator groups. In studies where both the intervention and comparator groups receive education, the monitoring intervention may show reduced effectiveness as the control group might also be expected to show improvement due to the education (saturation effects).

Table 69: Summary of questionnaires

Questionnaire	Reference	Number of items and Scale	Recall period	Established MID	Population for intended use
Asthma control questionnaires					
Asthma Control Test (ACT)	Developed by QualityMetric Inc. and GSK Nathan et al. Development of the asthma control test: a survey for assessing asthma control. J Allergy Clin Immunol. 2004 Jan;113(1):59-65[112].	5-items (activity limitations, shortness of breath, nocturnal symptoms, rescue medication, overall control in past 4 weeks) Each scored from 1 (worst) to 5 (best) Range 5-25 (better indicated by higher values)	Past 4 weeks	3.0	Adolescents and adults (12+ years)
Paediatric ACT (CACT)	Developed by QualityMetric Inc. and GSK Liu et al. Development and cross-sectional validation of the Childhood Asthma Control Test. J Allergy Clin Immunol. 2007; 119, 817-825[98].	7-item scale (4 child-reported and 3 caregiver reported) The child-completed items use a 4-point response scale and the caregiver-completed items use a 6-point response scale) Range 0-27 (better indicated by higher values)	Past 4 weeks	None established	Children 4-11 years
Asthma Control Questionnaire (ACQ)	Juniper et al. Development and validation of a questionnaire to measure asthma control. Eur Respir J 1999; 14: 902-907[79]	7 items (specific symptoms, timing of symptoms, activity limitation, rescue medications, lung function) Each scored from 0 (best) to 6 (worst) Range 0-6 (better indicated by lower values)	Past week	0.5 (score of 1.5 best discriminator between controlled and uncontrolled)	Children and adults (6+ years)
RCP 3 questions	Pearson MB (ed). Measuring clinical outcomes in asthma: a patient-focused approach. London: Royal College of	3 items Each item scored 0 (no) or 1 (yes) (better indicated by lower values)	Past week (or month)	None established	

Questionnaire	Reference	Number of items and Scale	Recall period	Established MID	Population for intended use
	Physicians; 1999.				
Asthma QOL questionnaires					
Asthma Quality of Life Questionnaire (AQLQ)	Juniper et al. Measuring quality of life in asthma. Am Rev Respir Dis 1993; 147: 832-838[77]	32 items 7 point scale, 1-7 (better indicated by higher values)	Past 2 weeks	0.5	Adults
Mini AQLQ	Juniper et al. Eur Respir J 1999; 14: 32-38[75]	15 items 7 point scale, 1-7 (better indicated by higher values)	Past 2 weeks	0.5	Adults
Paediatric AQLQ (child)	Juniper et al. Measuring quality of life in children with asthma. Quality of Life Research 1996; 5: 35-46[76]	23 items 7 point scale, 1-7 (better indicated by higher values)	Past week	0.5	Paediatrics
Paediatric AQLQ (carer PACQLQ)	Juniper et al. Measuring quality of life in the parents of children with asthma. Quality of Life Research 1996; 5: 27-34	13 items 7 point scale, 1-7 (better indicated by higher values)	Past week	0.5	Caregivers

Table 70: Summary of studies included in the review

Study	Intervention	Comparison	Population	Outcomes	
MEER 2009[184] RCT	Internet based self-management (treatment plan, online education and communication with nurse). Treatment algorithm tells patients how to adjust their treatment according to weekly ACQ score: - Four consecutive scores ≤0.5 : decrease treatment according to plan - Two scores >0.5 but <1: increase treatment	Asthma care according to Dutch guidelines (based on GINA), recommend medical review and treatment adjustment every 2 to 4 weeks in unstable asthma and once or twice yearly for controlled asthma.	ADULTS. Phys Dx asthma and ICS for at least 3 months in the previous year.	• QOL • Exacerbations • Asthma control • Lung function • Symptoms • ICS use	Both groups received a prior basic education session (but intervention group received additional web-

Study	Intervention	Comparison	Population	Outcomes	
	according to plan - One score ≥1 but <1.5: immediately increase according to plan - One score >1.5: immediately increase treatment and contact nurse.				based and face-to-face education
MEHUYS 2008[107] RCT	Pharmacist advice based on ACT score at 0, 1 and 3 month: -ACT <15: immediate referral to GP or specialist -ACT 15-19: review inhaler technique and check controller adherence -ACT >19: no advice, inform patient asthma is well-controlled	Usual pharmacist care	ADULTS with a prescription for asthma medication and treated for asthma >12 months, using controller meds	• QOL • Exacerbations • UHU • Asthma control • Rescue medication	ONLY INTERVENTION GROUP - additional personal education from the pharmacist at the start
RIKKERS 2012[143] RCT	Same as VAN DER MEER 2009	Same as VAN DER MEER 2009	CHILDREN 12-18 years, asthma not well controlled asthma as assessed by ACQ>0.75 and/or ATAQ <1.0	• QOL • Exacerbations • Asthma control • Lung function • Symptoms • ICS use	Same as VAN DER MEER 2009

Table 71: Clinical evidence summary: Children (5-16 years) with uncontrolled asthma: Monitoring questionnaires + treatment vs UC + treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with UC + treatment	Risk difference with Children with uncontrolled asthma: Monitoring control + treatment (95% CI)
QOL (< 6months) PAQLQ. Scale from: 1 to 7.	90 (1 study) 3 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean QOL in the control groups was **5.68**	The mean QOL in the intervention groups was **0.4 higher** (0.17 to 0.63 higher)
QOL (≥ 6months) PAQLQ. Scale from: 1 to 7.	90 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[1] due to risk of bias		The mean QOL in the control groups was **6.05**	The mean QOL in the intervention groups was **0.05 lower** (0.5 lower to 0.4 higher)
Exacerbations (≥ 6months) Course of OCS	75 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,3] due to risk of bias, imprecision	**RR 1.14** (0.41 to 3.22)	**Moderate** **150 per 1000**	**21 more per 1000** (from 89 fewer to 333 more)

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects	
Asthma control (< 6months) ACQ. Scale from: 0 to 6.	90 (1 study) 3 months	⊕⊕⊖⊖ **LOW**[1] due to risk of bias		The mean asthma control in the control groups was **1.19**	The mean asthma control in the intervention groups was **0.32 lower** (0.56 to 0.08 lower)
Asthma control (≥ 6months) ACQ. Scale from: 0 to 6.	90 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[1] due to risk of bias		The mean asthma control in the control groups was **0.79**	The mean asthma control in the intervention groups was **0.05 lower** (0.35 lower to 0.25 higher)
Lung function (< 6months) FEV1 L	90 (1 study) 3 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean lung function in the control groups was **2.9 L**	The mean lung function in the intervention groups was **0.23 higher** (0.08 to 0.38 higher)
Lung function (≥ 6months) FEV1 L	90 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean lung function in the control groups was **3.12 L**	The mean lung function in the intervention groups was **0.1 higher** (0.11 lower to 0.31 higher)
Symptom free days (< 6months) % over 2 weeks . Scale from: 0 to 100.	90 (1 study) 3 months	⊕⊖⊖⊖ **VERY LOW**[1,4] due to risk of bias, imprecision		The mean symptom free days in the control groups was **76 %**	The mean symptom free days in the intervention groups was **1.5 lower** (14.5 lower to 11.5 higher)
Symptom free days (≥ 6months) % over 2 weeks.	90 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,4] due to risk of bias, imprecision		The mean symptom free days in the control groups was **80 %**	The mean symptom free days in the intervention groups was **4 higher** (9.7 lower to 17.7 higher)
ICS use (< 6months) mean daily dose ug	90 (1 study) 3 months	⊕⊖⊖⊖ **VERY LOW**[1,4] due to risk of bias, imprecision		The mean ICS use in the control groups was **334 ug**	The mean ICS use in the intervention groups was **14 higher** (79 lower to 107 higher)
ICS use (≥ 6months) mean daily dose ug	90 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,4] due to risk of bias, imprecision		The mean ICS use in the control groups was **265 ug**	The mean ICS use in the intervention groups was **14 higher** (75 lower to 103 higher)

[1] The majority of the evidence was from studies at very high risk of bias
[2] 95% CI crosses one MID
[3] 95% CI for the absolute effect crosses one MID
[4] 95% CI crosses both MIDs

Table 72: Clinical evidence summary: Adults and young people (>16 years) overall: Monitoring questionnaires + treatment vs UC + treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with UC + treatment	Risk difference with Adults overall: Monitoring control + treatment (95% CI)
QOL (≥ 6months)	333	⊕⊕⊕⊖		The mean QOL in the control groups	The mean QOL in the intervention groups

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects	
AQLQ. Scale from: 1 to 7.	(2 studies) 6-12 months	**MODERATE**[1] due to risk of bias		was **5.89**	was **0.32 higher** (0.17 to 0.47 higher)
Exacerbations (≥ 6months) course of OCS	183 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision	**HR 1.18** (0.51 to 2.73)	**Moderate** **109 per 1000**	 **18 more per 1000** (from 52 fewer to 161 more)
Exacerbations (≥ 6months) ER, hospitalisation or OCS	333 (2 studies) 6-12 months	⊕⊖⊖⊖ **VERY LOW**[1,3,4] due to risk of bias, indirectness, imprecision	**RR 1.1** (0.61 to 1.99)	**Moderate** **112 per 1000**	 **11 more per 1000** (from 44 fewer to 111 more)
UHU (≥ 6months) ER or hospitalisation	150 (1 study) 6 months	⊕⊖⊖⊖ **VERY LOW**[4,5] due to risk of bias, imprecision	**RR 0.17** (0.02 to 1.46)	**Moderate** **71 per 1000**	 **59 fewer per 1000** (from 70 fewer to 33 more)
Asthma control (< 6months) ACT. Scale from: 5 to 25.	183 (1 study) 3 months	⊕⊕⊖⊖ **LOW**[5] due to risk of bias		The mean asthma control in the control groups was **20.0**	The mean asthma control in the intervention groups was **0.3 higher** (0.73 lower to 1.33 higher)
Asthma control (≥ 6months) ACQ . Scale from: 0 to 6.	183 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[1,6] due to risk of bias, imprecision		The mean asthma control in the control groups was **1.04**	The mean asthma control in the intervention groups was **0.47 lower** (0.64 to 0.3 lower)
Asthma control (≥ 6months) ACT. Scale from: 5 to 25.	150 (1 study) 6 months	⊕⊕⊖⊖ **LOW**[5] due to risk of bias		The mean asthma control in the control groups was **19.7**	The mean asthma control in the intervention groups was **0.5 higher** (0.86 lower to 1.86 higher)
Lung function (≥ 6months) FEV1 L	183 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[1,6] due to risk of bias, imprecision		The mean lung function in the control groups was **3.12 L**	The mean lung function in the intervention groups was **0.25 higher** (0.03 to 0.47 higher)
Symptom free days (≥ 6months) % over 2 weeks. Scale from: 0 to 100.	183 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[1,6] due to risk of bias, imprecision		The mean symptom free days in the control groups was **51.8 %**	The mean symptom free days in the intervention groups was **10.9 higher** (0.05 to 21.75 higher)
ICS use (≥ 6months) mean daily dose ug	183 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean ICS use in the control groups was **470 ug**	The mean ICS use in the intervention groups was **57 higher** (38 lower to 152 higher)
Rescue medication (< 6months) puffs/day	183 (1 study) 3 months	⊕⊕⊖⊖ **LOW**[1,6] due to risk of bias, imprecision		The mean rescue medication in the control groups was **1.3 puffs/day**	The mean rescue medication in the intervention groups was **0.62 lower** (1.21 to 0.03 lower)
Rescue medication (≥ 6months) puffs/day	150 (1 study) 6 months	⊕⊕⊕⊖ **MODERATE**[1] due to risk of bias		The mean rescue medication in the control groups was **0.9 puffs/day**	The mean rescue medication in the intervention groups was **0.23 lower** (0.66 lower to 0.2 higher)

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects

[1] The majority of the evidence was from studies at high risk of bias
[2] 95% CI crosses both the MIDs
[3] Evidence from one study with an indirect outcome (ER, hospitalisation or OCS)
[4] 95% CI for the absolute effect crosses one MID
[5] The majority of the evidence was from studies at very high risk of bias
[6] 95% CI crosses one MID

23.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow chart in Appendix E.

23.5 Evidence statements

Clinical

CHILDREN (5-16 years)

- No evidence was identified on mortality and unscheduled healthcare utilisation outcomes.
- Monitoring asthma control questionnaires vs usual monitoring was considered a clinically important benefit for QOL, asthma control questionnaire score and lung function (FEV1) at <6 months (evidence for all outcomes from 1 study, N=90, low and very low quality)
- Monitoring asthma control questionnaires vs usual monitoring resulted in no clinically important difference for QOL, asthma control questionnaire score, and lung function (FEV1), all at ≥6 months, and for symptom free days and ICS use, both at <6 months and ≥6 months (all evidence from 1 study, N=90, low and very low quality).
- Monitoring asthma control questionnaires vs usual monitoring resulted in an borderline clinically important difference for exacerbations at ≥ 6 months (1 study, N=75, very low quality)

ADULTS (>16 years)

- No evidence was identified for mortality.
- Monitoring asthma control questionnaires vs usual monitoring was considered a clinically important benefit for QOL (2 studies, N=333, moderate quality), UHU (1 study, N=150, very low quality), asthma control questionnaire score measured on the ACQ (1 study, N=183, low quality), lung function and symptom-free days (both from 1 study, N=183, low quality), all at ≥6 months and for use of rescue medication and <6 months and ≥6 months (low and moderate quality).
- Monitoring asthma control questionnaires vs usual monitoring resulted in no clinically important difference for Asthma control questionnaire score at < 6 months and ≥ 6 months measured on the ACT (1 study, low quality) and for ICS use at ≥6 months (1 study, N=183, very low quality).
- Monitoring asthma control questionnaires vs usual monitoring resulted in a borderline clinically important difference for exacerbations (assessed with course of OCS) (2 studies, N=333, very low quality) and exacerbations (assessed with ER, hospitalisation or course of OCS) (2 studies, N=333, very low quality), both at ≥ 6 months.

Economic

- No relevant economic evaluations were identified.

23.6 Recommendations and link to evidence

Recommendations	**39.Monitor asthma control at every review. If control is suboptimal:** • **confirm the person's adherence to prescribed treatment in line with recommendations 1.2.1, 1.2.2 and 1.2.3 on assessing adherence in the NICE guideline on medicines adherence**

	• **review the person's inhaler technique** • **review if treatment needs to be changed** • **ask about occupational asthma (see recommendation 10) and/or other triggers, if relevant.** **40. Consider using a validated questionnaire (for example, the Asthma Control Questionnaire or Asthma Control Test) to monitor asthma control in adults (aged 17 and over).**
Research recommendations	**1. What is the clinical and cost effectiveness of using validated quality of life questionnaires and the RCP 3 questions as tools to monitor asthma control in adults aged 17 years and over?** **2. What is the clinical and cost effectiveness of using validated paediatric questionnaires to monitor asthma control in children aged 5-16 years old with asthma?**
Relative values of different outcomes	The GC considered the following outcomes as critical for this review: mortality, quality of life, exacerbations, asthma control measured by questionnaires and unscheduled healthcare utilisation. The GC noted that exacerbations of asthma can lead to both unscheduled healthcare utilisation (ED visits, out-of-hours GP surgery visits or hospitalisation) and the need for a course of OCS. Therefore, the GC considered exacerbations requiring OCS treatment and exacerbations requiring unscheduled healthcare utilisation separately. The GC also considered the following important outcomes: lung function (FEV1), symptoms (symptom free days), dose of regular asthma therapy (ICS dose), rescue medication (SABA dose) and time off school or work. These outcomes were considered as important measures of asthma control for the patient. The GC acknowledged that regular asthma preventer therapy (ICS) is often under prescribed / used, and rescue medications (SABA) may be overprescribed / used[144].
Trade-off between clinical benefits and harms	For asthma control questionnaires, RCT evidence was identified for the ACQ and the ACT, but not the RCP 3 questions. In adults, monitoring asthma control questionnaires vs usual monitoring was considered a clinically important benefit for QOL, UHU and asthma control questionnaire score measured using the ACQ at >6 months. Evidence showed that there may be no clinically important difference between monitoring asthma control questionnaires vs usual monitoring in the asthma control questionnaire score measured using the ACT. There was evidence of an increase in the rate of exacerbations in the monitoring group; however, it was unclear if this represented a clinically important harm. The GC acknowledged that on some occasions OCS may be of benefit by preventing a more severe asthma attack requiring UHU. Evidence from one study showed that monitoring questionnaire scores resulted in a clinically important benefit in lung function, use of rescue medication and symptom-free days. These outcomes were particularly important to the GC and patient. The evidence suggested there may be no clinically important benefit in ICS use. The GC opinion was that asthma control is better captured using a questionnaire and the evidence suggests some longer-term benefit. In children, monitoring asthma control questionnaires vs usual monitoring was considered a clinically important benefit for QOL and asthma control questionnaire

	score at <6 months, but not at > 6 months. There was evidence of an increase in the rate of exacerbations in the monitoring group; however, it was unclear if this represented a clinically important harm. The GC acknowledged that use of OCS in exacerbations may be of benefit by preventing a more severe asthma attack requiring UHU. Evidence from one study showed that monitoring questionnaire scores resulted in a clinically important benefit in lung function at < 6 months but not at >6 months. The GC did not think the difference in symptom-free days and ICS use represented a clinically important difference. The evidence was suggestive of a benefit at <6 months follow-up, but not at >6 months follow-up. The GC acknowledged that good asthma control scores are associated with better outcomes. Due to the uncertainty of a longer term benefit, the GC recommended a future research recommendation in children on the effectiveness of monitoring asthma control using validated questionnaires. No evidence was identified in children aged 1-<5 years old.
Economic considerations	No economic evidence was found on symptoms scores. As the individual will be attending an annual asthma review anyway, the additional cost of monitoring asthma control with a validated questionnaire within this review will be negligible. However there may be additional costs to consider from using these questionnaires to increase or decrease medication usage. The GC's decision to recommend these questionnaires was based on clinical evidence showing that they were clinically effective and therefore changes in medication were providing benefit. Uncertainty in this evidence and the cost-effectiveness led the GC to make a 'consider' rather than 'offer' recommendation.
Quality of evidence	In adults, the evidence for the important and critical outcomes was of low and very low quality by GRADE criteria, with the exception of the QOL outcome of moderate quality. Only one study contributed to the evidence for the majority of outcomes, and all studies were of small sample size. The GC considered that the treatment algorithm in the Meer 2009 study was quite intensive. The GC did not think an ACQ score between 0.5-1.0 would always warrant an increase in treatment. The strength of the recommendation was based on the GC opinion (not the evidence alone) that a questionnaire should be used to capture symptom and control information. In children, all evidence was of very low and low quality. One study contributed to the evidence for each outcome and the majority of studies were of small sample size. Again, the GC considered that the treatment algorithm in the Rikkers 2012 study was quite intensive. The GC did not think an ACQ score between 0.5-1.0 would always warrant an increase in treatment. The GC noted that at 12 months there was an increase in asthma control in the comparator group, perhaps due to additional education or participation in a research study. This may have masked a benefit in the intervention group. The GC noted the different populations within the studies. In children, the participants were uncontrolled at the start of the study, whereas in adults, the participants were controlled at the start of the study. No evidence was identified in children aged 1-<5 years old.
Other considerations	The monitoring interventions reported in the studies were complex interventions involving different treatment algorithms, not the effect of control questionnaires in isolation. The GC noted that it was hard to look at the monitoring intervention outside of the clinical care provided. It was noted that some studies included additional educational components, and that the effect of control questionnaires may be saturated due to improved outcomes in the control group. The GC noted that

	the studies looked at the use of questionnaires within a particular context (for example, a pharmacist monitoring programme or treatment algorithm). Whilst the GC did not look at the individual evidence from prognostic studies of asthma control questionnaire scores as a risk factor for future outcomes, the GC was aware of and discussed the existence of prognostic evidence within the broader literature base showing that poor asthma control scores predict future risk. The GC consensus was that not enough information is gathered from just asking, 'how is your asthma today', and that asthma control questionnaires should be used at every asthma review. The GC discussed the NICE quality standards which recommend an asthma review annually. The GC was aware of validation studies for the QOL questionnaires and for the RCP 3 questionnaires, but not of any RCT studies of monitoring using these questionnaires to guide treatment. The GC made a future research recommendation for the use of QOL questionnaires and the RCP 3 questions for monitoring asthma control. The GC also made a future research recommendation to investigate the clinical and cost-effectiveness of using validated paediatric questionnaires to monitor asthma control in children aged 5-16 years old (please see appendix N for the full list of research recommendations made). Finally, the GC also noted that although they were considering evidence around methods of monitoring, any monitoring is futile unless it prompts action when the process reveals inadequate asthma control. This would apply to any of the monitoring techniques, but since questionnaires were considered first in the guideline development schedule, the GC agreed to add a consensus recommendation here, outlining the actions which might be necessary if a person's control has slipped, The GC agreed the factors that should be checked in this situation, as part of good practice:- to confirm the person's adherence to prescribed therapy for example using information on prescription refills; to review the person's inhaler technique; to review if treatment needs to be changed and to, if relevant, ask about occupational asthma and/or other triggers. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. One study was identified which the GC agreed did not suggest a change in the recommendations was warranted.

24 Monitoring: Lung function tests

24.1 Introduction

The aim of the review was to assess the clinical and cost-effectiveness of lung function measurements in the monitoring of asthma.

Airflow obstruction is a recognised characteristic abnormality in asthma. Guidelines for the management of asthma in children and adults emphasise the importance of objective assessment of lung function, in particular airflow obstruction.

Lung function does not correlate strongly with asthma symptoms in adults or children and many people with asthma are poor perceivers of changes in airway calibre. Evidence of airways obstruction is a poor prognostic factor for the outcome of asthma and a low FEV1 identifies patients at risk of asthma exacerbations, independent of symptom levels, especially if FEV1 is <60% predicted.

FEV1 is considered to be the "gold standard" measurement of airways obstruction due to its accurate, well standardised measurements, repeatability and reliable reference values.

PEF may provide some useful information however a normal PEF does not rule out significant airways obstruction and the variation in normal values, particularly in healthy children, is large, making comparison to reference values less helpful.

While the role of spirometry in the diagnosis and initial assessment of asthma is well established, its optimal role in the ongoing monitoring of asthma is still an area of uncertainty.

24.2 Review question: In people with asthma, what is the clinical and cost-effectiveness of using measures of pulmonary function assessing asthma control (for example, spirometry and peak expiratory flow) to monitor asthma?

For full details see review protocol in Appendix C.

Table 73: PICO characteristics of review question

Population	People with asthma (defined as physician Dx with objective test). If insufficient evidence is found we will consider evidence from studies where asthma is defined as physician Dx only or on asthma treatment (give details of how Dx was made). Not including severe asthma. All ages, stratified into the following 2 different groups: • Children/young people (5-16 years old) • Adults (>16 years old)
Intervention(s)	Monitoring lung function using the following tests, and using the outcomes to adjust management/therapy according to physician decision or personalised treatment plan (use of other interventions to be included if equal access in each group, eg both groups receive education in addition to monitoring): • Spirometry (FEV1; FEV1/FVC; Flow loop measures) • PEF
Comparison(s)	Comparison of adjustment of asthma therapy based on lung function tests to: • Usual care: eg clinical symptoms according to guidelines (including BTS/SIGN, GINA) • Asthma control or QOL questionnaires

	Comparison of adjustment of asthma therapy based on: • Spirometry versus PEF
Outcomes	Critical outcomes: • Mortality • Unscheduled healthcare utilisation (ED/A&E visit; hospital admissions; GP out of hours or walk-in centre) • Exacerbations (defined as need for course of oral steroids) • Asthma control questionnaires (ACT; CACT; ACQ; PACQ; RCP-3) • QoL (AQLQ; pAQLQ; St George's respiratory questionnaire) Important outcomes: • Lung function (FEV1, PEF) • Symptoms (annual symptom free days) • Dose of regular asthma therapy / preventer medication (ICS dose) • Rescue medication (SABA use) • Time off school or work
Study design	RCT

24.3 Clinical evidence

We searched for randomised trials comparing the effectiveness of monitoring lung function using spirometry or peak expiratory flow measures versus monitoring according to usual care (for example clinical symptoms) to guide asthma treatment and management.

Eleven studies were included in the review[3,24,29,38,39,80,93,100,180,192,199], summarised in Table 74 and Table 75. See also the study selection flow chart in Appendix D, study evidence tables in Appendix G, forest plots in Appendix J, GRADE tables in Appendix I and excluded studies list in Appendix K. Evidence in adults was available from eight studies[100 3,24,29,38,39,80,180]. Evidence in children and young people was available from four studies[29,93,192,199].

All studies were of self-management, with the action plans based on PEF readings versus action plans based on symptoms. Studies investigating the effectiveness of self-management plans based on PEF versus no self-management (the effectiveness of self-management plans) were excluded. No studies were identified monitoring spirometry. In three studies, the action plan in the intervention group was based on PEF and symptoms[100,192,199]. In the remaining studies, the action plan in the intervention group was based on PEF alone. One study also incorporated adjustment by the GP in addition to self-management[100].

A Cochrane systematic review was identified and included [20]. Studies included in this review were included and data extracted separately to incorporate additional studies and outcomes from the protocol.

Table 74: Summary of studies included in the review: Adults

Study	Intervention/comparison	Population	Outcomes	Follow up
Adams 2001[3]	Self-management action plan based on PEF / Self-management action plan based on symptoms	Adults 17-70 years Physician's diagnosis of asthma defined by ATS	Hospitalisation (mean days); ED visits (mean days); FEV1; (time off work) mean days	12 months
Buist 2006[24]	Self-management action plan based on PEF / Self-management action plan based on symptoms	50-92 years, moderate to severe asthma Physician-diagnosed asthma and had medication use suggestive of moderate-to-severe asthma; bronchodilator reversibility (> 8% of baseline FEV1)	QOL (AQLQ dichotomised); total asthma-related health care utilisation	2 years
Charlton 1990[29]	Self-management action plan based on PEF / Self-management action plan based on symptoms	Adults and children stratum separately Patients on the repeat prescribing register who were receiving prophylactic treatment for asthma	Exacerbations (OCS); rescue medications	12 months
Cote 1997[38]	Self-management action plan based on PEF / Self-management action plan based on symptoms	16 years and over, moderate to severe asthma The diagnosis had to be confirmed by either a documented reversibility greater than 15% in FEV1 or a methacholine PC20<8mg/ml	Exacerbations (OCS); Hospitalisation (mean events); ED visits (mean events); time off work (mean days)	12 months
Cowie 1997[39]	Self-management action plan based on PEF / Self-management action plan based on symptoms	Adults and adolescents Adult and adolescent patients who had received urgent treatment for their asthma in the preceding 12 months and used asthma medication	Visits for urgent treatment; hospital admissions	6 months
Kaya 2009[80]	Self-management action plan based on PEF / Self-management action plan based on symptoms	Adults Patients with persistent asthma receiving care for at least one year in asthma clinic	FEV1; PEF	12 months
Lopez-	Self-management action plan based on PEF,	17-65 years	Hospitalisation; ED visits;	12

Study	Intervention/comparison	Population	Outcomes	Follow up
Vina 2000[100]	symptoms & medications (and additional GP visits with adjustment based on spirometry, PEF and symptoms) / Self-management action plan based on symptoms (and additional GP visits with adjustment based on spirometry and symptoms)	ATS definition of asthma, with symptoms of episodic wheezing, cough and shortness of breath responding to bronchodilators, and documented BDR (>20% increase in FEV1 or PEF) or BHR (in patients with normal spirometry).	FEV1; time off work	months
Turner 1998[180]	Self-management action plan based on PEF / Self-management action plan based on symptoms	18-55 years Moderate to moderately severe asthma. BHR (methacholine <8mg/ml) and daily ICS.	QOL (SD not reported); Exacerbations (OCS); Hospitalisation; ED visits; unscheduled doctor visits; FEV1; PEF; time off work	6 months

Table 75: Summary of studies included in the review: Children

Study	Intervention/comparison	Population	Outcomes	Follow up
Charlton 1990[29]	Self-management action plan based on PEF / Self-management action plan based on symptoms	Adults and children stratum separately Patients on the repeat prescribing register who were receiving prophylactic treatment for asthma	Exacerbations (OCS); rescue medications	12 months
Letz 2004[93]	Self-management action plan based on PEF / Self-management action plan based on symptoms	6-12 years Diagnosed with mild to severe persistent asthma (symptoms at least 2 times per week, FEV1<80% and FEV1 or PEF variability 12% or greater). Diagnosis made on the basis of history, examination and pre/post-BD lung function testing.	Exacerbations (OCS)	3 months
Wensley 2004[192]	Self-management action plan based on PEF and symptoms / Self-management action plan based on symptoms	Children 7-14 years Physician diagnosis asthma and at least step 2 of BTS treatment guidelines.	Hospitalisation; ED visits; unscheduled doctor visits; FEV1; PEF; symptom free days; time off work	12 weeks
Yoos 2002[199]	Self-management action plan based on PEF and symptoms / Self-management	Children 6-19 years	FEV1	3 mont

Study	Intervention/comparison	Population	Outcomes	Follow up
	action plan based on symptoms	Diagnosis of asthma and more than 3 asthma-related healthcare visits in the past 12 months		hs

Table 76: Clinical evidence summary: Adults: Monitoring PEF versus symptom monitoring

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with PEF versus symptoms monitoring: adults (95% CI)
QOL ≥6 months AQLQ increase >0.5 points	262 (1 study) 2 years	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision	**RR 0.99** (0.73 to 1.35)	**Moderate** **391 per 1000**	**4 fewer per 1000** (from 106 fewer to 137 more)
QOL ≥6 months AQLQ decrease >0.5 points	262 (1 study) 2 years	⊕⊖⊖⊖ **VERY LOW**[1,2,3] due to risk of bias, imprecision	**RR 1.39** (0.67 to 2.88)	**Moderate** **86 per 1000**	**34 more per 1000** (from 28 fewer to 162 more)
Exacerbation ≥6 months need for OCS	152 (2 studies) 6-12 months	⊕⊖⊖⊖ **VERY LOW**[1,3,4] due to risk of bias, inconsistency, imprecision	**RR 1.28** (0.29 to 5.57)	**Moderate** **169 per 1000**	**47 more per 1000** (from 120 fewer to 772 more)
Exacerbations ≥6 months number of OCS courses	95 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[1] due to risk of bias		The mean exacerbations ≥6 months in the control groups was **0.9 courses of OCS**	The mean exacerbations ≥6 months in the intervention groups was **0.20 lower** (0.74 lower to 0.34 higher)
UHU ≥6 months Total asthma-related health care utilisation	294 (1 study) 2 years	⊕⊕⊕⊖ **MODERATE**[5] due to risk of bias		The mean uhu ≥6 months in the control groups was **1.5**	The mean uhu ≥6 months in the intervention groups was **0.11 lower** (0.59 lower to 0.37 higher)
UHU ≥6 months Hospitalisation	283 (3 studies) 6-12 months	⊕⊖⊖⊖ **VERY LOW**[3,5] due to risk of bias, imprecision	**RR 1.17** (0.31 to 4.43)	**Moderate** **22 per 1000**	**4 more per 1000** (from 15 fewer to 75 more)
UHU ≥6 months Number of hospital admissions	95 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean uhu ≥6 months in the control groups was **0.09**	The mean uhu ≥6 months in the intervention groups was **0.05 lower** (0.16 lower to 0.06 higher)
UHU ≥6 months days hospitalisation	88 (1 study) 12 months	⊕⊕⊕⊖ **MODERATE**[5] due to risk of bias		The mean uhu ≥6 months in the control groups was **0.1**	The mean uhu ≥6 months in the intervention groups was **0.03 lower** (0.21 lower to 0.15 higher)
UHU ≥6 months ED visits	192 (2 studies) 6-12 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision	**RR 3.78** (0.96 to 14.93)	**22 per 1000**	**60 more per 1000** (from 1 fewer to 303 more)
UHU ≥6 months Mean number of ED visits	183 (2 studies)	⊕⊕⊕⊖ **MODERATE**[5]		The mean uhu ≥6 months in the control groups was	The mean uhu ≥6 months in the intervention groups was **0.04 lower**

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects	
	12 months	due to risk of bias		**0.11**	(0.2 lower to 0.12 higher)
UHU ≥6 months Unscheduled doctors visit	183 (2 studies) 6 months	⊕⊖⊖⊖ **VERY LOW**[3,5,6] due to risk of bias, inconsistency, imprecision	**RR 0.77** (0.18 to 3.34)	**Moderate** **281 per 1000**	**65 fewer per 1000** (from 230 fewer to 658 more)
Rescue medication ≥6months requiring nebulised salbutamol	65 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,3] due to risk of bias, imprecision	**RR 1.98** (0.35 to 11.08)	**Moderate** **54 per 1000**	**53 more per 1000** (from 35 fewer to 544 more)
FEV1 L ≥6 months	88 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean fev1 l ≥6 months in the control groups was **2.71**	The mean fev1 l ≥6 months in the intervention groups was **0.26 lower** (0.61 lower to 0.09 higher)
FEV1 % ≥6 months Scale from: 0 to 100.	163 (2 studies) 6-12 months	⊕⊕⊖⊖ **LOW**[1] due to risk of bias		The mean fev1 % ≥6 months in the control groups was **84.1 %**	The mean fev1 % ≥6 months in the intervention groups was **0.10 higher** (0.92 lower to 1.12 higher)
PEF % best ≥6 months Scale from: 0 to 100.	63 (1 study) 6 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean pef % best ≥6 months in the control groups was **79.62 %**	The mean pef % best ≥6 months in the intervention groups was **5.31 higher** (1.91 lower to 12.53 higher)
Time off work ≥6 months	192 (2 studies) 6-12 months	⊕⊖⊖⊖ **VERY LOW**[1,3] due to risk of bias, imprecision	**RR 1.41** (0.62 to 3.21)	**Moderate** **83 per 1000**	**34 more per 1000** (from 32 fewer to 183 more)
Mean days off work ≥6 months	183 (2 studies) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean days off work ≥6 months in the control groups was **2.6 days**	The mean days off work ≥6 months in the intervention groups was **2.5 higher** (1.27 to 3.74 higher)

[1] The majority of the evidence was from studies at very high risk of bias
[2] 95% CI crosses one MID
[3] 95% CI crosses two MIDs
[4] Heterogeneity in the point estimates, I2=52%
[5] The majority of the evidence was from studies at high risk of bias
[6] Heterogeneity in the point estimates, I2=86%

Table 77: Clinical evidence summary: Children: Monitoring PEF versus symptom monitoring

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with PEF versus symptoms monitoring: children (95% CI)
Exacerbations <6months OCS	24 (1 study) 3 months	⊕⊖⊖⊖ **VERY LOW**[2,3] due to risk of bias,	**RR 1.00** (0.07 to 14.21)	**Moderate** **83 per 1000**	**0 fewer per 1000** (from 77 fewer to 1000 more)[1]

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects	
		imprecision			
Exacerbations ≥6months OCS	46 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[2] due to risk of bias	**OR 16.34** (3.25 to 82.24)	**Moderate**	
				0 per 1000	**370 more per 1000** (from 150 more to 590 more)[1]
UHU <6 months Hospitalisation	89 (1 study) 12 weeks	⊕⊖⊖⊖ **VERY LOW**[3,4] due to risk of bias, imprecision	**OR 7.56** (0.15 to 381.04)	**Moderate**	
				0 per 1000	**20 more per 1000** (from 40 fewer to 80 more)[1]
UHU <6 months Attendance at A&E	89 (1 study) 12 weeks	⊕⊖⊖⊖ **VERY LOW**[3,4] due to risk of bias, imprecision	**OR 7.56** (0.15 to 381.04)	**Moderate**	
				0 per 1000	**20 more per 1000** (from 40 fewer to 80 more)[1]
UHU(<6 months) Emergency GP visits	89 (1 study) 12 weeks	⊕⊖⊖⊖ **VERY LOW**[3,4] due to risk of bias, imprecision	**RR 0.93** (0.44 to 1.97)	**Moderate**	
				244 per 1000	**17 fewer per 1000** (from 137 fewer to 237 more)
Rescue meds ≥6 months requiring salbutamol	44 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[2,5] due to risk of bias, imprecision	**OR 14.15** (0.79 to 252.1)	**Moderate**	
				0 per 1000	**120 more per 1000** (from 50 fewer to 280 more)[1]
FEV1 % best (<6 months)	202 (2 studies) 12 weeks	⊕⊕⊖⊖ **LOW**[2] due to risk of bias		The mean fev1 % best (<6 months) in the control groups was **88.5 %**	The mean fev1 % best (<6 months) in the intervention groups was **0.39 higher** (0.21 lower to 0.98 higher)
PEF % best (<6 months)	89 (1 study) 12 weeks	⊕⊖⊖⊖ **VERY LOW**[2,3] due to risk of bias, imprecision		The mean pef % best (<6 months) in the control groups was **80.6 %**	The mean pef % best (<6 months) in the intervention groups was **2.8 higher** (2.15 to 3.45 higher)
Time off school (<6 months)	89 (1 study) 12 weeks	⊕⊖⊖⊖ **VERY LOW**[2,3] due to risk of bias, imprecision	**RR 1.18** (0.64 to 2.18)	**Moderate**	
				289 per 1000	**52 more per 1000** (from 104 fewer to 341 more)

[1] Manual risk difference calculation due to no events in one group
[2] The majority of the evidence was from studies at very high risk of bias
[3] 95% CI crosses 2 MIDs
[4] The majority of the evidence was from studies at high risk of bias
[5] 95% CI crosses one MID

24.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow chart in Appendix E.

24.5 Evidence statements

Clinical

ADULTS (>16 years): monitoring PEF vs conventional monitoring

- No evidence was identified for mortality, asthma control questionnaires or QOL (as a continuous outcome).
- Monitoring PEF vs conventional monitoring was considered a clinically important harm for asthma exacerbations (2 studies, N=152, very low quality), UHU ED visits,(2 studies, N=192, very low quality) and number of people requiring nebulised salbutamol (1 study, N=65, very low quality), all at ≥6 months and for lung function (FEV1) at ≥6 months when measured in litres (1 study, N=88, very low quality). However, this benefit was not seen when lung function was measured as %pred, there was no difference between the two groups (2 studies, N=163, very low quality).
- Monitoring PEF vs conventional monitoring was considered a clinically important benefit for UHU GP visits (2 studies, N=183, very low quality) and lung function when measured as PEF %best (1 study, N=63, very low quality), both at ≥6 months.
- Monitoring PEF vs conventional monitoring resulted in no clinically important difference for UHU hospitalisation (3 studies, N=283, very low quality) and time off work (2 studies, N=192, very low quality), both at ≥6 months.
- Evidence was also available for asthma exacerbations (mean number of OCS courses per person) and for UHU (mean number of total asthma related visits, mean number of hospitalisations, mean number of days in hospital and mean number of ED visits) as continuous outcomes. However, for all these outcomes, it is unclear whether the lower absolute values in the PEF monitoring group represent a clinical benefit when reported on a continuous scale.

CHILDREN (>16 years): monitoring PEF vs conventional monitoring

- No evidence was identified for mortality, asthma control questionnaires or QOL.
- Monitoring PEF vs conventional monitoring resulted in no clinically important difference for asthma exacerbations (1 study, N=24, very low quality) and lung function measured as both FEV1 %best (2 studies, N=202, low quality) and PEF %best (1 study, N=89, very low quality), all at <6 months.
- Monitoring PEF vs conventional monitoring was considered a clinically important harm for asthma exacerbations (1 study, N=46, low quality) and time off school (1 study, N=89, very low quality), both at <6 months, and for use of rescue medications (1 study, N=44, very low quality) at ≥6 months.
- Monitoring PEF vs conventional monitoring resulted in a borderline clinically important difference for UHU hospitalisations, ED visits and GP visits (1 study, N=89, very low quality), all at <6 months.

Economic

- No relevant economic evaluations were identified.

24.6 Recommendations and link to evidence

Recommendations	**41. Monitor asthma control at each review in adults, young people and children aged 5 and over using either spirometry or peak flow variability testing.**
Relative values of different outcomes	The GC considered the following outcomes as critical for this review: mortality, quality of life, exacerbations, asthma control measured by questionnaires and unscheduled healthcare utilisation. The GC noted that exacerbations of asthma can lead to both unscheduled healthcare utilisation (ED visits, out-of-hours GP surgery visits or hospitalisation) and the need for a course of OCS. Therefore, the GC considered exacerbations requiring OCS treatment and exacerbations requiring unscheduled healthcare utilisation separately. The GC also considered the following important outcomes: lung function (FEV1), symptoms (symptom scores and symptom-free days), dose of regular asthma therapy (ICS dose), rescue medication requirement (SABA dose) and time off school or work. These outcomes were considered as important measures of asthma control for the patient. The GC acknowledged that regular asthma preventer therapy (ICS) is often under prescribed / used, and rescue medications (SABA) may be overprescribed / used[144].
Trade-off between clinical benefits and harms	In adults, monitoring including PEF vs. monitoring without PEF was associated with a clinically important benefit for UHU (unscheduled GP visits) and a clinically important harm for ED visits and asthma exacerbations (OCS use). The GC acknowledged that on some occasions OCS may be of benefit by preventing a more severe asthma attack requiring UHU. There was no clinically important difference in hospitalisations. For the majority of the outcomes, evidence was only available from two studies, and no evidence was available for mortality, asthma control questionnaires or QOL. In children, monitoring including PEF vs monitoring without PEF was associated with no clinically important benefit in asthma exacerbations at less than 6 months and a clinically important harm for asthma exacerbations at more than 6 months. There was a trend towards more hospitalisations and ED visits in the PEF monitoring group and fewer GP visits, but the GC considered the clinical harms and benefits to be unclear. For the majority of the outcomes, evidence was only available from one study, and no evidence was available for mortality, asthma control questionnaires or QOL. The GC noted that all of the studies compared self-management using PEF monitoring with self-management using symptom monitoring. No studies were identified comparing monitoring of PEF vs monitoring of symptoms by a GP. Also, no studies were identified monitoring spirometry. According to the GC, some outcomes favoured PEF monitoring, whereas others favoured symptom monitoring. The GC agreed that all people with asthma should have a self-management action plan, but there is no evidence to show PEF-monitoring plans are better than symptom-monitoring plans. The GC also discussed the small increase in hospitalisations and the need for nebulised salbutamol. This may be an appropriate increase in outcomes in some people and may reflect the PEF self-management monitoring intervention recognising poor control at an earlier stage. No evidence was available to assess the utility of monitoring spirometry to measure asthma control. The GC debated the importance of monitoring spirometry and the additional information that it provided over and above PEF. Given the relative ease

	of monitoring spirometry and the additional information that it provides, the GC agreed that spirometry should be measured at every review. Spirometry provides additional information on the level of airways obstruction and can be compared to the previous best measurement or predicted measurement based on age and height of the individual. As no evidence was identified comparing PEF or spirometry monitoring by a GP at each asthma review, the GC made a consensus recommendation on the basis of current best practice that either spirometry (FEV1) or PEF should be used at every asthma review to monitor asthma control in children aged over 5 years. Children under 5 years are unable to perform these tests.
Economic considerations	No relevant economic evaluations were identified. The cost of providing the equipment, such as peak-flow meters, to monitor asthma is likely to be negligible. The main cost-consequence of monitoring using lung function tests is the impact it has on medication usage. If monitoring using lung function tests produces false results which increase medication usage then this will stand as an inefficient use of NHS resources as costs will increase with no added health benefits. On the other hand, if accurate, monitoring using lung function tests could reduce medication usage and provide cost savings. The clinical evidence showed a reduction in unscheduled healthcare utilisation. Although oral corticosteroid usage was higher in the lung function test group compared to the no lung function test group, the GC considered that this may be due to the poor control being identified early and thus preventing expensive hospitalised exacerbations. The GC considered it likely that using lung functions tests to monitor asthma control is cost-effective at a £20,000 per QALY threshold.
Quality of evidence	In adults, for the comparison of monitoring PEF vs conventional monitoring, evidence ranged from very low to moderate quality. For the majority of the outcomes, evidence was only available from one or two studies. In children, all the evidence was of very low and low quality. For the majority of the outcomes, evidence was only available from one study and the studies were of small sample size. The GC noted that the majority of the evidence was from older studies.
Other considerations	None. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

25 Monitoring: Fractional exhaled nitric oxide (FeNO)

25.1 Introduction

Asthma can be divided into extrinsic asthma (atopic or allergic), intrinsic asthma (non-atopic) and occupational asthma. Atopy is defined as a genetic predisposition to produce immunoglobulin E (IgE) against common environmental aeroallergens such as house dust mites, animal dander, pollens and moulds. Approximately 80% of people with asthma are atopic compared with 30% of the general population. Atopic asthma is characterised by Th2 lymphocyte driven inflammation within the airways.

Exhaled nitric oxide (NO) mainly originates from the respiratory epithelium and is produced by inducible NO synthase (iNOS). In patients with asthma, iNOS expression is upregulated by interleukin-4 and -13, both archetypal Th2 cytokines. Thus exhaled NO primarily signals Th2 lymphocyte driven inflammation in the bronchial mucosa and consequently has potential utility in the monitoring of airways inflammation in asthma.

25.2 Review question: In people with asthma, what is the clinical and cost-effectiveness of using fractional exhaled nitric oxide (FeNO) measures for monitoring asthma control?

For full details see review protocol in Appendix C.

Table 78: PICO characteristics of review question

Population	People with asthma (defined as physician Dx with objective test). If insufficient evidence is found we will consider evidence from studies where asthma is defined as physician Dx only or on asthma treatment (give details of how Dx was made). Not including severe asthma. All ages, stratified into the following 2 different groups: • Children/young people (5-16 years old) • Adults (>16 years old)
Intervention(s)	Monitoring FeNO and adjustment of management/therapy according to physician decision or personalised treatment plan (use of other interventions to be included if equal access in each group, eg both groups receive education in addition to monitoring) Only use validated methods of measuring FeNO (eg 50ml/s flow rate).
Comparison(s)	Comparison of adjustment of asthma therapy based on FeNO to: • Usual care: eg clinical symptoms (with or without PEF) according to guidelines (including BTS/SIGN, GINA) • Asthma control questionnaires or QOL questionnaires • Lung function tests (spirometry or PEFv) • Blood eosinophils • Challenge tests Comparison of different frequencies of monitoring using FeNO.
Outcomes	Critical outcomes: • Mortality

	• Unscheduled healthcare utilisation (ED/A&E visit; hospital admissions; GP out of hours or walk-in centre) • Exacerbations (defined as need for course of oral steroids) • Asthma control questionnaires (ACT; CACT; ACQ; PACQ; RCP-3) • QoL (AQLQ; pAQLQ; St George's respiratory questionnaire) Important outcomes: • Lung function (FEV_1, PEF) • Symptoms (annual symptom free days) • Dose of regular asthma therapy / preventer medication (ICS dose) • Rescue medication (SABA use) • Time off school or work

25.3 Clinical evidence

We searched for randomised trials comparing FeNO monitoring versus conventional monitoring, in patients with asthma. A Cochrane systematic review was identified[128]. Studies included in this Cochrane review were included individually and data extracted separately in order to incorporate additional outcomes from the protocol and additional, more recently published studies.

Fifteen studies (14 RCTs and the Cochrane review) were included in the review[26,42,56,71,125,125,127,128,130,131,138,156,164,172,173,187], all compared FeNO monitoring versus conventional monitoring. These studies are summarised in Table 79 and Table 80 below. See also the study selection flow chart in Appendix D, study evidence tables in Appendix G, forest plots in Appendix J, GRADE tables in Appendix I and excluded studies list in Appendix K.

Evidence in adults was available from six studies[26,71,138,156,164,172], one of which was in pregnant women[138], and evidence in children from eight studies[41,56,125,130,131,173,187]. For the conventional monitoring group, some studies monitored symptoms alone and other studies used algorithms for treatment adjustment based on symptoms, lung function and bronchodilator (BD) use. For the FeNO monitoring group, studies monitored FeNO alone; FeNO and symptoms; or FeNO, symptoms, lung function and BD use. Two studies included patients with severe asthma: Peirsman[125] included 6% (N= 3) patients with severe asthma in the control group and 4% (N=2) severe asthma in the intervention group; Pike[131] recruited patients with moderate to severe asthma.

Table 79: Summary of studies included in the review: ADULTS

Study	Intervention/comparison (frequency of adjustment)	Population	Outcomes	Follow-up and frequency of adjustment
Calhoun 2012[26]	Monitoring FeNO / monitoring symptoms and lung function (every 6 weeks) FeNO cut-off: <22ppb step-down; 22-35ppb maintain; >35 step up	Adults Physician diagnosis with BDR or AHR	QOL; exacerbations; asthma control questionnaires; SABA use; ICS use; lung function; time off work	6, 12 and 36 weeks
Honkoop 2014[71]	Monitoring FeNO and symptoms (ACQ score) / monitoring symptom control questionnaire and treatment (every 3 months) FeNO cut-off: <25ppb; 22-35ppb; >50 step up	Adults Physician diagnosis according to the Dutch national guidelines	Exacerbation; ACQ; UHU (ED visit & hospitalisation; lung function	12 months
Powell 2011[138]	Monitoring FeNO / monitoring symptoms according to clinical guideline algorithm (monthly) FeNO cut off: <16ppb; 16-29ppb; >29 ppb	Adults (non-smoking pregnant women, >18 years) Physician diagnosis and using inhaled therapy for asthma within the last year. The diagnosis was confirmed by a respiratory physician's diagnostic interview.	QOL; exacerbation (mixed, moderate to severe, defined as events for which the patient sought medical attention (an unscheduled visit to a doctor, presentation to the emergency room or admission to hospital, or use of ICS)); ACQ; SABA use; ICS use; lung function; symptom free days	4-6 months
Shaw 2007[156]	Monitoring FeNO and symptoms (ACQ score) / monitoring symptoms (ACQ score) (monthly to 4 months then every 2 months) FeNO cut-off: <16ppb (or <26ppb twice) step-down; >26 ppb step up	Adults GP diagnosis of asthma.	Exacerbations; ICS use	12 months
Smith 2005[164]	Monitoring FeNO (additional safety buffer if asthma deteriorated in absence of FeNO rise) / monitoring symptoms, BD use and lung function	12 – 75 years	Exacerbations; SABA use; ICS use; lung function; symptom free days	12 months

Study	Intervention/comparison (frequency of adjustment)	Population	Outcomes	Follow-up and frequency of adjustment
	(every 2 months) FeNO cut-off: 35ppb (safety buffer to step-up if deteriorating asthma in the absence of a rise in FeNO)	Chronic asthma		
Syk 2013[172]	Monitoring FeNO and symptoms / monitoring symptoms, beta agonist use, lung function (at 1, 2, 4 and 8 months) Men - FeNO step-down: <21 ppb; no change 21-25 ppb; step-up; ≥26 ppb / Women – FeNo step-down: <19; no change 19-23; step-up ≥24	18-64 years Physician diagnosis of asthma and atopic	Exacerbation; ACQ; SABA use; ICS use; lung function	12 months

Table 80: Summary of studies included in the review: CHILDREN

Study	Intervention/comparison (frequency of adjustment)	Population	Outcomes	Follow-up and frequency of adjustment
De jongste 2009[42]	Monitoring FeNO and symptoms / monitoring symptoms (measured daily and adjusted based on 3 week mean) FeNO step-down: symptom score low, FeNO low; maintain: symptom score high, FeNO low; step-up performed in every other case. Cut-off: 20 ppb for children aged 6-10 years and 25 ppb for older children	Children 6-18 years Diagnosed according to GINA	QOL; exacerbation; UHU; SABA use; ICS use; lung function; symptom free days	30 weeks
Fritsch 2006[56]	Monitoring symptoms, beta agonist use, lung function, and FeNO / monitoring symptoms, beta agonist use and lung function (every 6 weeks)	Children 6-18 years Diagnosed according to ATS	Exacerbation; ICS use	6 months

Study	Intervention/comparison (frequency of adjustment)	Population	Outcomes	Follow-up and frequency of adjustment
	Step-down: FEV1% predicted: ≥80%, no or mild symptoms, and beta agonist use <6 puffs over last 14 days; step-up performed in every other case. Further adjustment based on FeNO cut-off >20 ppb			
Peirsman 2013[125]	Monitoring FeNO and symptoms / monitoring symptoms, beta agonist use, lung function (every 3 months) FeNO cut-off: 20 ppb	Children Mild to severe allergic asthma according to GINA guidelines	Exacerbation; UHU; ICS use; lung function; symptom free days; time off school	12 months
Petsky 2014[127]	Monitoring FeNO and atopic status / monitoring symptoms (monthly to 4 months then every 2 months) FeNO cut-off: ≥10 ppb in children with no positive skin prick test (SPT), ≥12 ppb in children with one positive SPT, and ≥20 ppb in children with ≥2 positive SPT	Children aged >4 years with persistent asthma	Exacerbation; QOL; UHU; ICS use; lung function	12 months
Pijnenburg 2005[130]	Monitoring FeNO and symptoms / monitoring symptoms (every 3 months) FeNO step-down: ≤30 ppb and symptom score ≤14; maintain: ≤30 ppb and symptom >14; step-up: >30 ppb, regardless of symptoms.	Children Atopic asthma, and fulfilled ATS criteria for asthma.	Exacerbation; ICS use; lung function	12 months
Pike 2012[131]	Monitoring FeNO and symptoms / monitoring symptoms, BD use and lung function (every 2 months) FeNO cut-off: step-up: ≥25 ppb or FeNO more than twice baseline; maintain: >15 ppb to ≤25 ppb; step-down: ≤15 ppb.	6-17 years Clinical diagnosis of asthma with BDR or PEFv	UHU; ICS use	12 months
Szefler 2008[173]	Monitoring FeNO, symptoms, BD use and lung	12 to 20 years	Exacerbation; UHU; ACQ; ICS use;	46 weeks

Study	Intervention/comparison (frequency of adjustment)	Population	Outcomes	Follow-up and frequency of adjustment
	function / monitoring symptoms, BD use and lung function (every 6-8 weeks) Control level determined based on all the above. FeNO control level one: 0-20ppb; two: 20.1-30ppb; three: 30.1-40ppb; four: >40ppb.	Asthma	lung function; symptoms free days; time off school	
Verini 2010[187]	Monitoring FeNO, symptoms, BD use and lung function / monitoring symptoms, BD use and lung function (every 6 months) FeNO cut-off: ≥12ppb	6-17 years Diagnosis made according to ATS-ERS criteria	ICS use; SABA use	12 months

Table 81: Clinical evidence summary: [FeNO versus Conventional Monitoring] Adults

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with FeNO versus conventional monitoring ADULTS (95% CI)
UHU (ED visit) ≥6 months	415 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision	**OR 0.68** (0.12 to 3.98)	**Moderate** **14 per 1000**	**4 fewer per 1000** (from 12 fewer to 39 more)
UHU (hospitalisation) ≥6 months	415 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision	**OR 0.52** (0.05 to 5.07)	**Moderate** **10 per 1000**	**5 fewer per 1000** (from 9 fewer to 39 more)
Exacerbation (OCS) ≥6 months	393 (3 studies) 52 weeks	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision	**RR 0.84** (0.56 to 1.26)	**Moderate** **313 per 1000**	**50 fewer per 1000** (from 138 fewer to 81 more)
Exacerbation (OCS) ≥6 months	342 (1 study) 9 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision	**HR 0.91** (0.39 to 2.11)	**Moderate**	-[3]
Exacerbation (OCS) ≥6 months	415 (1 study)	⊕⊖⊖⊖ **VERY LOW**[1,2]	**OR 0.64** (0.27 to	**Moderate**	-[3]

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects	
	12 months	due to risk of bias, imprecision	1.56)		
Exacerbation (mixed UHU or OCS use) <6 months	220 (1 study) 4-6 months	⊕⊕⊖⊖ **LOW**[1,6] due to risk of bias, indirectness	**RR 0.61** (0.41 to 0.90)	**413 per 1000**	**161 fewer per 1000** (from 41 fewer to 244 fewer)
AQLQ (≥ 6months) Asthma Quality of Life Questionnaire. Scale from: 1 to 7.	227 (1 study) 6 weeks	⊕⊕⊖⊖ **LOW**[1] due to risk of bias		The mean aqlq (≥ 6months) in the control groups was **0.02 change score**	The mean aqlq (≥ 6months) in the intervention groups was **0 higher** (0.22 lower to 22 higher)[5]
ACQ ≥6 months Asthma Control Questionnaire. Scale from: 0 to 6.	644 (2 studies) 9-12 months	⊕⊕⊕⊖ **MODERATE**[1] due to risk of bias		The mean acq ≥6 months in the control groups was **0.03 change score**	The mean acq ≥6 months in the intervention groups was **0.05 lower** (0.13 lower to 0.04 higher)
ACQ (<6 months) Asthma Control Questionnaire. Scale from: 0 to 6.	220 (1 study) 4-6 months	⊕⊕⊕⊕ **HIGH**		The acq (mean) 4-6 months in the control groups was **0.72**	The acq (mean) 4-6 months in the intervention groups was **0.16 lower** (0.36 lower to 0.04 higher)
ACQ (at exacerbation <6 month) Asthma Control Questionnaire. Scale from: 0 to 6.	220 (1 study) 4-6 months	⊕⊕⊕⊕ **HIGH**		The acq (mean at exacerbation) 4-6 months in the control groups was **1.97**	The acq (mean at exacerbation) 4-6 months in the intervention groups was **0.05 higher** (0.18 lower to 0.28 higher)
ACQ (at unscheduled doctor visits <6 month) Asthma Control Questionnaire. Scale from: 0 to 6.	220 (1 study) 4-6 months	⊕⊕⊕⊕ **HIGH**		The acq (mean at unscheduled doctor visit) 4-6 months in the control groups was **2.01**	The acq (mean at unscheduled doctor visit) 4-6 months in the intervention groups was **0.02 lower** (0.21 lower to 0.25 higher)
ACQ (clinically important improvement, ≥0.5) ≥6 months Asthma Control Questionnaire	155 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision	**RR 1.39** (0.86 to 2.26)	**Moderate** **257 per 1000**	**100 more per 1000** (from 36 fewer to 324 more)
FEV1 %pred Scale from: 0 to 100.	736 (3 studies) 9-12 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean fev1 %pred in the control groups was **82.3 %**	The mean fev1 %pred in the intervention groups was **0.45 higher** (0.69 lower to 1.59 higher)
FEV1, litres ≥6 months	166 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[1] due to risk of bias		The mean fev1, litres ≥6 months in the control groups was **-0.006 litres change score**	The mean fev1, litres ≥6 months in the intervention groups was **0.03 lower** (0.11 lower to 0.06 higher)
PEF am (L/min) ≥6 months	321 (2 studies) 9-12 months	⊕⊕⊖⊖ **LOW**[1] due to risk of bias		The mean pef am (l/min) ≥6 months in the control groups was **403 L/min**	The mean pef am (l/min) ≥6 months in the intervention groups was **2 higher** (10.39 lower to 14.39 higher)

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects	
PEF pm (L/min) ≥6 months	227 (1 study) 9 months	⊕⊕⊖⊖ **LOW**[1] due to risk of bias		The mean pef pm (l/min) ≥6 months in the control groups was **-13.3 L/min change score**	The mean pef pm (l/min) ≥6 months in the intervention groups was **3.8 higher** (10 lower to 17.6 higher)
ICS use ≥6 months fluticasone or BDP equivalent	212 (2 studies) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,2,6] due to risk of bias, indirectness, imprecision		The mean ics use ≥6 months in the control groups was **768 mcg**	The mean ics use ≥6 months in the intervention groups was **0.53 standard deviations lower** (0.8 to 0.25 lower)
Rescue medication (puffs/day) ≥6 months	321 (2 studies) 9-12 months	⊕⊖⊖⊖ **VERY LOW**[1,6] due to risk of bias, indirectness		The mean rescue medication (puffs/day) ≥6 months in the control groups was **0.4**	The mean rescue medication (puffs/day) ≥6 months in the intervention groups was **0.06 lower** (0.12 lower to 0 higher)
% symptom free days ≥6 months Scale from: 0 to 100.	94 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean % symptom free days ≥6 months in the control groups was **63.7 %**	The mean % symptom free days ≥6 months in the intervention groups was **5.6 higher** (8.51 lower to 19.71 higher)
Time of work (number of people) ≥6 months	229 (1 study) 9 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision	**OR 2** (1.17 to 3.41)	**Moderate**	-[3]

[1] Downgraded by one increment if the majority of the evidence was at high risk of bias, and downgraded by two increments if the majority of the evidence was at very high risk of bias
[2] Downgraded by one increment if the confidence interval crossed one MID or by two increments if the confidence interval crossed both MIDs
[3] Control group event rate not reported
[4] Control group event rate not reported
[5] 97.5% CI reported and extracted
[6] Downgraded by one/two increments because: the majority of the evidence had indirect outcomes

Table 82: Clinical evidence summary: [FeNO versus Conventional Monitoring] Children

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with FeNO versus conventional monitoring CHILD (95% CI)
UHU (unscheduled visits) ≥6 months	581 (2 studies) 46-52 weeks	⊕⊖⊖⊖ **VERY LOW**[1,2] due to inconsistency, imprecision	**RR 0.67** (0.29 to 1.55)	**Moderate** **299 per 1000**	**99 fewer per 1000** (from 212 fewer to 164 more)
UHU (hospitalisation) ≥6 months	725 (4 studies) 46-52 weeks	⊕⊖⊖⊖ **VERY LOW**[2,3] due to risk of	**RR 0.97** (0.48 to 1.95)	**Moderate** **34 per 1000**	**1 fewer per 1000** (from 18 fewer to 32 more)

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects	
		bias, imprecision			
UHU (number of children ≥1 emergency room admin) ≥6 months	91 (1 study) 52 weeks	⊕⊖⊖⊖ **VERY LOW**[2,3] due to risk of bias, imprecision	**RR 0.51** (0.1 to 2.65)	**Moderate** **87 per 1000**	**43 fewer per 1000** (from 78 fewer to 144 more)
Exacerbation (OCS) ≥6 months	927 (6 studies) 43 weeks	⊕⊕⊕⊖ **MODERATE**[2] due to imprecision	**RR 0.74** (0.61 to 0.9)	**Moderate** **192 per 1000**	**50 fewer per 1000** (from 19 fewer to 75 fewer)
Asthma control (ACT score) ≥6 months ACT. Scale from: 5 to 25.	494 (1 study) 46 weeks	⊕⊕⊕⊕ **HIGH**		The mean asthma control (act score) ≥6 months in the control groups was **21.83**	The mean asthma control (act score) ≥6 months in the intervention groups was **0.06 higher** (0.27 lower to 0.39 higher)
PACQLQ (Pediatric Asthma Caregiver) ≥6 months Pediatric Asthma Care Quality of Life Questionnaire. Scale from: 1 to 7.	147 (1 study) 30 weeks	⊕⊕⊖⊖ **LOW**[3] due to risk of bias		The mean pacqlq (pediatric asthma caregiver) ≥6 months in the control groups was **6.2**	The mean pacqlq (pediatric asthma caregiver) ≥6 months in the intervention groups was **0 higher** (0.24 lower to 0.24 higher)
FEV1 % pred ≥6 months Scale from: 0 to 100.	579 (2 studies) 46-52 weeks	⊕⊕⊕⊖ **MODERATE**[2] due to imprecision		The mean fev1 % pred in the control groups was **95.5 %**	The mean fev1 % pred in the intervention groups was **0.94 higher** (0.31 lower to 2.19 higher)
ICS dose ≥6 months fluticasone	494 (1 study) 46 weeks	⊕⊕⊕⊖ **MODERATE**[4] due to indirectness		The mean ics dose in the control groups was **570 mcg (estimated from graph)**	The mean ics dose in the intervention groups was **118.9 higher** (48.5 to 189.3 higher)
% symptom free days ≥6 months Scale from: 0 to 100.	147 (1 study) 30 weeks	⊕⊖⊖⊖ **VERY LOW**[2,3] due to risk of bias, imprecision			The mean % symptom free days in the intervention groups was **0.3 higher** (10 lower to 10.6 higher)
Number of symptom days in last 2 weeks; ≥6 months	494 (2 studies) 46 weeks	⊕⊕⊕⊕ **HIGH**		The mean number of symptom days in last 2 weeks; ≥6 months in the control groups was **1.89**	The mean number of symptom days in last 2 weeks; ≥6 months in the intervention groups was **0.04 higher** (0.21 lower to 0.29 higher)
Number of patients not using inhaled corticosteroids or anti-leukotrienes ≥6 months	64 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[2,3] due to risk of bias, imprecision	**RR 0.33** (0.07 to 1.53)	**Moderate** **188 per 1000**	**126 fewer per 1000** (from 175 fewer to 100 more)
Rescue medication (no. of patients needed beta-agonist due to symptoms) ≥6 months	64 (1 study)	⊕⊖⊖⊖ **VERY LOW**[2,3]	**RR 0.62** (0.42 to	**Moderate** **813 per 1000**	**309 fewer per 1000**

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects	
	12 months	due to risk of bias, imprecision	0.9)		(from 81 fewer to 472 fewer)
Number of school days missed in last 2 weeks; ≥6 months	494 (1 study) 46 weeks	⊕⊕⊕⊕ **HIGH**		The mean number of school days missed in last 2 weeks; ≥6 months in the control groups was **0.23 days**	The mean number of school days missed in last 2 weeks; ≥6 months in the intervention groups was **0.04 lower** (0.12 lower to 0.04 higher)
Time off (school - number of children missed school) ≥6 months	92 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[2,3] due to risk of bias, imprecision	**RR 0.83** (0.4 to 1.73)	**Moderate** **261 per 1000**	**44 fewer per 1000** (from 157 fewer to 191 more)

[1] Downgraded by one/two increments because: heterogeneity, I2=50%, p=0.04
[2] Downgraded by one increment if the confidence interval crossed one MID or by two increments if the confidence interval crossed both MIDs
[3] Downgraded by one increment if the majority of the evidence was at high risk of bias, and downgraded by two increments if the majority of the evidence was at very high risk of bias
[4] Downgraded by one/two increments because: the majority of the evidence had indirect outcomes

Table 83: Clinical evidence summary: [FeNO versus Conventional Monitoring] Adults – outcomes reported with median values

ICS: ADULTS	
Dose of regular therapy	
Calhoun (9 mo) ≥6mo = ICS, expressed as equivalent dose to BDP	As means (without variance – cannot calculate effect estimates)
Powell (4-6 mo) <6mo =BDP equivalent ICS dose	Median (IQR) - FeNO group: 2000 (0 to 400). Control group: 0 (0 to 800.)
Shaw (12 mo) ≥6mo = ICS, expressed as equivalent dose to BDP	As means
Smith (12 mo) ≥ 6mo = fluticasone equivalent dose	As means
Syk (12 mo) ≥ 6mo = Budesonide equivalent dose	FeNO group: 0 (-400 - 400), N= 86. Control group: 0 (-200 - 200), N= 78. [p=.945]

Table 84: Clinical evidence summary: [FeNO versus Conventional Monitoring] Children – outcomes reported with median values

ICS: CHILDREN	
Dose of regular therapy	
De Jongste (30 wk) ≥6mo = budesonide	Medians. FeNO group: 200 mcg/day (0-500). Control group: 200 mcg/day (100-500). P=<0.0001.
Fritsch (6 mo) ≥6mo = Fluticasone and budesonide permitted (unclear)	Medians. FeNO group: 316 mcg (200-500 mcg). Control group: 241 mcg (26-607 mcg).
Perisman (12 mo) ≥6mo	Medians. FeNO group: +100 mcg (0, +400). Control group: 0 mcg (-200, +80). P=0.016

ICS: CHILDREN	
= Budesonide or equivalent	
Petsky (12 mo) ≥6mo = fluticasone	Medians. FeNO group: 400, (250-600). Control group: 200, (100-400).
Pijnenburg (3 mo) <6mo = budesonide	As means (cannot calculate effect estimate). Mean daily ICS dose increased between visits 1 (0 months) and 2 (3 months) by 169 mcg (95% CI, 80-259; p<0.001) in the FeNO group and 172 mcg (95% CI, 92-251; p<0.001) in the symptom group.
Pike (12 mo) ≥6mo = Beclometasone, fluticasone and budesonide permitted (unclear)	Medians. FeNO group: 800 (400-1000), N=34. Control group: 500 (400-1000), N=43. P=0.0543.
Szefler (46 we) ≥6mo = fluticasone	As means
Verini (12 mo) ≥6mo = ICS unclear	Only reported as number not using ICS or anti-LTs

[Median range - Budesonide, FeNO: 100-800 mcg. Control: 0-500 mcg; Fluticasone, FeNO: 316-800 mcg, Control: 200-500 mcg]

GENERAL: other outcomes reported as medians: Adults	
Rescue Meds	
Powell (4-6 months) <6mo = beta agonist use in past week	Median (IQR) - Feno group: 0 (0 to 3). Control group: 1 (0 to 5). P-value 0.024.
Syk (12 mo) ≥6mo = SABA use per week, at 8-12 months	Medians. FeNO group: 1.56 (0.06-5.18). Control group: 0.94 (0.03-2.81)

GENERAL: other outcomes reported as medians: Adults	
% Symptom free	
Powell (4-6 months)<6mo	Median (IQR) – FeNO group: 7 (4 to 7). Control: 6 (2 to 7).
Quality of Life – AQLQ-M score	
Powell (4-6 months)<6mo	FeNO group: 0.81 (0.38 to 1.63). Control group: 0.75 (0.38 to 1.25).
Lung Function	
Powell (4-6 months)<6mo =FEV1 (L)	Mean (95%CI) – FeNO group: 3.09 (3.0 to 3.17). Control group: 3.01 (2.91 to 3.10).
Powell (4-6 months)<6mo = FEV1 (%)	Mean (95% CI) – FeNO group: 96.4% (94.31 to 98.46). Control group: 94.4% (91.84 to 96.96).

GENERAL: other outcomes reported as medians: Children	
% Symptom free	
Perisman (12 mo) ≥6mo	Medians. FeNO group: 83.7 (27.1-91.9). Control group: 79.6 (51.7-94.0).
Rescue Meds	
De Jongste (30 wk) ≥6mo = beta agonist puffs per 3 weeks	The median number of rescue beta-agonist puffs per 3 weeks was similar at baseline [2 (0-19) in the FeNO group and 2 (0-21) in the control group] and decreased to 0 (0-19) and 1 (0-19), respectively.

25.4 Economic evidence

Published literature

One economic evaluation was identified with the relevant comparison and has been included in this review.[66] This is summarised in the economic evidence profile below (Table 85) and the economic evidence table in Appendix H.

Three economic evaluations relating to this review question were identified but were excluded due to methodological limitations and the availability of more applicable evidence.[18,71,140] These are listed in Appendix L, with reasons for exclusion given.

See also the economic article selection flow chart in Appendix E.

Table 85: Economic evidence profile: FeNO monitoring versus standard care

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost-effectiveness	Uncertainty
Harnan et al. 2013 (UK) [66]	Directly applicable[(a)]	Potentially serious limitations[(b)]	The economic evaluation was conducted using a Markov model with a lifetime horizon. The evaluation compared monitoring using 4 FeNO measurements per year plus annual clinical review to annual clinical review only (standard care) each year of the patient's life. Separate analyses were run for children and for adults.	Children: £2288[(c)] Adults: £81	Children: 0.0560 QALYs Adults: 0.0379 QALYs	Children: £45,213 per QALY gained Adults: £2,146 per QALY gained	Probabilistic sensitivity analysis (PSA) showed that FeNO monitoring on adults had an 82% chance of being cost-effective at a £20,000 per QALY threshold. A variety of one-way sensitivity analyses were conducted on the adult analysis. The result was most sensitive to how long the impact of FeNO monitoring lasted. If the impacts on exacerbation rates and change in inhaled-corticosteroids (ICS) dosage lasted less than 30 years (as opposed to lifetime in the base case) then FeNO monitoring was no longer cost-effective with an ICER of £29,707 per QALY gained. PSA on the result for children showed that FeNO monitoring had a 1% chance of being cost-effective at a £20,000 per QALY threshold. The same one-way sensitivity analyses were conducted for children. The result was most sensitive to how long the impact of FeNO monitoring on exacerbation rates and ICS dosage lasted. If it lasted less than 5 years then FeNO monitoring was considered cost-effective with an ICER of £7598 per QALY gained. This is because ICS dosage was increased following monitoring and that increased costs.

(a) Cost–utility analysis conducted using an NHS perspective

(b) Both analyses for children and adults are based on single RCT trials. Only quality of life improvements from reduced exacerbations are considered and impacts on mortality are not considered, however these limitations are due to a lack of clinical evidence rather than methodological choices. Strong assumptions imposed regarding extrapolating treatment effects over a lifetime horizon. Model results are very sensitive to changes in core parameters.

(c) The significant difference in cost between adults and children is due to FeNO monitoring reducing medication costs for adults but increasing medication costs for children.

Unit costs

Although UK economic evidence was available, the unit cost of FeNO monitoring was presented to the GC for considerations. This is reported in Table 87 below.

Table 86: Annual cost of FeNO monitoring

Item	Unit cost	Quantity per year	Annual cost	Source
20 minute practice nurse visit	£14.66	4	£58.64	PSSRU 2013[40]
Marginal cost of using FeNO equipment[(a)]	£4.82–£7.07	4	£19.28–£28.28	Harnan et al. 2013[66]
		Annual total	£77.92–£86.92	

(a) The cost varies depending on whether NIOX VERO, NIOX MINO or No breath equipment is used

Economic considerations

One study by Honkoop et al[71,72] was excluded from the economic review as the uncertainty surrounding the health benefits derived from the monitoring strategies was too uncertain to produce a reliable ICER. The reason being QALYs were likely rounded, but also quality of life was valued using a Dutch EQ-5D tariff. However this study was a within-trial analysis that provided useful data on the potential cost impact of FeNO monitoring, therefore these costs were presented to the GC for consideration and are shown below in Table 87. Note the societal costs from the paper have been excluded as the NICE reference case only uses costs that are incurred by the NHS. These costs had a strong influence over the cost-effectiveness of FeNO in the study. Non-asthma related costs have also been excluded as there was no reason to believe why these costs would be influenced by FeNO monitoring and they are likely to drastically fluctuate year on year. This was shown by the large confidence intervals surrounding these costs.

Table 87: Incremental costs of FeNO monitoring compared to monitoring using Asthma Control Questionnaire

Item	Incremental cost
Asthma medication costs	-£78
Marginal cost of using FeNO equipment	£73
Asthma related healthcare visits	-£40
Annual total	-£45

Table 87 above shows that FeNO monitoring resulted in costs that were £45 lower from FeNO monitoring. However as the study was not conducted in a UK setting it is difficult to say whether the asthma related healthcare visit costs would remain the same in a UK setting, no information was given on resource use. These costs are only gathered from one year of follow-up. Medication useage fluctuated significantly in the study and there was no indication that medication useage would remain the same after one year of follow-up.

25.5 Evidence statements

Clinical

Children

- No evidence was identified for mortality.
- Monitoring FeNO vs conventional monitoring was considered a clinical benefit for asthma exacerbations (OCS use; 6 studies, N=927, moderate quality), UHU all unscheduled visits (2 studies, N=581, very low quality) and use of rescue medication (1 study, N=64, very low quality), all at ≥6 months.
- Monitoring FeNO vs conventional monitoring resulted in a borderline clinically important difference for UHU emergency room admissions (1 study, N=91, very low quality) and time off school (1 study, N=92, very low quality), both at ≥6 months.
- Monitoring FeNO vs conventional monitoring resulted in no clinically important difference for UHU hospitalisations (4 studies, N=725, very low quality), asthma control questionnaire score (1 study, N=494, high quality), QOL (1 study, N=147, low quality), lung function measured as FEV1 %pred (2 studies, N=579, moderate quality), symptom free days reported as both the % of days and number of days (very low and high quality, respectively) and days off school when reported as the mean number of days (1 study, N=494, high quality), all at ≥6 months.
- Although evidence from one study showed fewer patients in the FeNO group required inhaled corticosteroids of anti-leukotreines, in general, there was no clinically important difference in the mean or median dose of ICS between groups.

Adults

- No evidence was identified for mortality.
- Monitoring FeNO vs conventional monitoring resulted in a clinically important benefit for asthma exacerbations at ≥6 months (3 studies, N=393, very low quality) and at <6 month in pregnant women (1 study, N=220, low quality). More people in the FeNO monitoring group had a clinically important improvement in ACQ score (1 study, N=155, very low quality), however, there was no clinically important difference in the mean ACQ score.
- Monitoring FeNO vs conventional monitoring resulted in no clinically important difference for UHU ED visits or hospitalisations (1 study, N=415, very low quality), QOL (1 study, N=227, low quality), asthma control questionnaire score (2 studies, N=644, moderate quality), lung function when measured as FEV1 litres, FEV1 %pred or PEF L/min (low to very low quality), use of rescue medications (2 studies, N=321, very low quality) and symptom free days (1 study, N=94, very low quality), all at ≥6 months.

Economic

- One cost–utility analysis found that for monitoring asthma
  - FeNO was cost-effective in adults compared to standard care (ICER: £2146 per QALY gained).
  - FeNO was not cost-effective in children compared to standard care (ICER: £45,213 per QALY gained).

  This analysis was assessed as directly applicable with potentially serious limitations.

25.6 Recommendations and link to evidence

Recommendations	**42.Do not routinely use FeNO to monitor asthma control.**

	43. Consider FeNO measurement as an option to support asthma management in people who are symptomatic despite using inhaled corticosteroids. (This recommendation is from NICE's diagnostics guidance on measuring fractional exhaled nitric oxide concentration in asthma.)
Research recommendations	**1. Which patient groups are likely to benefit from FeNO monitoring to guide asthma management, for example, individuals with atopy, frequent asthma attacks, poor adherence?** **2. What is the clinical and cost effectiveness of FeNO-guided monitoring of asthma in real-world settings?**
Relative values of different outcomes	The GC considered the following outcomes as critical for this review: mortality, quality of life, exacerbations, asthma control measured by questionnaires and unscheduled healthcare utilisation. The GC noted that exacerbations of asthma can lead to both unscheduled healthcare utilisation (ED visits, out-of-hours GP surgery visits or hospitalisation) and the need for a course of OCS. Therefore, the GC considered exacerbations requiring OCS treatment and exacerbations requiring unscheduled healthcare utilisation separately. The GC also considered the following important outcomes: lung function (FEV_1), symptoms (symptom-free days), dose of regular asthma therapy (ICS dose), rescue medication (SABA dose) and time off school or work. These outcomes were considered as important measures of asthma control for the patient. The GC acknowledged that regular asthma preventer therapy (ICS) is often under-prescribed/used, and rescue medications (SABA) may be over-prescribed/used[144].
Trade-off between clinical benefits and harms	In children, monitoring FeNO vs conventional monitoring was considered to confer a clinically important benefit in reduction in exacerbations (defined as the need for course of oral corticosteroids). Evidence from six studies suggested fewer children required a course of OCS in the FeNO monitoring group. There was also a clinical benefit in UHU when reported as all unscheduled visits, and rescue medication use in the FeNO monitoring group. However, there was no clinically important difference in UHU hospitalisations, asthma control questionnaire score, quality of life, dose of regular therapy (inhaled corticosteroid dose), lung function and symptom-free days. Whilst there were reports of fewer UHU emergency room visits and less time off school the differences in these parameters between the FeNO group and the control group were small and insuffient for the GC to justify a recommendation in favour of FeNO. For the majority of the outcomes, evidence was only available from one study, and no evidence was available for mortality. It is the view of the GC that whilst evidence exists (some of which is of high quality, with one study indicating a positive result), overall it is insufficient to support the use of FeNO in children. The evidence for paediatrics is dominated by two studies (de Jongste 2009 and Szefler 2008). Analysis suggests that there is evidence of a trend towards clinical benefit in exacerbations (i.e. fewer children requiring a course of OCS in the FeNO monitoring group); however, there is essentially no difference in clinically meaningful outcomes between the FeNO and the control-group patients. The GC noted a different method of administering OCS in the US study (Szefler 2008) from that normally found in the UK. This study involved parent-initiated use of OCS rather than doctor-initiated, and may represent a different population to the UK and influence the magnitude of effect seen in this outcome. Furthermore, it was noted that the cumulative steroid burden (OCS and ICS) is still higher in people who had

	FeNO monitoring, and those in the FeNO group received higher doses of ICS. The GC also acknowledged that, on some occasions, OCS may be of benefit by preventing a more severe asthma attack requiring UHU. The GC concluded that there is insufficient evidence to support routine FeNO for monitoring in children, and targeted research in this area is required to identify whether FeNO monitoring is beneficial in certain subgroups, for example, in those with severe asthma. Severe asthma is, however, outside the scope of this review. Monitoring FeNO may also benefit those patients with symptoms despite taking ICS, for example, with relative steroid-resistant Th2 high disease, but there is no evidence to support this currently; further research is required. Research in these areas would require very large participant numbers in order to subgroup patients. In adults, monitoring FeNO vs conventional monitoring was considered to confer a clinically important benefit for asthma exacerbations due to fewer people requiring a course of OCS in the FeNO monitoring group. The GC acknowledged that on some occasions OCS may be of benefit by preventing a more severe asthma attack requiring UHU. There was no clinically important difference in quality of life, mean asthma control questionnaire score, UHU (ED visits and hospitalisations), lung function measured by FEV1 or PEF, symptom free days and dose of regular therapy (inhaled corticosteroids). For the majority of the outcomes, evidence was only available from one study, and no evidence was available for mortality. As with children, the GC concluded that there is little evidence to support FeNO for routine monitoring in adults. There was some evidence to suggest a small clinical benefit of FeNO monitoring on the outcome of exacerbations. The GC discussed the definition for the exacerbation outcome – as stated in the protocol, as usual with a dichotomous outcome, the review considered the number of patients requiring a course of oral corticosteroids. However, it was noted that the number of oral corticosteroids courses prescribed may provide different and valuable information, and would perhaps be a more responsive measure.
Economic considerations	One health economic study was presented to the GC. This study reported different results for children and adults and separately. The study concluded that for adults FeNO monitoring could be cost-effective with an ICER of £2,146 per QALY gained. However this result is contingent on various strong assumptions that the GC debated. The main assumption underpinning the result is that the benefits derived from FeNO monitoring in terms of exacerbation and ICS dosage will be life-long. The GC agreed that this assumption is unlikely to hold in reality as any potential benefits would be seen early on and as time passes the benefits would decrease as there would be less need for medicine titration once experience had shown the optimum dose in each individual person. The result was also contingent on the idea that inhaled-corticosteroid usage would be lower for the remainder of the individual's life. As the health benefit from FeNO monitoring is small this benefit has a significant impact on the cost-effectiveness of FeNO monitoring. The GC agreed that lower ICS usage is unlikely to hold for the remainder of the individual's life especially since this result was extrapolated from one RCT with a two-year follow up time. In the sensitivity analysis the economic study showed that slight changes to these assumptions, such as making FeNO benefit only last for 30 years or less, was enough to drive the ICER above £20,000 per QALY. Due to considerable uncertainty in the result and contention over the assumptions used, the GC agreed that this evidence was not certain enough to recommend FeNO monitoring for adults given its proven low health benefit in the general asthma community, as shown in our clinical review. The study concluded for children that FeNO monitoring was not cost-effective with

	an ICER of £45,213 per QALY gained. This result was based on the same strong assumptions as in the adult model. As this study was based on one non-UK RCT and the assumptions were unlikely to hold in reality, the GC was cautious about the validity of this evidence. However the clinical evidence in our review showed little to no benefit of FeNO monitoring in children and changing the assumptions on the duration of effectiveness would increase the ICER even more, so the GC had no reason to believe the conclusions of the study were wrong and concluded that FeNO monitoring is not cost-effective in children with asthma. The GC also considered the cost implications derived from the study by Honkoop. The GC noted that the only reliable information was the cost derived from the difference in medication usage between the monitoring strategies. This cost was calculated in the study as being £78. This is £5 greater than the cost of the FeNO monitoring used in the study. This suggests that FeNO is the slightly cheaper alternative. The study also showed FeNO produced further cost savings from reduced healthcare visits which resulted in a further £40 cost-saving. However the study used a Dutch healthcare perspective and gave no information on resource use, making it difficult to extrapolate this value to a UK setting. The GC's main concern was that these data were extracted over a 12 month time frame; as asthma is a variable disease, a longer time-frame would be needed to show if over time these costs remained consistently different. FeNO monitoring costs £77 - £87 per patient per year. It can also have a large impact on resource utilisation by increasing or reducing ICS usage. Given there was no strong clinical evidence that showed significant health benefits the GC noted that FeNO monitoring was unlikely to be cost-effective as a routine management strategy for all people with asthma. However the GC noted that in a specific severe sub-group of patients the health benefits could be much higher. Therefore in these people FeNO monitoring could be a cost-effective management strategy and therefore identifying this subgroup through was research was the GCs top priority.
Quality of evidence	In children, for the comparison of monitoring FeNO tests vs conventional monitoring, evidence for four of the critical outcomes ranged from very low to high quality (specifically: unscheduled healthcare utilisation, very low; quality of life, low; exacerbation, moderate; asthma control questionnaire, high). Evidence for all other outcomes was of very low quality (symptom-free days, rescue medication, and time off school), moderate quality (dose of regular therapy (ICS), lung function), and high quality (number of symptom-free days in last 2 weeks, number of school days missed). For the majority of the outcomes, evidence was only available from one study with a follow-up of 30-52 weeks. In adults, for the comparison of monitoring FeNO vs conventional monitoring, evidence for all outcomes was of low and very low quality. For the majority of the outcomes, evidence was only available from one study with a follow-up of 36-52 weeks. The economic evidence was assessed as directly applicable with potentially serious limitations.
Other considerations	It is the view of the GC that there is very little additional benefit for monitoring FeNO compared with conventional monitoring. The GC was aware that the NICE DAP recommended FeNO monitoring "*as an option to support asthma management (in conjunction with the British guideline on the management of asthma) in people who are symptomatic despite using inhaled corticosteroids*". It is important to note that the current review has included three additional studies in children and one additional study in adults. It should also be noted that the NICE DAP considered the definition of exacerbation as that stated in the study. In this review, however, the definition of exacerbation was defined and agreed by the GC a priori as the number

	of people requiring a course of OCS. Even taking the study definition of exacerbations into account, there was only a trend towards benefit from FeNO monitoring (in adults) in the NICE DAP. No meta-analysis was performed in children in the NICE DAP. Although evidence from published studies does not suggest a significant benefit for monitoring FeNO compared with conventional monitoring, the GC noted its potential importance in identifying adherence. FeNO levels are very much predicated on whether or not patients are adhering to treatment, and FeNO may be potentially useful in monitoring whether or not they are taking inhaled steroids as prescribed. The GC noted the heterogeneity between studies with regards to the algorithms used for treatment adjustment in the intervention and control arms. The change in outcomes would be dependent on the algorithm and cut-off values used, which complicates the assessment of the effectiveness of FeNO as a monitoring strategy. In the conventional monitoring group, some studies monitored symptoms alone and other studies used algorithms for treatment adjustment based on symptoms, lung function and bronchodilator (BD) use. For the FeNO monitoring group, studies monitored FeNO alone; FeNO and symptoms; or FeNO, symptoms, lung function and BD use. Similarly, different FeNO cut-off levels were employed across the studies (<20ppb, <25ppb, <35ppb, etc.). Furthermore, there was heterogeneity in the definition of the 'severe asthma exacerbations' outcome between individual studies. For this review, severe asthma exacerbations were considered as requiring a course of OCS. Other study definitions (for example, asthma exacerbations resulting in hospitalisation or increased use of SABA) were reported in the review under the relevant protocol outcomes. The GC noted that in the studies included, the frequency of visits (around 4 times per year) was the same in the intervention and control groups. The GC also discussed whether FeNO monitoring may be of benefit in patients in specialist centres. The GC therefore made future research recommendations to investigate which subgroups of patients are likely to benefit from FeNO monitoring to guide asthma management, e.g. individuals with atopy, frequent asthma attacks and/or those with poor adherence, and the clinical and health economic benefit of FeNO-guided monitoring of asthma in real-world settings (please see appendix N for the full list of research recommendations made). Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. One study was identified which the GC agreed did not suggest a change in the recommendations was warranted.

26 Monitoring: Peripheral blood eosinophil count

26.1 Introduction

Eosinophils are a form of white blood cells produced by the bone marrow. Their exact role in health has yet to be determined, but it is believed that they play a role in fighting parasitic infections and primarily reside within the lining of the gut.

Biopsies taken from the lungs of people with asthma have frequently demonstrated increased numbers of eosinophils and the number of eosinophils is also often increased in sputum samples taken from asthma sufferers. Measurement of sputum eosinophil numbers have been used to aid the diagnosis and management of asthma. However, this is a time consuming procedure, which is only performed in a specialist setting. Eosinophils travel in the blood stream from the bone marrow to the lung, it is therefore logical to investigate whether measurement of blood eosinophils is a useful tool for monitoring asthma control.

26.2 Review question: In people with asthma, what is the clinical and cost-effectiveness of using the peripheral blood eosinophil count for monitoring asthma control?

For full details see review protocol in Appendix C.

Table 88: PICO characteristics of review question

Population	People with asthma (defined as physician Dx with objective test or recurrent persistent wheeze in <5 years). If insufficient evidence is found we will consider evidence from studies where asthma is defined as physician Dx only or on asthma treatment (give details of how Dx was made). Not including severe asthma. All ages, stratified into the following 3 different groups: • Children (1-<5 years old) • Children/young people (5-16 years old) • Adults (>16 years old)
Intervention(s)	Monitoring peripheral blood eosinophil count and adjustment of management/therapy according to physician decision or personalised treatment plan (use of other interventions to be included if equal access in each group, eg both groups receive education in addition to monitoring)
Comparison(s)	Comparison of adjustment of asthma therapy based on peripheral blood eosinophil count to: • Usual care: eg clinical symptoms (with or without PEF) according to guidelines (including BTS/SIGN, GINA) • Asthma control questionnaires or QOL questionnaires • Lung function tests (spirometry or PEFv) • Challenge tests Comparison of different frequencies of monitoring using blood eosinophil count.
Outcomes	Critical outcomes: • Mortality • Unscheduled healthcare utilisation (ED/A&E visit; hospital admissions; GP out of hours or walk-in centre) • Exacerbations (defined as need for course of oral steroids)

	• Asthma control questionnaires (ACT; CACT; ACQ; PACQ; RCP-3) • QoL (AQLQ; pAQLQ; St George's respiratory questionnaire) Important outcomes: • Lung function (FEV1, PEF) • Symptoms (annual symptom free days) • Dose of regular asthma therapy / preventer medication (ICS dose) • Rescue medication (SABA use) • Time off school or work
Study design	RCTs

26.3 Clinical evidence

No relevant clinical evidence was identified from RCTs (from full papers or conference abstracts) looking at the effectiveness of using the peripheral blood eosinophil count for monitoring asthma control.

26.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow chart in Appendix E.

26.5 Evidence statements

Clinical

- No relevant clinical studies were identified.

Economic

- No relevant economic evaluations were identified.

26.6 Recommendations and link to evidence

Recommendations	**No clinical recommendation.**
Research recommendations	**3. What is the clinical and cost effectiveness of using blood eosinophils as a tool to monitor asthma control?**
Relative values of different outcomes	The GC considered the following outcomes as critical for this review: mortality, quality of life, exacerbations, asthma control measured by questionnaires and unscheduled healthcare utilisation. The GC noted that exacerbations of asthma can lead to both unscheduled healthcare utilisation (ED visits, out-of-hours GP surgery visits or hospitalisation) and the need for a course of OCS. Therefore, the GC considered exacerbations requiring OCS treatment and exacerbations requiring unscheduled healthcare utilisation separately. The GC also considered the following important outcomes: lung function (FEV1),

	symptoms (symptom-free days), dose of regular asthma therapy (ICS dose), rescue medication (SABA dose) and time off school or work. These outcomes were considered as important measures of asthma control for the patient. The GC acknowledged that regular asthma preventer therapy (ICS) is often underprescribed / used, and rescue medications (SABA) may be overprescribed / used (NRAD[144]).
Trade-off between clinical benefits and harms	No RCT evidence was identified.
Economic considerations	As no clinical evidence was found, there was no way to determine the cost-effectiveness of using blood eosinophils for monitoring asthma.
Quality of evidence	No RCT evidence was identified.
Other considerations	The GC was aware of and discussed the existence of prognostic studies within the broader literature base showing the association between peripheral blood eosinophils and future outcomes. However, these studies do not show that interventions based on the peripheral blood eosinophil level improve patient outcomes. There is some evidence for monitoring sputum eosinophil levels in asthma[60], but the practical difficulties of doing this have prevented translation into routine practice. Blood eosinophil meaurement is relatively quick and easy to perform, correlates loosely with sputum eosinophilia, and there are prognostic studies within the literature suggesting that the blood eosinophil count may be predictive of future outcomes. The GC was therefore keen to make a future research recommendation to explore whether improved patient outcomes might result when the blood eosinophil count is used to monitor asthma control and guide treatment (please see appendix N for the full list of research recommendations made). Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

27 Monitoring: Challenge tests

27.1 Introduction

Asthma is characterised by excessive constriction of the smooth muscle in the airways in response to a variety of stimuli, including inhaled allergens, viral infections, cold air, smoke and other irritants. The degree of bronchoconstriction induced by an appropriate exposure is considerably greater in people with asthma than in those without, although bronchoconstriction can be induced in healthy people if high enough stimuli are provided. The 'twitchiness' of the airways can be assessed by 'bronchial challenge tests', in which the individual is exposed to progressively higher levels of constriction-inducing stimuli and the level of bronchoconstriction (usually assessed as FEV1) is assessed. Exposure usually occurs through incrementally greater exposure to an inhaled constrictor, usually as a 'direct' challenge (e.g. with a nebulized bronchconstrictor molecule, such as methacholine or histamine) or as an 'indirect' challenge (e.g. hypertonic saline, mannitol inhalation, eucapnic hyperventilation or exercise), measuring the level of exposure required to produce a 10 or 20% fall in FEV1. The result is usually expressed as the concentration or cumulative dose of an exposure resulting in a specified fall in FEV1.

The test needs to be done in a controlled pulmonary function laboratory setting by a qualified technician with appropriate equipment and protocols and resuscitation facilities, as there is a small risk of severe bronchoconstriction. The test will usually take approximately 30 minutes, and is usually mildly unpleasant to the patient, in that it involves induced bronchoconstriction (which is relieved by inhaled bronchodilator at the end of the test). The test is unsuitable for younger children. Varieties of this test are available in most hospital lung departments, but are rarely available to GPs currently. However, the clinical and cost-effectiveness of using indirect challenge tests to monitor asthma control is currently uncertain.

27.2 Review question: In people with asthma, what is the clinical and cost-effectiveness of using indirect challenge tests with mannitol or direct challenge tests with histamine or methacholine for monitoring asthma control?

For full details see review protocol in Appendix A.

Table 89: PICO characteristics of review question

Population	People with asthma (defined as physician Dx with objective test). If insufficient evidence is found we will consider evidence from studies where asthma is defined as physician Dx only or on asthma treatment (give details of how Dx was made). Not including severe asthma. All ages, stratified into the following 2 different groups: • Children/young people (5-16 years old) • Adults (>16 years old)
Intervention(s)	Monitoring using indirect or direct challenge tests and using the outcomes to adjust management/therapy according to physician decision or personalised treatment plan (use of other interventions to be included if equal access in each group, eg both groups receive education in addition to monitoring): • Indirect challenge test with mannitol • Direct challenge test with methacholine or histamine
Comparison(s)	Comparison of adjustment of asthma therapy based on indirect or direct challenge tests

	to: • Usual care: eg clinical symptoms according to guidelines (including BTS/SIGN, GINA) • Asthma control questionnaires or QOL questionnaires • Lung function tests (spirometry or PEFv) Comparison of adjustment of asthma therapy based on: • Indirect vs direct challenge tests • Comparison of different frequencies of monitoring using challenge tests
Outcomes	Critical outcomes: • Mortality • Unscheduled healthcare utilisation (ED/A&E visit; hospital admissions; GP out of hours or walk-in centre) • Exacerbations (defined as need for course of oral steroids) • Asthma control questionnaires (ACT; CACT; ACQ; PACQ; RCP-3) • QoL (AQLQ; pAQLQ; St George's respiratory questionnaire) Important outcomes: • Lung function (FEV1, PEF) • Symptoms (annual symptom free days) • Dose of regular asthma therapy / preventer medication (ICS dose) • Rescue medication (SABA use) • Time off school or work
Study design	• RCTs • Exclude observational cohort studies and NRS unless limited evidence from RCTs

27.3 Clinical evidence

We searched for randomised trials comparing the effectiveness of monitoring using indirect or direct challenge tests vs. monitoring according to usual care to guide asthma treatment and management. Four studies were included in the review[86,97,118,166] these are summarised in Table 90 below. See also the study selection flow chart in Appendix D, clinical evidence tables in Appendix G, forest plots in Appendix J, GRADE tables in Appendix I and excluded studies list in Appendix K.

Three involved adults[86,97,166] and one involved children[118]; all compared challenge tests versus no challenge tests for monitoring. For the monitoring intervention, three studies used methacholine challenge test[86,118,166], including the study in children, and one study used mannitol challenge test[97]. No evidence was found for monitoring mannitol challenge tests in children. For the comparator group, some studies monitored symptoms alone and other studies used algorithms for treatment adjustment based on more than one clinical parameter (for example, symptoms, BD use and lung function). Evidence from these studies is summarised in the clinical evidence summary below in Table 91, Table 92 and Table 93.

Table 90: Summary of studies included in the review

Study	Intervention/comparison	Population	Outcomes	Comments
Koenig 2008[86]	BHR (methacholine PC20)/ Starting dose of treatment and adjustment of dose (at each 8 week visit for 40 weeks) based on severity class (without BHR as a clinical measure).	12 years of age and older Either historical documentation of reversible airways disease within the last 24 months or an increase in FEV1 of at least 12% within 30 min of inhalation of 2 puffs (180 mcg) of salbutamol.	Mortality, exacerbations (not defined), ICS dose, lung function (FEV1 and PEF), % symptom-free days	
Lipworth 2012[97]	Treatment adjusted based on mannitol AHR only, every 2 months for 12 months/ Treatment adjusted according to BTS guidelines	18 to 65 years old History of mild to moderate persistent asthma	AQLQ, exacerbations (OCS), lung function (FEV1 and PEF)	Step down of treatment before study
Nuijsink 2007[118]	Treatment adjusted on the basis of methacholine AHR and symptom score/ Treatment adjusted on the basis of symptom score only	Children aged 6–16 years Documented clinical history of moderate persistent asthma, according to GINA guidelines.	Exacerbations (OCS), daily ICS dose, lung function (FEV1), % symptom-free days	
Sont 1999[166]	Treatment adjusted at each 3 month visit based on severity class or methacholine AHR/ Treatment adjusted at each 3 month visit based on severity class only	18 to 50 years old History of episodic chest tightness and wheezing in the previous year and visiting a chest physician for their asthma.	Lung function (FEV1)	Did not report severe exacerbations due to infrequent events, only mild exacerbations reported

Table 91: Clinical evidence summary: ADULTS Methacholine challenge test versus no challenge test for asthma monitoring

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with ADULTS Methacholine challenge test versus no challenge test (95% CI)
Mortality (≥6 months)	212 (1 study) 40 weeks	⊕⊖⊖⊖ **VERY LOW**[2,3] due to risk of bias, imprecision	**OR 7.53** (0.15 to 379.61)	**Moderate** **0 per 1000**	**10 more per 1000** (from 20 fewer to 40 more)[1]
Asthma exacerbations (≥6 months)	212 (1 study) 40 weeks	⊕⊖⊖⊖ **VERY LOW**[2,3,4] due to risk of bias, indirectness, imprecision	**RR 0.86** (0.52 to 1.42)	**Moderate** **243 per 1000**	**34 fewer per 1000** (from 117 fewer to 102 more)
Rescue medications (≥6 months) Salbutamol puffs/day	212 (1 study) 40 weeks	⊕⊕⊕⊖ **MODERATE**[5] due to risk of bias		The mean rescue medications (≥6 months) in the control groups was **-0.7 puffs/day (change score)**	The mean rescue medications (≥6 months) in the intervention groups was **0.1 lower** (0.58 lower to 0.38 higher)
ICS use >6months mean daily dose (mcg; fluticasone propionate)	212 (1 study) 40 weeks	⊕⊕⊕⊖ **MODERATE**[5] due to risk of bias		The mean ics use >6months in the control groups was **254 mcg**	The mean ics use >6months in the intervention groups was **131.2 higher** (83.57 to 178.83 higher)
FEV1 (≥6 months) L	279 (2 studies) 40-104 weeks	⊕⊕⊖⊖ **LOW**[5,6] due to risk of bias, inconsistency		The mean fev1 (≥6 months) in the control groups was **0.05 L change score**	The mean fev1 (≥6 months) in the intervention groups was **0.04 lower** (0.09 lower to 0.16 higher)
% symptom free days (≥6 months) Scale from: 0 to 100.	212 (1 study) 40 weeks	⊕⊕⊕⊖ **MODERATE**[5] due to risk of bias		The mean % symptom free days (≥6 months) in the control groups was **18.1 % change score**	The mean % symptom free days (≥6 months) in the intervention groups was **5.1 lower** (20.06 lower to 9.86 higher)
PEF am (≥6 months) L/min	212 (1 study) 40 weeks	⊕⊕⊕⊖ **MODERATE**[5] due to risk of bias		The mean pef am (≥6 months) in the control groups was **407 L/min**	The mean pef am (≥6 months) in the intervention groups was **8.6 lower** (17.20 lower to 0 higher)
PEF pm (≥6 months) L/min	212 (1 study) 40 weeks	⊕⊕⊖⊖ **LOW**[5,7] due to risk of bias, imprecision		The mean pef pm (≥6 months) in the control groups was **22.4 L/min change score**	The mean pef pm (≥6 months) in the intervention groups was **6 lower** (29.96 lower to 17.96 higher)

[1] Manual calculation of absolute effect as zero events in the control group
[2] The majority of the evidence was from studies at very high risk of bias due to allocation concealment and missing data
[3] 95% CI crosses 2 MIDs
[4] Evidence from one study - exacerbations not defined
[5] The majority of the evidence was from studies at high risk of bias due to allocation concealment
[6] Point estimates show statistical heterogeneity I2=72% P<0.06. Only 2 studies so random effects model used.
[7] 95% CI crosses one MID

Table 92: Clinical evidence summary: ADULTS Mannitol challenge test versus no challenge test for asthma monitoring

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with ADULTS Mannitol challenge test versus no challenge test (95% CI)
AQLQ (≥6 months) mini AQLQ. Scale from: 1 to 7.	119 (1 study) 52 weeks	⊕⊕⊖⊖ **LOW**[1,2] due to risk of bias, indirectness		The mean aqlq (≥6 months) in the control groups was **5.9**	The mean aqlq (≥6 months) in the intervention groups was **0.06 higher** (0.3 lower to 0.42 higher)
Asthma exacerbations (≥6 months)	119 (1 study) 52 weeks	⊕⊖⊖⊖ **VERY LOW**[2,3,4] due to risk of bias, indirectness, imprecision	**RR 0.88** (0.44 to 1.76)	**Moderate** **224 per 1000**	**27 fewer per 1000** (from 125 fewer to 170 more)
Rescue medications (≥6 months) Salbutamol puffs/day	119 (1 study) 52 weeks	⊕⊖⊖⊖ **VERY LOW**[1,2,5] due to risk of bias, indirectness, imprecision		The mean rescue medications (≥6 months) in the control groups was **0.67 puffs/day**	The mean rescue medications (≥6 months) in the intervention groups was **0.31 lower** (0.73 lower to 0.11 higher)
ICS use >6months mean daily dose (mcg; ciclesonide)	119 (1 study) 52 weeks	⊕⊕⊖⊖ **LOW**[1,2] due to risk of bias, indirectness		The mean ics use >6months in the control groups was **208 mcg**	The mean ics use >6months in the intervention groups was **306 higher** (241.71 to 370.29 higher)
FEV1% (≥6 months)	119 (1 study) 52 weeks	⊕⊕⊖⊖ **LOW**[1,2,6] due to risk of bias, indirectness		The mean fev1% (≥6 months) in the control groups was **88 %**	The mean fev1% (≥6 months) in the intervention groups was **0.3 higher** (8.21 lower to 8.81 higher)
PEF% (≥6 months) Scale from: 0 to 100.	119 (1 study) 52 weeks	⊕⊕⊖⊖ **LOW**[1,2] due to risk of bias, indirectness		The mean pef% (≥6 months) in the control groups was **94.3 %**	The mean pef% (≥6 months) in the intervention groups was **2.7 lower** (13.17 lower to 7.77 higher)
PEF am (≥6 months) L/min	119 (1 study) 52 weeks	⊕⊖⊖⊖ **VERY LOW**[1,2,4] due to risk of bias, indirectness, imprecision		The mean pef am (≥6 months) in the control groups was **411.1 L/min**	The mean pef am (≥6 months) in the intervention groups was **1.5 higher** (34.7 lower to 37.7 higher)

[1] The majority of the evidence was from studies at high risk of bias due to blinding
[2] Patients initially underwent step-down of their existing treatment. Patients on combination inhalers were switched to an equivalent dose of the same ICS only (unclear whether LABA was continued).
[3] The majority of the evidence was from studies at high risk of bias due to missing data
[4] 95% CI crosses 2 MIDs
[5] 95% CI crosses one MID
[6] The majority of the evidence was from studies at high risk of bias due to baseline differences

Table 93: Clinical evidence summary: CHILDREN Challenge test versus no challenge test for asthma monitoring

Outcomes	No of Participants	Quality of the evidence (GRADE)	Relative effect	Anticipated absolute effects	
				Risk with Control	Risk difference with CHILDREN Challenge test

	(studies) Follow up		(95% CI)		versus no challenge test (95% CI)
Asthma exacerbations (≥6 months) OCS course	206 (1 study) 2 years	⊕⊖⊖⊖ **VERY LOW**[1,2,3] due to risk of bias, indirectness, imprecision	**RR 0.96** (0.51 to 1.79)	**Moderate** **164 per 1000**	**7 fewer per 1000** (from 80 fewer to 130 more)
ICS dose Mean daily dose for treatment period	175 (1 study) 2 years	⊕⊖⊖⊖ **VERY LOW**[2,4,5] due to risk of bias, indirectness, imprecision		The mean ics dose in the control groups was **478 mcg**	The mean ics dose in the intervention groups was **84 higher** (10.66 to 157.34 higher)
FEV1% (≥6 months) Scale from: 0 to 100.	185 (1 study) 2 years	⊕⊕⊖⊖ **LOW**[2,4] due to risk of bias, indirectness		The mean fev1% (≥6 months) in the control groups was **93 %**	The mean fev1% (≥6 months) in the intervention groups was **6 higher** (1.2 lower to 10.8 higher)
% symptom free days (≥6 months) in last 3 months of treatment. Scale from: 0 to 100.	175 (1 study) 2 years	⊕⊖⊖⊖ **VERY LOW**[2,3,4] due to risk of bias, indirectness, imprecision		The mean % symptom free days (≥6 months) in the control groups was **71 %**	The mean % symptom free days (≥6 months) in the intervention groups was **1.1 lower** (10.1 lower to 7.9 higher)

[1] No explanation was provided
[2] Patients initially underwent step-down of their existing treatment.
[3] 95% CI crosses both MIDs
[4] The majority of the evidence was at high risk of bias due to allocation concealment and baseline differences
[5] 95% CI crosses one MID

27.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow chart in Appendix E.

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs are provided below to aid consideration of cost-effectiveness.

Table 94: Unit costs of monitoring with challenge tests

Item	Quantity[(a)]	Unit cost	Cost (per year)	Source of unit cost
Secondary care respiratory medicine outpatient visit	3	£150	£450	NHS reference costs[46]
Bronchial reactivity studies - Lab costs	3	£177	£531	NHS reference costs[46]
Total			£981	

(a) Based on GC opinion.

Economic considerations

Using the incremental cost of monitoring using challenge tests, the QALY increase which would be required for challenge tests to be considered cost-effective at a £20,000 per QALY threshold can be calculated as:

$$Change\ in\ QALYs = \frac{Change\ in\ cost}{£20{,}000}$$

Therefore, if it costs £981 to monitor using challenge tests each year, then this strategy would need to generate 0.04905 extra QALYs for each year monitoring occurs. This could be achieved by improving quality of life by 0.04905 per year or producing less quality of life per year but improving life expectancy.

$$\frac{£981}{£20{,}000} = 0.04905$$

To help put this figure into context we can consider the disutility and costs associated with an exacerbation.

Table 95: Disutility a patient experiences with an exacerbation

Severity of exacerbation	Quality of life decrease during exacerbation	Duration of exacerbation (years)	Disutility (QALYs)	Cost of exacerbation
Severe	0.56	0.08	0.0448	£873.75
Non-severe	0.32	0.01	0.0032	£38.33

Source: Harnan et al[66], NHS reference costs[46]

Based on these figures, monitoring using challenge tests would have to greatly decrease the number of exacerbations per year for it to be cost-effective.

27.5 Evidence statements

Clinical

ADULTS (>16 years): monitoring methacholine challenge tests vs conventional monitoring

- No evidence was identified for UHU, QOL or asthma control questionnaires.
- Monitoring methacholine challenge tests vs conventional monitoring resulted in a borderline clinically important difference for mortality at ≥6 months (1 study, N=212, very low quality)
- Monitoring methacholine challenge tests vs conventional monitoring was considered a clinically important benefit for asthma exacerbations at ≥6 months (1 study, N=212, very low quality)
- Monitoring methacholine challenge tests vs conventional monitoring resulted in no clinically important difference for use of rescue medications (1 study, N=212, moderate quality), lung function measured as FEV1 litres or PEF (low to moderate quality) and symptom free days (1 study, N=212, moderate quality), all at ≥6 months.Evidence of moderate quality was available from 1 study demonstrating a higher mean ICS dose in the methacholine challenge test monitoring group.

ADULTS (>16 years): monitoring mannitol challenge tests vs conventional monitoring

- No evidence was identified for mortality, UHU or asthma control questionnaires.
- Monitoring mannitol challenge tests vs conventional monitoring was considered a clinically important benefit for asthma exacerbations (1 study, N=119, very low quality) and use of rescue medications (1 study, N=119, very low quality), both at ≥6 months.
- Monitoring mannitol challenge tests vs conventional monitoring resulted in no clinically important difference forQOL at (1 study, N=119, low quality) and lung function measured using FEV1 or PEF (low to very low quality), all at ≥6 months.
- Evidence of low quality was available from 1 study demonstrating a higher mean ICS dose in the mannitol challenge test monitoring group.

CHILDREN (5-16 years): monitoring methacholine challenge tests vs conventional monitoring

- No evidence was identified for mortality, UHU, QOL or asthma control questionnaires.
- Monitoring methacholine challenge tests vs conventional monitoring resulted in no clinically important difference for asthma exacerbations (1 study, N=206, very low quality), lung function measured using FEV1 (1 study, N=185, low quality) and symptom free days (1 study, N=185, very low quality), all at ≥6 months.
- Evidence of very low quality was available from 1 study demonstrating a higher mean ICS dose in the methacholine challenge test monitoring group.

Economic

- No relevant economic evaluations were identified.

27.6 Recommendations and link to evidence

Recommendations	44.Do not use challenge testing to monitor asthma control.
Relative values of different outcomes	The GC considered the following outcomes as critical for this review: mortality, quality of life, exacerbations, asthma control measured by questionnaires and unscheduled healthcare utilisation.

	The GC noted that exacerbations of asthma can lead to both unscheduled healthcare utilisation (ED visits, out-of-hours GP surgery visits or hospitalisation) and the need for a course of OCS. Therefore, the GC considered exacerbations requiring OCS treatment and exacerbations requiring unscheduled healthcare utilisation separately. The GC also considered the following important outcomes: lung function (FEV1), symptoms (symptom-free days), dose of regular asthma therapy (ICS dose), rescue medication (SABA dose) and time off school or work. These outcomes were considered as important measures of asthma control for the patient. The GC acknowledged that regular asthma preventer therapy (ICS) is often underprescribed / used, and rescue medications (SABA) may be overprescribed / used.[144]
Trade-off between clinical benefits and harms	Evidence was available on mortality in adults from a single study of methacholine challenge testing, reporting one death in the monitoring group (reported as unrelated to asthma). Monitoring methacholine challenge tests demonstrated a clinically important benefit for asthma exacerbations. For the majority of the outcomes, evidence was only available from one study, and no evidence was available for UHU, QOL or asthma control questionnaire scores. There was no clinically important difference in SABA use, FEV1, the % symptom-free days or PEF. The mean dose of fluticasone propionate was 131.2mcg higher in the challenge test monitoring group at 40 weeks. In children, monitoring methacholine challenge tests vs conventional monitoring resulted in no clinically important difference for asthma exacerbations at ≥6 months. Evidence was only available from one study, and no evidence was available for mortality, UHU, QOL or asthma control questionnaire scores. There was no clinically important difference in FEV1 or % symptom-free days. The mean ICS dose was 84mcg higher in the challenge test monitoring group at 2 years. In adults, monitoring mannitol challenge tests vs conventional monitoring was considered to confer a clinically important benefit for asthma exacerbations and SABA use at 1 year, but there was no clinically important difference in the QOL, FEV1 or PEF between monitoring groups. Evidence was only available from one study, and no evidence was available for mortality, UHU or asthma control questionnaire scores. The mean ICS dose (using ciclesonide) was 306mcg higher in the challenge test monitoring group at 1 year. No evidence was found for monitoring mannitol challenge tests in children. The GC considered there to be a small benefit in asthma exacerbations in adults when monitoring methacholine or mannitol challenge tests; however, this was at the expense of a higher steroid load. The GC acknowledged that on some occasions OCS may be of benefit by preventing a more severe asthma attack requiring UHU. The evidence was limited and the majority of the evidence was of low quality. The GC agreed that the evidence was not strong enough to recommend challenge testing for monitoring of asthma.
Economic considerations	The cost of challenge test monitoring was estimated to be £981 per year. This includes the cost of three respiratory outpatient visits and the cost of conducting the challenge test each review. At this cost, the intervention would need to provide an additional 0.04905 QALYs per year of monitoring to be considered cost-effective. The GC was presented with severe and non-severe exacerbation disutilities to aid in their consideration of cost-effectiveness. It was noted that challenge tests would need to reduce severe exacerbations by at least one every year to be considered cost-effective at a £20,000 per QALY threshold. This level of benefit was not shown in the

	clinical review. The clinical evidence did not suggest that these benefits were achievable and therefore the GC agreed that challenge tests would not be cost-effective in routine care.
Quality of evidence	In adults, for the comparison of monitoring methacholine challenge tests vs. conventional monitoring, evidence for both of the critical outcomes was of very low quality by GRADE criteria. In particular, there was little evidence on mortality and the single study which reported this did not have the power to show a significant difference. Evidence for all other outcomes was of low or moderate quality. For the majority of the outcomes, evidence was only available from one study with a long follow-up of 40-104 weeks. In children, for the comparison of monitoring methacholine challenge tests vs conventional monitoring, evidence for all outcomes was of low and very low quality. Evidence was only available from one study with a long follow-up of 2 years. In adults, for the comparison of monitoring mannitol challenge tests vs conventional monitoring, evidence for all outcomes was of low and very low quality. Evidence was only available from one study with a long follow-up of 1 year.
Other considerations	The GC noted the heterogeneity between studies with regards to the algorithms used for treatment adjustment based on the challenge tests. The change in ICS dose and outcomes would be dependent on the algorithm and cut-off values used, so it is difficult to assess the effectiveness of challenge tests as a treatment strategy. This would also be dependent on the baseline severity of asthma and current ICS treatment level. The GC acknowledged that in certain high-risk people challenge testing may possibly have a benefit but concluded that this should not be recommended for general use. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

28 Monitoring adherence to treatment

28.1 Introduction

The regular (daily) use of inhaled corticosteroids (ICS) is advised for all patients other than those with mild, infrequent symptoms and low risk of exacerbation, with additional regular maintenance therapy added for those failing to achieve control with standard doses of ICS alone. There is strong evidence of a favourable risk-benefit ratio for regular ICS in reducing symptoms, improving quality of life and reducing risks of asthma attacks, hospitalisations and death. However, despite these proven benefits, non-adherence to treatment is common. On average patients prescribed regular ICS receive prescriptions for less than half the number of inhalers they need for regular treatment each year. Non-adherence is associated with poor outcomes and increased risk in patients of all levels of asthma severity, including those with the most difficult to control asthma.

Non-adherence occurs for a variety of reasons, some intentional and some non-intentional, often relating to patient beliefs, health literacy and to clinician-patient communication. When recognised, poor adherence can be improved through various communication and management strategies, including shared decision-making and personal asthma action plans. GP computerised repeat prescribing systems allow an objective record of refill prescriptions for ICS and other medication to be accessed by clinicians, and can be assessed as part of a structured asthma review. This review investigates the best method of monitoring adherence to treatment.

28.2 Review question: In people with asthma, what is the clinical and cost-effectiveness of monitoring adherence to treatment?

For full details see review protocol in Appendix A.

Table 96: PICO characteristics of review question

Population / Target Condition	People with asthma (defined as physician Dx with objective test or recurrent persistent wheeze in <5 years). If insufficient evidence is found we will consider evidence from studies where asthma is defined as physician Dx only or on asthma treatment. All ages, stratified into the following 3 different groups: • Children (1-<5 years old) • Children/young people (5-16 years old) • Adults (>16 years old)
Intervention(s)	Monitoring adherence/compliance/concordance using the following methods and provide patient feedback or intervention to improve: • Adherence with repeat therapy (using prescription and refill data) • Electronic monitoring inhalers (to monitor inhaler use) • Prednisolone levels (serum and urine – when on prednisolone) • MARS questionnaire (medication adherence rating scale) • FeNO levels (comes down if patients are taking their inhalers) • Theophylline levels (when on theophylline)
Comparison(s)	• No monitoring of adherence • Usual care • Comparison of different frequencies of monitoring adherence
Outcomes	Critical outcomes: • Mortality

	• Unscheduled healthcare utilisation (ED/A&E visit; hospital admissions; GP out of hours or walk-in centre) • Exacerbations (defined as need for course of oral steroids) • Asthma control questionnaires (ACT; CACT; ACQ; PACQ; RCP-3) • QoL (AQLQ; pAQLQ; St George's respiratory questionnaire) • Adherence Important outcomes: • Lung function (FEV1, PEF) • Symptoms (annual symptom free days) • Dose of regular asthma therapy / preventer medication (ICS dose) • Rescue medication (SABA use) • Time off school or work
Study design	RCTs

28.3 Clinical evidence

We searched for randomised trials comparing the effectiveness of monitoring adherence with feedback vs. no monitoring of adherence/usual care to guide asthma treatment and management.

Four studies were included in the review[25,119,122,196], these are summarised in Table 97 below. See also the study selection flow chart in Appendix D, study evidence tables in Appendix G, forest plots in Appendix J, GRADE tables in Appendix I and excluded studies list in Appendix K. Three studies used an electronic monitoring device in the inhaler to monitor adherence to treatment plus feedback from the physician. One study used prescription refill data to monitor adherence, with physician access to this data during review. No relevant RCTs were identified that monitored adherence using prednisolone levels, MARS questionnaires, FeNO levels or theophylline levels compared to no monitoring.

In children age 5-16 years, evidence comparing monitoring adherence plus feedback vs. no monitoring to guide ongoing management was available from two studies[25,122] summarised in the clinical evidence summary (Table 98). These studies were in children with uncontrolled asthma and evidence was available for outcomes reported at both <6months and ≥6 months.

In adults age >16 years, evidence comparing monitoring adherence plus feedback vs. no monitoring to guide ongoing management was available from two studies[119,196] summarised in the clinical evidence summary (Table 99). These studies were in adults with mixed level of asthma control and evidence was available for outcomes reported at both <6months and ≥6 months.

In children age 1-<5 years, no relevant clinical studies comparing monitoring adherence plus feedback vs no monitoring were identified.

Table 97 also summarises additional education interventions received by the intervention or comparator groups. In studies where both the intervention and comparator groups receive education, the monitoring intervention may show reduced effectiveness as the control group might also be expected to show improvement due to the education (saturation effects).

One study[119] was downgraded due to including an atypically high number of patients with severe asthma. Other limitations of the studies included a small sample size and short follow-up period in some studies. Adherence monitoring using the two methods reported (electronic recording of actuations or refill prescriptions) are indirect measures of adherence and do not necessarily ensure that the patient is taking the prescribed dose. A further limitation of Williams 2010 is that not all the patients in the intervention group had their adherence data viewed by their physician.

Table 97: Summary of studies included in the review

Study	Intervention	Comparison	Population	Outcomes	
BURGESS 2010[25] RCT	Electronic monitoring device. Adherence shared with child and carer and incorporated into the management plan (direct feedback from physician)	Usual care (no adherence feedback)	CHILDREN (6-14 years). Unstable asthma (not well controlled despite preventative medication)	• Adherence • Exacerbation • Rescue medication	In both groups: personalised asthma education and generic written information. Personalised asthma management plan devised, assessment of inhaler technique.
ONYIRIMBA 2003[119] RCT	Electronic monitoring device. Received direct feedback on ICS use from the clinician investigator and discussion of techniques to improve adherence (in addition to standard asthma care)	Usual care (no adherence feedback)	ADULTS. Moderate to severe asthma with regular ICS and low socioeconomic status	• QOL • Lung function	In both groups: if necessary, ICS switched to a twice daily regime; 3 week intensive asthma education (four 30-60min sessions) by a nurse and/or respiratory therapist blinded to the patient group. Physician input, therapy adjustment and implementation of a management plan based on PEF or symptoms at these sessions.
OTSUKI 2009[122] RCT	Electronic monitoring device. Feedback of adherence, goal-setting and reinforcement of adherence goals and strategies for self-monitoring of medication use	Usual care (no adherence feedback)	CHILDREN (2-12 years). Phys Dx asthma and 2 ED visits or 1 hospitalisation in last year	• Adherence (self-reported) • Adherence (refill) • Exacerbation • UHU	In both groups: Home-based asthma education programme (five 30min home visits by trained asthma educators; review of asthma regime; training in inhaler technique; development of asthma action plan and other education materials).
WILLIAMS 2010[196] Cluster RCT	Prescription refill adherence. Physicians provided with adherence information when reviewing and writing prescriptions.	Usual care (no adherence feedback)	ADULTS and CHILDREN (5-56 years). At least one asthma Dx and on ICS.	• Adherence • Exacerbation • UHU	In both groups: GP practices in both groups received information on the most recent asthma guidelines, and methods for discussing nonadherence with their patients.

Table 98: Clinical evidence summary: Children (5-16 years) with uncontrolled asthma: Monitoring adherence + feedback vs no monitoring.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with UC + treatment	Risk difference with Children with uncontrolled asthma: Monitoring adherence + treatment (95% CI)
Adherence <6months % of prescribed doses measured by	26 (1 study)	⊕⊖⊖⊖ **VERY LOW**[1,2]		The mean adherence <6months in the control groups was	The mean adherence <6months in the intervention groups was

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects	
the electronic inhaler	4 months	due to risk of bias, imprecision		**55.3 %**	**28.9 higher** (8.62 to 49.18 higher)
Adherence ≥6months Number of canister refills (100% adherence = 3.0). Scale from: 0 to 3.	157 (1 study) 18 months	⊕⊕⊖⊖ **LOW**[1] due to risk of bias		The mean adherence ≥6months in the control groups was **0.6 canister refills**	The mean adherence ≥6months in the intervention groups was **0.02 lower** (0.29 lower to 0.25 higher)
Adherence (self-reported) ≥6months % self-reported adherence in previous 6 months. Scale from: 0 to 100.	157 (1 study) 18 months	⊕⊕⊖⊖ **LOW**[1] due to risk of bias		The mean adherence (self-reported) ≥6months in the control groups was **85.4 %**	The mean adherence (self-reported) ≥6months in the intervention groups was **1.95 higher** (5.87 lower to 9.77 higher)
Exacerbation < 6months need for OCS	26 (1 study) 4 months	⊕⊖⊖⊖ **VERY LOW**[1,3] due to risk of bias, imprecision	**RR 2.57** (0.31 to 21.59)	**Moderate** **83 per 1000**	**130 more per 1000** (from 57 fewer to 1000 more)
Exacerbation ≥6 months no. of OCS courses in 6 months	157 (1 study) 18 months	⊕⊕⊕⊖ **MODERATE**[4] due to risk of bias		The mean exacerbation ≥6 months in the control groups was **0.74 courses of OCS**	The mean exacerbation ≥6 months in the intervention groups was **0.22 higher** (0.19 lower to 0.63 higher)
UHU ≥6 months Hospitalisations in previous 6 months	157 (1 study) 18 months	⊕⊕⊕⊖ **MODERATE**[4] due to risk of bias		The mean uhu ≥6 months in the control groups was **12**	The mean uhu ≥6 months in the intervention groups was **0 higher** (4.8 lower to 4.8 higher)
Rescue medication < 6months Reliever medication 3 or more times a week	26 (1 study) 4 months	⊕⊖⊖⊖ **VERY LOW**[1,3] due to risk of bias, imprecision	**OR 6.92** (0.41 to 118.14)	**Moderate** **0 per 1000**	**140 more per 1000** (from 7 more to 360 more)[5]

[1] The majority of the evidence was from studies at very high risk of bias
[2] 95% CI crosses one MID
[3] 95% CI crosses both MIDs
[4] The majority of the evidence was from studies at high risk of bias
[5] Manual calculation of absolute risk difference as no events in the control group

Table 99: Clinical evidence summary: Adults (>16 years) overall: Monitoring adherence + feedback vs no monitoring.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with UC + treatment	Risk difference with Adults overall: Monitoring adherence + treatment (95% CI)
Adherence ≥6months % adherence to prescription refills in previous 3 months. Scale from: 0 to 100.	0 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean adherence in the control groups was **23.3 %**	The mean adherence in the intervention groups was **2 lower** (8.61 lower to 4.61 higher)
QOL <6months AQLQ. Scale from: 1 to 7.	19 (1 study) 10 weeks	⊕⊖⊖⊖ **VERY LOW**[3,4,5] due to risk of bias,		The mean QOL in the control groups was **4.51**	The mean QOL in the intervention groups was **0.37 higher**

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects	
		indirectness, imprecision			(0.08 to 0.66 higher)
Exacerbation ≥6months course of OCS	2698 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[1] due to risk of bias	**HR 1.07** (0.89 to 1.29)	**Moderate** **220 per 1000**	**13 more per 1000** (from 22 fewer to 54 more)
UHU (hospitalisation) ≥6months	2698 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,6] due to risk of bias, imprecision	**HR 0.86** (0.32 to 2.31)	**Moderate** **8 per 1000**	**1 fewer per 1000** (from 6 fewer to 11 more)
UHU (ED visit) ≥6months	2698 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,5] due to risk of bias, imprecision	**HR 1.22** (0.83 to 1.79)	**Moderate** **81 per 1000**	**17 more per 1000** (from 13 fewer to 59 more)
Lung function <6months FEV1 L	19 (1 study) 10 weeks	⊕⊖⊖⊖ **VERY LOW**[2,3,4] due to risk of bias, indirectness, imprecision		The mean lung function in the control groups was **0.16 L**	The mean lung function in the intervention groups was **0.12 lower** (7.31 lower to 7.07 higher)

[1] The majority of the evidence was from studies at very high risk of bias
[2] 95% CI crosses both MIDs
[3] The majority of the evidence is from studies at very high risk of bias
[4] Population indirectness: includes severe asthma
[5] 95% CI crosses one MID
[6] 95% CI crosses both the MIDs but only downgraded by one as the 95% CI for the absolute effect is small

28.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow chart in Appendix E.

28.5 Evidence statements

Clinical

CHILDREN (5-16 years)

- No evidence was identified on mortality, quality of life or asthma control questionnaire outcomes.
- Monitoring adherence vs no monitoring was considered a clinically important benefit for adherence measured by the percentage of doses registered by the inhaler (1 study, N=26, very low quality), at <6 months
- Monitoring adherence vs no monitoring was considered a clinically important harm for asthma exacerbations (number of patients requiring OCS, 1 study, N=26, very low quality) and use of rescue medications (1 study, N=26, very low quality), both at <6 months.
- Monitoring adherence vs no monitoring resulted in no clinically important difference for adherence measured by the number of canister refills (1 study, N=157, low quality), adherence measured by percentage self-reported adherence (1 study, N=157, low quality) and UHU hospitalisations (1 study, N=157, moderate quality), all at ≥6 months.
- Monitoring adherence vs no monitoring resulted in a borderline clinically important difference for asthma exacerbations at ≥6 months (mean number of OCS courses in 6 months, 1 study, N=157, moderate quality).

ADULTS (>16 years)

- No evidence was identified for mortality or asthma control questionnaire outcomes.
- Monitoring adherence vs no monitoring resulted in no clinically important difference for adherence measured percentage of prescription refills (1 study, N=2698, very low quality) and UHU hospitalisations (1 study N=2698, very low quality), both and ≥6 months and for lung function measured using FEV1 (1 study, N=19, very low quality), at <6 months.
- Monitoring adherence vs no monitoring resulted in a borderline clinically important difference for asthma exacerbations (number of patients requiring OCS, 1 study N=2698, low quality) and UHU ED visits (1 study N=2698, very low quality), both at ≥6 months and for QOL (1 study, N=19, very low quality) at <6 months.

Economic

- No relevant economic evaluations were identified.

28.6 Recommendations and link to evidence

Recommendations	**No clinical recommendation.**
Research recommendations	**4. What is the clinical and cost effectiveness of using electronic alert systems designed to monitor and improve adherence with regular**

	inhaled maintenance therapy in people with asthma?
Relative values of different outcomes	The GC considered the following outcomes as critical for this review: adherence to regular ICS treatment, mortality, quality of life, exacerbations, asthma control measured by questionnaires and unscheduled healthcare utilisation. The GC acknowledged that the aim of studies monitoring adherence to regular ICS treatment is often to improve adherence itself. Therefore, the outcome of adherence to treatment itself was considered to be an important outcome for this question. Whilst this does not directly provide evidence of asthma control, poor adherence to treatment is associated with poor outcomes. The GC noted that exacerbations of asthma can lead to both unscheduled healthcare utilisation (ED visits, out-of-hours GP surgery visits or hospitalisation) and the need for a course of OCS. Therefore, exacerbations requiring OCS treatment and exacerbations requiring unscheduled healthcare utilisation were considered separately. The following important outcomes were considered: lung function (FEV1), symptoms (symptom-free days), dose of regular asthma therapy (ICS dose), rescue medication (SABA dose) and time off school or work. These outcomes were considered as important measures of asthma control for the patient. The GC acknowledged that regular asthma preventer therapy (ICS) is often underprescribed / used, and rescue medications (SABA) may be overprescribed / used[144].
Trade-off between clinical benefits and harms	In adults, whilst the effect estimate did not reach the established MID, the GC agreed that an AQLQ QOL score of 0.37 higher may represent a clinically important benefit, although evidence was from one small study. Evidence showed no convincing difference between monitoring and no monitoring of the adherence to treatment and rate of hospitalisation. There was evidence of an increase in the rate of exacerbations and ED visits in the monitoring group; however, it was unclear if this represented a clinically important harm. The GC acknowledged that on some occasions OCS may be of benefit by preventing a more severe asthma attack requiring UHU. Evidence from one small study with N=19 showed that monitoring adherence resulted in no clinically important difference in lung function. No evidence was available on mortality and asthma control using validated questionnaires. In children, evidence from one study showed that monitoring adherence resulted in a clinically important benefit in adherence measured by the number of prescribed doses administered by the electronic inhaler.. There was evidence from one small study of an increase in the rate of exacerbations in the monitoring group, which may represent a clinically important harm. Evidence showed no convincing difference between monitoring and no monitoring in the rate of unscheduled healthcare utilisation, and adherence to treatment (measured by self-report or the percentage of refills taken). No evidence was identified on mortality, quality of life or asthma control questionnaire outcomes. No evidence was identified in children aged 1-<5 years old. Overall, the GC agreed that there was not enough available evidence to weigh up the benefits and harms and make a recommendation to monitor adherence to treatment. Monitoring adherence and providing feedback to the patient does not have any direct safety implications to the patient. The GC believed it to be self-evident that good adherence to treatment would be associated with better outcomes, but future research is needed to establish the clinical effectiveness of systems that alert healthcare professionals to poor patient adherence.

Economic considerations	No economic evaluations were identified. Prescribing all individuals with electronic monitoring devices would increase NHS costs. The GC agreed, however, that the evidence on the use of these devices was weak and gave no true indication of how effective and therefore cost-effective these devices were. Although in some reviews monitoring adherence showed that the use of oral steroids increased, the GC considered that this occurred because an individual's poor asthma control was identified early and treated accordingly. Therefore, although it may appear that monitoring adherence increases NHS costs, in fact, it may reduce costs in the long run and improve health outcomes. Overall, the clinical evidence gave no strong indication of how cost-effective monitoring adherence could be or how it should be conducted and therefore the GC agreed that this area would benefit greatly from a future research recommendation.
Quality of evidence	Evidence for each outcome was only available from one study and the majority of the evidence was of low and very low quality by GRADE criteria. In adults, all evidence was of low and very low quality. Lung function and QOL outcomes were only available from one study with a small sample size. The other outcomes were also only available from one study: a large study with a cluster randomised design. In this study, it was noted that only a proportion of physicians accessed the adherence data during the patient review. The GC noted that one study (Onyrimba 2003) stated that patients could be switched to a twice-daily ICS regimen if necessary, suggesting some patients started on a once-daily regimen. The GC questioned the applicability and directness of this study as only certain ICS drugs can be used once daily. In children, the evidence was of low and very low quality with the exception of the outcomes for UHU and exacerbations. However, the exacerbations outcome was reported as the mean number of OCS courses in 6 months and the GC was uncertain about how to interpret the evidence for exacerbations as a continuous outcome. Only one study contributed to the evidence for each outcome and the studies were of small sample size. No evidence was identified in children aged 1-<5 years old. The GC believed the uncertainty in the available evidence for all outcomes was sufficient to justify delaying a recommendation to await further research.
Other considerations	The monitoring interventions reported in the studies were complex interventions which did not just monitor adherence in isolation. The GC noted that it was hard to look at monitoring adherence in isolation outside of the clinical care provided. It was noted that some studies included an educational component in the control group, and that the effect of monitoring adherence may be saturated due to improved outcomes in the control group. Whilst the GC did not look at the individual evidence from prognostic studies of adherence as a risk factor for future outcomes, the GC was aware of and discussed key prognostic studies showing that poor adherence predicts future risk[54,145,170,171]. The GC also considered the NRAD audit[144], which reported poor adherence in a large number of those who died from asthma, and an association between non-adherence to preventer inhaled corticosteroids and increased risk of poor asthma control.

	The GC considered that adherence to preventer treatment is an important area of asthma care and should be regularly monitored in all patients. They were disappointed not to be able to make a recommendation about how best to do this. The GC decided to make a high-priority research recommendation to investigate the clinical and cost-effectiveness of using electronic alert systems designed to monitor and improve adherence with regular inhaled maintenance therapy in people with asthma, and discussed the alignment of the research recommendation with the recommendations from the NRAD report. The NRAD recommends that electronic surveillance of prescribing in primary care should be introduced as a matter of urgency to alert clinicians to patients being prescribed excessive quantities of short-acting reliever inhalers, or too few preventer inhalers. Further details on the high-priority research recommendation made can be found in appendix N, along with the full list of research recommendations made. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. Two studies were identified which the GC agreed did not suggest a change in the recommendations was warranted.

29 Monitoring inhaler technique

29.1 Introduction

The selection of an appropriate inhaler device is an important part of pharmacotherapy for asthma management. With all inhalers, correct technique is essential for ensuring appropriate (or proper) delivery of treatment. There should be proper understanding of, and training in, inhaler technique for patients, parents and/or carers. It is essential for healthcare professionals such as GPs, practice nurses, asthma nurse specialists, health visitors, school nurses, hospital doctors and nurses, community and hospital pharmacists and pharmacy technicians dealing with people with asthma-related medical problems to have an equally good understanding, so that they can provide education and support.

This review investigates the best method of monitoring inhaler technique, and the frequency with which this should be applied.

29.2 Review question: In people with asthma, what is the optimal frequency and method for monitoring inhaler technique?

For full details see review protocol in Appendix C.

Table 100: PICO characteristics of review question

Population	People with asthma (defined as physician Dx with objective test or recurrent persistent wheeze in <5 years). If insufficient evidence is found we will consider evidence from studies where asthma is defined as physician Dx only or on asthma treatment (give details of how Dx was made). Not including severe asthma. All ages, stratified into the following 3 different groups: • Children (1-<5 years old) • Children/young people (5-16 years old) • Adults (>16 years old)
Intervention(s)	Monitoring inhaler technique using the following methods and provide patient feedback or intervention to improve inhaler technique: • Electronic devices to monitor inhaler technique • Visual monitoring by doctor, nurse or pharmacist
Comparison(s)	• No monitoring of inhaler technique • Comparison of different frequencies of monitoring inhaler technique • Monitoring using electronic devices vs monitoring by visual inspection
Outcomes	Critical outcomes: • Mortality • Unscheduled healthcare utilisation (ED/A&E visit; hospital admissions; GP out of hours or walk-in centre) • Exacerbations (defined as need for course of oral steroids) • Asthma control questionnaires (ACT; CACT; ACQ; PACQ; RCP-3) • QoL (AQLQ; pAQLQ; St George's respiratory questionnaire) Important outcomes: • Lung function (FEV1, PEF) • Symptoms (annual symptom free days) • Dose of regular asthma therapy / preventer medication (ICS dose) • Rescue medication (SABA use)

	• Time off school or work
Study design	RCTs

29.3 Clinical evidence

We searched for randomised trials comparing the effectiveness of monitoring inhaler technique with feedback vs. no monitoring of inhaler technique. We also searched for randomised trials comparing the effectiveness of monitoring inhaler technique using different methods (visual inspection by a healthcare professional with verbal feedback or monitoring inhaler technique using electronic devices with feedback).

Four studies were included in the review[4,5,14,15], these are summarised in Table 101 below. See also the study selection flow chart in Appendix D, study evidence tables in Appendix G, forest plots in Appendix J, GRADE tables in Appendix I and excluded studies list in Appendix K.

Evidence was identified for the following strata and comparisons:

In adults:

- Comparison 1: visual monitoring of inhaler technique with verbal feedback plus use of electronic training device vs. visual monitoring of inhaler technique with verbal feedback alone (2 studies, one in primary care[5] and one in secondary care[4]).
- Comparison 2: visual monitoring of inhaler technique (and PEF meter technique) by pharmacist plus feedback vs. monitoring of PEF meter technique only (1 study[14,15]).

In children:

- Comparison 1: visual monitoring of inhaler technique with verbal feedback plus use of electronic training device vs. visual monitoring of inhaler technique with verbal feedback alone (1 study[5]).

For comparison 1 in both adults and children, the aim of monitoring in both studies was to slow down the inhalation rate in people with poor inhaler technique due to fast inhalation flow rate (IFR). For the study which provided evidence for both adults and children[5], the population included those with poor inhaler technique (defined as poor hand-lung coordination and an IFR ≥90 l/min). The other study in adults[4] was in a population with good coordination but poor inhaler technique defined as a fast IFR ≥90 l/min.

Table 101: Summary of studies included in the review

Study	Intervention/comparison	Population	Outcomes	Comments
Al-showair 2007[4]	Verbal training + 2Tone Trainer (2TT) vs verbal training alone. Verbal training on the most desirable inhalation technique + 2Tone Trainer every morning and night to obtain the one-tone sound and to use the same inhalation procedure when using their MDI.	Adults - Secondary care Identified with poor inhaler technique (good coordination but inhaled too fast IFR ≥90l/min).	QOL Lung function	1 visit, 6 weeks follow-up (intervention group encouraged to practice with 2TT twice daily before taking their MDI).
Ammari 2013 [5]	Verbal training + 2TT vs verbal training alone. Verbal training on the most desirable inhalation technique + 2Tone Trainer every morning and night to obtain the one-tone sound and to use the same inhalation procedure when using their MDI.	Adults - Primary care Identified with poor inhaler technique (defined as poor hand-lung coordination and an IFR ≥90l/min).	QOL Lung function	1 visit, 6 weeks follow-up (intervention group encouraged to practice with 2TT twice daily before taking their MDI).
		Children - Primary care Identified with poor inhaler technique (defined as poor hand-lung coordination and an IFR ≥90l/min).	QOL Lung function	
Basheti 2007 [14,15]	Monitoring PEF meter and inhaler technique + feedback vs monitoring PEF meter technique only Pharmacy trained to deliver education on PEF meter technique and inhaler technique. Assessed inhaler technique using checklists and then educated using 'show and tell' for each step on the checklist. Incorrect steps on the checklist were highlighted and attached to the patient's inhaler using a label.	Adults – Community pharmacy Doctor diagnosed asthma; use of ICS with Turbuhaler or Diskus with or without LABA.	**<6 months** QOL Lung function **≥6 months** QOL Lung function	Training at 0, 1, 2, 3 and 6 months. Follow-up 6 months.

Table 102: ADULTS: Monitoring inhaler technique compared to no monitoring for asthma

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with No monitoring	Risk difference with ADULTS: Monitoring inhaler technique (95% CI)

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects	
Lung function <6 months PEF Min%Max (higher is less variability). Scale from: 0 to 100.	97 (1 study) 3 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean lung function <6 months in the control groups was **77.6 %**	The mean lung function <6 months in the intervention groups was **6.2 higher** (2.68 to 9.72 higher)
Lung function ≥6 months PEF Min%Max (higher is less variability). Scale from: 0 to 100.	97 (1 study) 6 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean lung function ≥6 months in the control groups was **74.4 %**	The mean lung function ≥6 months in the intervention groups was **4.5 higher** (0.79 to 8.21 higher)
QOL <6 months Marks AQLQ. Scale from: 0 to 10.	97 (1 study) 3 months	⊕⊕⊖⊖ **LOW**[1] due to risk of bias		The mean qol <6 months in the control groups was **1.35**	The mean qol <6 months in the intervention groups was **0.55 lower** (0.77 to 0.33 lower)
QOL ≥6 months Marks AQLQ. Scale from: 0 to 10.	97 (1 study) 6 months	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean qol ≥6 months in the control groups was **1.3**	The mean qol ≥6 months in the intervention groups was **0.5 lower** (0.74 to 0.26 lower)

[1] The evidence was from one study at very high risk of bias for this outcome

[2] 95% CI crosses one MID

Table 103: ADULTS: Monitoring (verbal and electronic) compared to verbal monitoring only for asthma

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Verbal monitoring only	Risk difference with ADULTS: Monitoring (verbal and electronic) (95% CI)
QOL <6 months mini AQLQ. Scale from: 1 to 7.	105 (2 studies) 6 weeks	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean qol <6 months in the control groups was **4.2**	The mean qol <6 months in the intervention groups was **0.38 higher** (0.02 lower to 0.79 higher)
Lung function <6 months FEV1 L	71 (1 study) 6 weeks	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean lung function <6 months in the control groups was **2.16 L**	The mean lung function <6 months in the intervention groups was **0.23 lower** (0.55 lower to 0.09 higher)
Lung function <6 months FEV1 % pred. Scale from: 0 to 100.	34 (1 study) 6 weeks	⊕⊕⊖⊖ **LOW**[2,3] due to risk of bias, imprecision		The mean lung function <6 months in the control groups was **87.2 %**	The mean lung function <6 months in the intervention groups was **9.1 higher** (3.71 lower to 21.91 higher)

[1] The majority of the evidence was from studies at very high risk of bias for this outcome

[2] 95% CI crosses one MID

[3] The majority of the evidence was from studies at high risk of bias for this outcome

Table 104: CHILDREN: Monitoring (verbal and electronic) compared to verbal monitoring only for asthma

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Verbal monitoring only	Risk difference with CHILDREN: Monitoring (verbal and electronic) (95% CI)
Lung function <6 months FEV1 % pred. Scale from: 0 to 100.	12 (1 study) 6 weeks	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, imprecision		The mean lung function <6 months in the control groups was **94.1 %**	The mean lung function <6 months in the intervention groups was **3.2 lower** (15.27 lower to 8.87 higher)
QOL <6 months PAQLQ. Scale from: 1 to 7.	12 (1 study) 6 weeks	⊕⊖⊖⊖ **VERY LOW**[2,3] due to risk of bias, imprecision		The mean qol <6 months in the control groups was **-0.391 change score**	The mean qol <6 months in the intervention groups was **0.03 higher** (0.66 lower to 0.72 higher)

[1] The evidence was from one study at high risk of bias for this outcome

[2] 95% CI crosses both MIDs

[3] No explanation was provided

29.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

See also the economic article selection flow chart in Appendix E.

29.5 Evidence statements

Clinical

ADULTS (>16 years): monitoring inhaler technique compared to no monitoring

- No evidence was identified for mortality, UHU, exacerbations or asthma control questionnaires.
- Monitoring inhaler technique vs. no monitoring resulted in a borderline clinically important difference forQOL at <6 months and ≥6 months (1 study, N=97, low to very low quality) and PEF variability at <6 months and ≥6 months (1 study, N=97, very low quality)

ADULTS (>16 years): monitoring inhaler technique (verbal and electronic trainer device) compared to verbal monitoring only

- No evidence was identified for mortality, UHU, exacerbations or asthma control questionnaires.
- Monitoring inhaler technique (verbal and electronic trainer device) vs. verbal monitoring only was considered a clinically important benefit for QOL at (2 studies, N=105, very low quality) and lung function measured using FEV1 %pred (1 study, N=34, low quality), but not when measured using FEV1 litres where a clinical harm was observed (1 study, N=71, very low quality), all at <6 months.

CHILDREN (5-16 years): monitoring inhaler technique (verbal and electronic trainer device) compared to verbal monitoring only

- No evidence was identified for mortality, UHU, exacerbations or asthma control questionnaires.
- Monitoring inhaler technique (verbal and electronic trainer device) vs. verbal monitoring only resulted in no clinically important difference for QOL and lung function measured using FEV1 (both 1 study, N=12, very low quality), both at <6 months.

Economic

- No relevant economic evaluations were identified.

29.6 Recommendations and link to evidence

Recommendations	**45. Observe and give advice on the person's inhaler technique:** • **at every consultation relating to an asthma attack, in all care settings** • **when there is deterioration in asthma control** • **when the inhaler device is changed** • **at every annual review** • **if the person asks for it to be checked.**
Research recommendations	**5. What is the current frequency and the current method being used to check the inhaler technique of people with asthma? What is the optimal**

	frequency and the best method of checking inhaler technique to improve clinical outcomes for people with asthma?
Relative values of different outcomes	The GC considered the following outcomes as critical for this review: mortality, quality of life, exacerbations, asthma control measured by questionnaires and unscheduled healthcare utilisation. The GC noted that exacerbations of asthma can lead to both unscheduled healthcare utilisation (ED visits, out-of-hours GP surgery visits or hospitalisation) and the need for a course of OCS. Therefore, exacerbations requiring OCS treatment and exacerbations requiring unscheduled healthcare utilisation were considered separately. The GC also considered the following important outcomes: lung function (FEV1), symptoms (symptom-free days), dose of regular asthma therapy (ICS dose), rescue medication (SABA dose) and time off school or work. These outcomes were considered as important measures of asthma control for the patient. The GC acknowledged that regular asthma preventer therapy (ICS) is often under-prescribed / used, and rescue medications (SABA) may be overprescribed / used[144].
Trade-off between clinical benefits and harms	<u>Monitoring inhaler technique</u> RCT evidence was only identified from one study in adults for the comparison of monitoring inhaler technique vs. no monitoring. There was a difference in QOL at both 3 months and 6 months. QOL scores on the Marks AQLQ scale were lower in the monitoring group, indicating better QOL; however, there is no established MID for this scale and the GC was unsure of the clinical importance. There was less PEF variability in the monitoring group at both 3 and 6 months, which may represent a clinically important benefit. No evidence was available for mortality, UHU, exacerbations or asthma control questionnaire scores. <u>Monitoring inhaler technique plus the use of an electronic training device</u> RCT evidence was also identified comparing monitoring of inhaler technique plus the use of an electronic training device vs monitoring inhaler technique alone in both adults and children. The aim of monitoring in these studies was to decrease the inhalation rate in people with poor inhaler technique due to fast inhalation flow rate. For the study which provided evidence for both adults and children, the population included those with poor inhaler technique (defined as poor hand-lung coordination and an IFR ≥90 l/min). The other study in adults was in a population with good coordination but poor inhaler technique defined as a fast IFR ≥90 l/min. In adults, monitoring inhaler technique (verbal and electronic trainer device) compared to verbal monitoring resulted in a clinically important benefit in QOL at 6 weeks. Evidence was available from two studies. For the additional use of the electronic training device, one study showed a clinically important harm in the secondary outcome of FEV1 at 6 weeks and the other study showed a clinically important benefit in FEV1 at 6 weeks. No evidence was available for mortality, UHU, exacerbations or asthma control questionnaire scores. In children and young people, monitoring inhaler technique (verbal and electronic trainer device) compared to verbal monitoring resulted in no clinically important difference in QOL at 6 weeks. Monitoring using the electronic training device resulted in a lower FEV1 at 6 weeks, but this difference was not of clinical benefit.
Economic considerations	No economic evidence was found on monitoring inhaler technique. The cost of monitoring inhaler technique is negligible as this could be carried out as part of routine visits. The clinical review showed that in adults monitoring inhaler technique (verbal and electronic trainer device) compared to verbal monitoring resulted in a clinically

	important benefit in QOL at 6 weeks. Therefore the GC considered monitoring inhaler technique likely to be cost-effective.
Quality of evidence	In adults, for monitoring inhaler technique vs. no monitoring, evidence for the important and critical outcomes was of low and very low quality by GRADE criteria. Only one study contributed to the evidence for all outcomes, and this study was of small sample size. In adults, for monitoring inhaler technique (verbal and electronic trainer device) compared to verbal monitoring only, evidence for the important and critical outcomes was of low and very low quality. Evidence was available from two small studies with a short follow-up time of 6 weeks. In children and young people, for monitoring inhaler technique (verbal and electronic trainer device) compared to verbal monitoring only, evidence for the important and critical outcomes was of very low quality. Only one very small study contributed to the evidence for all outcomes with a short follow-up time of 6 weeks.
Other considerations	The GC discussed the importance of good inhaler technique. Whilst the GC did not look at the evidence from prognostic studies of poor inhaler technique as a risk factor for future outcomes, the GC was aware of and discussed the existence of prognostic studies within the broader literature base showing that poor inhaler technique predicts future risk. The GC was in agreement with the NICE Quality Standard for asthma that inhaler technique should be assessed after every attack, with every change of inhaler device and at every annual asthma review. They also agreed that it is important to check inhaler technique when there is deterioration in asthma control or when it is requested by the patient. They made a recommendation based on consensus agreement (not the evidence alone). The GC agreed it was unethical to make a future research recommendation to compare monitoring inhaler technique vs. no monitoring of inhaler technique and hence made a recommendation on appropriate timing for checking inhaler technique. The GC had concerns that, for asthma reviews performed by phone, people would not have their inhaler technique checked in the situations recommended. Therefore, the GC chose the wording of the recommendation carefully, that inhaler technique should be 'observed'. The GC agreed that healthcare professionals need to be regularly trained in inhaler technique in order to monitor inhaler technique effectively. The GC was interested in the best method of monitoring inhaler technique. Due to the absence of evidence, the GC made a high-priority future research recommendation to assess the best method for monitoring inhaler technique. Further details on the high-priority research recommendation made can be found in appendix N. The GC was aware of additional observational studies using different methods of monitoring inhaler technique, and the effect on the outcome of inhaler technique score itself. However, studies of this sort were excluded as they do not report the relevant efficacy outcome and were short-term observational studies. The GC concurred that RCT evidence was needed to assess the long-term benefit of different methods of monitoring inhaler technique. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. No new information was identified for this review question.

30 Monitoring: Tele-healthcare

30.1 Introduction

Tele-healthcare is the utilisation of information and communication technologies by patients and healthcare professionals to deliver clinical care, health promotion or to carry out research where the participants are not in the same location. The information shared between participants may be stored and used later or may be used interactively to make a diagnosis, to monitor a condition or to enable the patient to adjust a clinical management plan. Tele-healthcare has the potential to improve monitoring of asthma by increasing accessibility of care for patients and supporting effective self-management, reducing cost and detecting exacerbations or loss of asthma control sooner. However, there are also risks involved with the use of tele-healthcare for monitoring asthma, and the benefits and harms need to be considered.

30.2 Review question: In people with asthma, what is the clinical and cost-effectiveness of tele-healthcare to monitor asthma control?

For full details see review protocol in Appendix C.

Table 105: PICO characteristics of review question

Population	Children and adults with clinician-diagnosed asthma
Intervention(s)	Tele-healthcare interventions (review divided into two sections): • Tele-healthcare with healthcare professional involvement • Tele-healthcare with no involvement from a healthcare provider
Comparison(s)	Usual care or any other control intervention
Outcomes	Critical outcomes: • Mortality • Exacerbations requiring hospitalisation • Exacerbations (defined as need for course of oral steroids) • Unscheduled healthcare visits • QoL (AQLQ; pAQLQ; St George's respiratory questionnaire) • Asthma control questionnaires (ACT; CACT; ACQ; PACQ; RCP-3) Important outcomes: • Lung function (FEV1, PEF) • Symptoms (annual symptom free days)
Study design	Full reports of parallel randomised controlled trials

Tele-health interventions with healthcare professional involvement and interventions with no involvement from a healthcare provider were dealt with separately (see section 30.3.2). The cost-effectiveness of fully automated interventions is likely to be very different to those which include personalised feedback from a health professional. Both reviews had the same study inclusion criteria with respect to population, comparison, outcomes and study design.

30.3 Clinical evidence

30.3.1 Tele-healthcare with healthcare professional involvement

We searched for randomised controlled trials comparing tele-healthcare interventions delivered with input from a healthcare provider with usual care or a control intervention.

Twenty-five studies met the review eligibility criteria;[12,13,28,31,48,50,61-64,74,83,99,121,132,133,139,142,146,154,184,190,194,198,200] these are summarised in Table 106 below. See also the study selection flow chart in Appendix D, study evidence tables in Appendix G, forest plots in Appendix J, GRADE tables in Appendix I and excluded studies list in Appendix K.

Thirteen of the 16 adult studies and eight of the 10 children studies reported data that could be included in meta-analysis for one or more of the clinical outcomes.

Studies were analysed in three separate comparisons which were identified as addressing three questions relating to the use of tele-healthcare. These comparisons were not pre-specified in the review protocol, but were necessary in order to make the analyses clinically meaningful and were constructed prior to extraction and pooling of outcome data.

1. **How do consultations conducted with tele-healthcare compare to face-to-face reviews?**
   Studies assessed the feasibility of replacing face-to-face monitoring in clinics with tele-healthcare reviews.
2. **Is tele-healthcare monitoring better than paper-based self-monitoring?**
   These studies isolated the effect of using tele-healthcare systems (e.g. an electronic diary or program) by controlling for the non-specific effects of self-monitoring.
3. **Do broad tele-healthcare packages improve health outcomes?**
   The aim of these studies was to test a complete monitoring package delivered solely or predominantly with tele-healthcare. The studies did not isolate the effect of tele-health components from non-specific effects of increased contact with healthcare services.

Tele-health interventions within each of the comparisons varied with respect to the length and type of tele-health intervention, qualifications of the health provider involved and the extent of their input, and participant inclusion and exclusion criteria. There were not enough studies to reliably explore the effect of these moderators within each comparison, so important differences have been summarised narratively.

In adults aged > 16 years, evidence for comparison one was available in four studies[61,132,133,142]. Three studies compared telephone consultations with face-to-face equivalents, either with an asthma nurse or a doctor for either six or 12 months. One study compared a six-month PEF monitoring and email advice intervention with a clinic-based equivalent. Evidence for comparison 2 was available from four studies with four to 12 months' follow-up[99,121,146,184]. Comparison 2 studies used primarily mobile phone-based interventions (SMS or smart-phone software) aimed at symptom monitoring by an asthma nurse or clinician compared with paper symptom diaries. Two studies included in comparison 1[132,142] also compared their tele-health interventions with a usual care control group which were included in comparison 3 with five other studies ranging from three months to a year[12,50,139,190,194]. Comparison 3 studies were the most varied; all used usual care or similarly minimal control groups, but interventions included monitoring and advice programs *via* telephone, internet or SMS, and hospital discharge telephone monitoring.

In children aged 5 to 16 years, evidence for comparison one was available from one study[28] comparing internet-based case management and education from a paediatrician with face-to-face sessions for one year. One child study[74] provided evidence for comparison two, comparing a three-

month internet monitoring program, PEF diary and physician feedback with a PEF diary alone. Evidence for comparison three was available from six child studies with follow-ups ranging from three months to a year[48,63,198], comparing a range of tele-healthcare packages to usual care.

No relevant studies were found comparing tele-health interventions with usual care or a control intervention for children aged less than five.

Table 106: Summary of studies included in the review

Study	Intervention/comparison	Population	Outcomes	Follow-up
Baptist 2013[12]	Asthma calls plus face to face sessions with a health educator *vs.* non-asthma calls	• Older adults • N=70 • 65+ years • Daily controller meds • 83% predicted FEV1	• Dichot. ACQ • Hospitalisation • GP visits • FEV1	6 and 12 months
Barbanel 2003[13]	Training course and follow-up pharmacist calls *vs.* routine care	• Adults • N=24 • 18-65 years • All taking ICS	• Withdrawal	6 months
Chan 2007[28]	Internet-based case management and education from a paediatrician *vs.* face-to-face sessions	• Children/adolescents • N=120 • 6-17 years • Persistent asthma	• Hospitalisations • ED visits • PAQLQ • FEV1 % predicted	12 months
Chatkin 2006[31]	Calls from physician to improve adherence *vs.* routine care	• Adults/adolescents • N=271 • 12+ years • Mod./severe asthma	• Adherence measures	Unknown follow-up
Deschildre 2012[48]	Daily FEV1 transmission via internet with physician feedback *vs.* routine care	• Children/adolescents • N=50 • 6-16 years • Severe allergic asthma, frequent exacerbations	• Hospitalisations • Oral steroid use	12 months
Donald 2008[50]	Post hospital discharge telephone PEF and symptom monitoring by nurse *vs.* routine care	• Adults • N=71 • 18-55 years • Previous asthma admission	• Hospitalisations • ED visits • GP visits • Oral steroid use • Absence	12 months
Gruffydd-Jones 2005[61]	6-monthly telephone monitoring by asthma nurse *vs.* usual 6-monthly clinic consultations	• Adults • N=194 • 17-70 years	• AQLQ • ACQ • Costs	6 and 12 months
Guendelman 2002[62]	Internet management and education program with asthma nurse *vs.* paper symptom diary	• Children/adolescents • N=134 • 8-16 years • Persistent asthma	• Hospitalisations • ED visits	3 months

Study	Intervention/comparison	Population	Outcomes	Follow-up
Gustafson 2012[63]	Automated management software with monthly calls from nurse *vs.* routine care	• Children • N=301 • 4-12 years • Controller meds and poor adherence	• ACQ	12 months
Halterman 2012[64]	Internet communication, prescription, and symptom monitoring by asthma nurse *vs.* routine care	• Children • N=100 • 3-10 years • Persistent symptoms	• Hospitalisations • ED visits • GP visits • AQLQ • School absence	8 months
Jan 2007[74]	Internet monitoring program and PEF diary with physician email or phone feedback *vs.* PEF diary and routine care	• Children • N=164 • 6-12 years	• PEF	3 months
Khan 2004[83]	Post-discharge telephone follow-up from nurse *vs.* written materials	• Children • N=310 • 1-15 years • Recent ED discharge	• Hospitalisation • ED visits • Parent QoL	6 months
Liu 2011[99]	Mobile phone software with electronic diary reviewed by staff daily *vs.* written asthma diary	• Adults • N=89 • Mean 52 years • Mod./severe asthma	• Mortality • Hospitalisations • ED visits • FEV1 and PEF • SF12	6 months
Ostojic 2005[121]	Written asthma diary and PEF send via text daily with weekly instructions from a specialist *vs.* diary only	• Adults • N=16 • Mean 25 years Moderate asthma • All using LABA/ICS	• Hospitalisations • FEV1	4 months
Pinnock 2003[133]	Telephone review *vs.* face-to-face review, both with the asthma nurse	• Adults • N=278 • 18+ years • Asthma for 1 year +	• Hospitalisation • ED visits • Oral steroid use • GP visits • AQLQ	Variable follow-up, pragmatic design
Pinnock 2007[132]	Pre-arranged phone or face-to-face review *vs.* face-to-face only *vs.* usual care (no review)	• Adults • N=1728 • 12+ years (mean 44)	• AQLQ • ACQ • Costs	12 months
Prabhakaran 2009[139]	SMS monitoring and education with the asthma nurse *vs.* education with no SMS monitoring	• Adults • N=120 • 21+ years • Previous asthma admission	• Hospitalisation • ED visits • Dichot. ACT	3 months

Study	Intervention/comparison	Population	Outcomes	Follow-up
Rasmussen 2005[142]	Electronic PEF diary and advice via email *vs.* face-to-face specialist instruction with PEF *vs.* usual GP contact	• Adults • N=300 • 18-45 years	• Hospitalisation • ED visits • GP visits • FEV1	12 months
Ryan 2012[146]	Twice daily mobile phone symptom, drug, and PEF transmission with immediate feedback *vs.* paper monitoring	• Adults • N=288 • 12+ years (mean 49) • Poorly controlled asthma	• Hospitalisation • ED visits • GP visits • Oral steroid use • AQLQ • ACQ	6 months
Seid 2012[154]	Tailored SMS plus in-person motivational interviewing *vs.* education without tailored SMS	• Adolescents • N=26 • 12-18 years • Mod./severe asthma	• PedsQL	1 and 3 months
Van der Meer 2009[184]	Daily symptom and FEV1 reporting via internet or SMS plus communication with an asthma nurse *vs.* diary only	• Adults • N=200 • 18-50 years • ICS for > 3 months in the past year • No OCS therapy	• AQLQ	12 months
Vollmer 2006[190]	Three phone calls with tailored advice *vs.* routine care	• Adults • N=6948 • 18+ years	• AQLQ	10 months
Willems 2007[194]	Internet daily PEF monitoring with feedback from asthma nurse *vs.* routine care	• Adults and children • N=109 • 7+ years (mean=28)	• AQLQ • ED visits	12 months
Xu 2010[198]	Symptom monitoring and advice in fortnightly follow-up calls from nurse specialist *vs.* routine care	• Children/adolescents • N=121 • 3-16 years • Recent exacerbation	• Hospitalisations • ED visits • Oral steroids • School absence	6 months
Young 2012[200]	Telephone pharmacist consultations *vs.* routine care	• Adults • N=98 • 19+ years	• Withdrawal	Unknown follow-up

Table 107: Adult comparison 1: Tele-health services versus face-to-face equivalents for adults with asthma

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with face-to-face equivalents	Risk difference with Tele-health services (95% CI)
Quality of life Asthma Quality of Life Questionnaire. Scale from: 1 to 7.	960 (3 studies) 12 months	⊕⊕⊕⊖ **MODERATE**[2] due to risk of bias		The mean quality of life in the control groups was **5.35 units**[1]	The mean quality of life in the intervention groups was **0.01 lower** (0.17 lower to 0.14 higher)
UHU hospitalisation	451 (2 studies) 6 months[4]	⊕⊖⊖⊖ **VERY LOW**[2,5,6] due to risk of bias, imprecision	**OR 0.14** (0 to 7.06)[3]	**Moderate** **6 per 1000**	**5 fewer per 1000** (from 6 fewer to 35 more)
UHU ED visit	451 (2 studies) 6 months[4]	⊕⊖⊖⊖ **VERY LOW**[2,5,6] due to risk of bias, imprecision	**OR 7.75** (0.48 to 124.9)[3]	**Moderate** **0 per 1000**	**-**
Exacerbations requiring oral steroids	278 (1 study)	⊕⊖⊖⊖ **VERY LOW**[2,6] due to risk of bias, imprecision	**RR 1.72** (0.42 to 7.04)	**Moderate** **21 per 1000**	**15 more per 1000** (from 12 fewer to 127 more)
Asthma control Asthma Control Questionnaire. Scale from: 0 to 6.	682 (2 studies) 12 months	⊕⊕⊕⊖ **MODERATE**[2] due to risk of bias		The mean asthma control in the control groups was **1.33 units**[7]	The mean asthma control in the intervention groups was **0.11 lower** (0.27 lower to 0.04 higher)
UHU GP visits	451 (2 studies) 6 months[4]	⊕⊕⊖⊖ **LOW**[2,6,8] due to risk of bias, imprecision	**RR 0.86** (0.56 to 1.32)	**Moderate** **132 per 1000**	**18 fewer per 1000** (from 58 fewer to 42 more)
Change in FEV1 (mL)	173 (1 study) 6 months	⊕⊖⊖⊖ **VERY LOW**[2,5,9] due to risk of bias, imprecision		The mean change in fev1 (ml) in the control groups was **35 mL**	The mean change in fev1 (ml) in the intervention groups was **152 higher** (54 to 250 higher)
Withdrawal	672 (3 studies) 6-12 months	⊕⊖⊖⊖ **VERY LOW**[6,10] due to inconsistency, imprecision	**RR 0.78** (0.32 to 1.9)	**Moderate** **120 per 1000**	**26 fewer per 1000** (from 82 fewer to 108 more)

[1] Weighted endpoint mean of the control groups
[2] Studies could not use blinding to control for performance or detection bias
[3] Very rare events - Peto odds ratio used
[4] Pinnock 2003 was a pragmatic trial of variable intervention duration, but did not contribute any events to the analysis
[5] Evidence of sub-optimal randomisation procedures and imputation of missing values, and selective reporting
[6] 95% CI crosses both the MIDs
[7] Endpoint mean in the control group of Pinnock 2007, the larger of the two included trials (Gruffydd-Jones reported mean change)

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects

[8] While there were several issues with one of the studies in the analysis, it only accounted for 6.6% of the analysis weight.
[9] 95% CI crossed an MID
[10] Heterogeneity was high (I^2 = 79%)

Table 108: Adult comparison 2: Tele-monitoring versus paper-based monitoring for adults with asthma

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Paper-based monitoring	Risk difference with Tele-monitoring (95% CI)
Quality of life Asthma Quality of Life Questionnaire. Scale from: 1 to 7.	384 (2 studies) 6-12 months	⊕⊖⊖⊖ **VERY LOW**[1,2,3] due to risk of bias, inconsistency, imprecision		The mean quality of life in the control groups was **5.45 units**	The mean quality of life in the intervention groups was **0.21 higher** (0.09 lower to 0.5 higher)
UHU hospitalisation	386 (3 studies) 4-6 months	⊕⊖⊖⊖ **VERY LOW**[4,5,6] due to risk of bias, inconsistency, imprecision	**RR 0.60** (0.13 to 2.86)	**Moderate**	
				22 per 1000	**9 fewer per 1000** (from 19 fewer to 41 more)
UHU ED visit	370 (2 studies) 6 months	⊕⊖⊖⊖ **VERY LOW**[6,7,8] due to risk of bias, inconsistency, imprecision	**RR 0.89** (0.02 to 33.53)	**Moderate**	
				130 per 1000	**14 fewer per 1000** (from 127 fewer to 1000 more)
Exacerbations requiring oral steroids	281 (1 study) 6 months	⊕⊕⊖⊖ **LOW**[6] due to imprecision	**RR 0.94** (0.59 to 1.49)	**Moderate**	
				213 per 1000	**13 fewer per 1000** (from 87 fewer to 104 more)
Asthma control Asthma Control Questionnaire. Scale from: 0 to 6.	478 (2 studies) 6-12 months	⊕⊖⊖⊖ **VERY LOW**[1,3,9] due to risk of bias, inconsistency, imprecision		The mean asthma control in the control groups was **1.56 units**	The mean asthma control in the intervention groups was **0.24 lower** (0.72 lower to 0.24 higher)
UHU GP visits	281 (1 study) 6 months	⊕⊕⊕⊖ **MODERATE**[3] due to imprecision	**RR 1.25** (0.89 to 1.76)	**Moderate**	
				291 per 1000	**73 more per 1000** (from 32 fewer to 221 more)
Change in FEV1 (mL)	200 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[3,10] due to risk of bias, imprecision		The mean change in fev1 (ml) in the control groups was **-10 mL**	The mean change in fev1 (ml) in the intervention groups was **250 higher** (33.36 to 466.64 higher)
PEF (L/min)	89 (1 study) 6 months	⊕⊕⊖⊖ **LOW**[3,7] due to risk of bias, imprecision		The mean pef (l/min) in the control groups was **343.5 Litres per minute**	The mean pef (l/min) in the intervention groups was **39.2 higher** (16.58 to 61.82 higher)
Withdrawal	624 (4 studies) 4-12 months	⊕⊕⊖⊖ **LOW**[6] due to imprecision	**RR 1.01** (0.73 to 1.39)	**Moderate**	
				152 per 1000	**2 more per 1000** (from 41 fewer to 59 more)

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects

[1] One study analysed complete cases and did not blind participants, investigators or outcome assessors, which carried the majority of the analysis weight.
[2] Heterogeneity was high (I^2 = 53%)
[3] 95% CI crosses one of the MIDs
[4] Only one study used any blinding procedures (outcome assessors), and there were uncertainties regarding allocation concealment
[5] Heterogeneity was not statistically significant (I^2= 42%), but point estimates are very different
[6] 95% CIs cross both MIDs
[7] Study carrying the most weight did not blind outcome assessors (and could not blind participants and investigators), and dropout was high in both groups
[8] Heterogeneity was high (I^2= 80%)
[9] Heterogeneity was very high (I squared = 91%)
[10] No blinding of outcome assessors (and unable to blind participants and investigators). Only complete cases were analysed.

Table 109: Adult comparison 3: Tele-health packages versus nothing (usual care) for adults with asthma

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Nothing (usual care)	Risk difference with Tele-health packages (95% CI)
Quality of life Asthma Quality of Life Questionnaire. Scale from: 1 to 7.	1633 (3 studies) 10-12 months	⊕⊕⊕⊖ **MODERATE**[1] due to risk of bias		The mean quality of life in the control groups was **5.18 units**	The mean quality of life in the intervention groups was **0.08 higher** (0.03 lower to 0.20 higher)
UHU hospitalisation	404 (4 studies) 6-12 months	⊕⊕⊕⊖ **MODERATE**[1,3] due to risk of bias	**OR 0.16** (0.05 to 0.56)[2]	**Moderate** **56 per 1000**	**47 fewer per 1000** (from 24 fewer to 53 fewer)
UHU ED visit	415 (4 studies) 6-12 months	⊕⊖⊖⊖ **VERY LOW**[1,4,5] due to risk of bias, imprecision	**RR 0.82** (0.38 to 1.8)	**Moderate** **65 per 1000**	**12 fewer per 1000** (from 40 fewer to 52 more)
Exacerbations requiring oral steroids	60 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,5] due to risk of bias, imprecision	**RR 0.94** (0.67 to 1.3)	**Moderate** **724 per 1000**	**43 fewer per 1000** (from 239 fewer to 217 more)
Asthma control Asthma Control Questionnaire. Scale from: 0 to 6.	556 (1 study) 12 months	⊕⊕⊕⊕ **HIGH**		The mean asthma control in the control groups was **1.24 units**	The mean asthma control in the intervention groups was **0.04 lower** (0.2 lower to 0.12 higher)
UHU GP visits	295 (3 studies) 6-12 months	⊕⊖⊖⊖ **VERY LOW**[1,5,6,7] due to risk of bias, inconsistency, imprecision	**RR 0.96** (0.39 to 2.37)	**Moderate** **389 per 1000**	**16 fewer per 1000** (from 237 fewer to 533 more)
Change in FEV1 (mL)	165 (1 study) 6 months	⊕⊕⊖⊖ **LOW**[1,8] due to risk of bias, imprecision		The mean change in fev1 (ml) in the control groups was **4 mL**	The mean change in fev1 (ml) in the intervention groups was **183 higher** (85 to 281 higher)

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects	
Symptom days per month Scale from: 0 to 30.	608 (1 study)	⊕⊕⊕⊖ **MODERATE**[8] due to imprecision		The mean symptom days per month in the control groups was **7.3 days**	The mean symptom days per month in the intervention groups was **0.6 higher** (0.82 lower to 2.02 higher)
Symptom nights per month Scale from: 0 to 30.	608 (1 study)	⊕⊕⊕⊖ **MODERATE**[8] due to imprecision		The mean symptom nights per month in the control groups was **3.8 nights**	The mean symptom nights per month in the intervention groups was **0.1 lower** (1.21 lower to 1.01 higher)
Withdrawal	512 (5 studies) 6-12 months	⊕⊖⊖⊖ **VERY LOW**[1,5] due to risk of bias, imprecision	**RR 0.81** (0.51 to 1.29)	**Moderate** **111 per 1000**	**21 fewer per 1000** (from 54 fewer to 32 more)

[1] Issues across studies with blinding, completeness of outcome data, and allocation concealment
[2] Very rare events - Peto odds ratio used
[3] Confidence intervals were wide but did not cross an MID
[4] Point estimates were varied but statistical heterogeneity was low - random effects sensitivity did not change imprecision
[5] 95% CI crossed both MIDs
[6] Heterogeneity was high (I^2 = 66%)
[7] One study was only recruited older adults (53% of analysis weight)
[8] 95% CIs crossed an MID

Table 110: Child comparison 1: Tele-health services versus face-to-face equivalents for children with asthma

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with face-to-face equivalents	Risk difference with Tele-health services (95% CI)
Quality of life - child Paediatric Asthma Quality of Life Questionnaire. Scale from: 1 to 7.	120 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[1,2] due to risk of bias, imprecision		The mean quality of life - child in the control groups was **5.8 units**	The mean quality of life - child in the intervention groups was **0.3 higher** (0.11 lower to 0.71 higher)
Quality of life - caregiver Paediatric Asthma Quality of Life Questionnaire. Scale from: 1 to 7.	120 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[1,2] due to risk of bias, imprecision		The mean quality of life - caregiver in the control groups was **6.2 units**	The mean quality of life - caregiver in the intervention groups was **0.2 higher** (0.12 lower to 0.52 higher)
UHU hospitalisation	120 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,3] due to risk of bias, imprecision	**RR 1** (0.06 to 15.62)	**Moderate** **17 per 1000**	**0 fewer per 1000** (from 16 fewer to 249 more)
UHU ED visit	120 (1 study) 12 months	⊕⊖⊖⊖ **VERY LOW**[1,3] due to risk of bias, imprecision	**RR 2** (0.38 to 10.51)	**Moderate** **33 per 1000**	**33 more per 1000** (from 20 fewer to 314 more)

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects	
FEV1 % predicted	120 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[1,2] due to risk of bias, imprecision		The mean fev1 % predicted in the control groups was **92.2 %**	The mean fev1 % predicted in the intervention groups was **5.2 higher** (1.48 lower to 11.88 higher)

[1] No blinding and unbalanced attrition
[2] 95% CI crosses an MID
[3] 95% CI crosses both MIDs

Table 111: Child comparison 2: Tele-monitoring versus paper-based monitoring for children with asthma

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Paper-based monitoring	Risk difference with Tele-monitoring (95% CI)
Change in morning PEF (L/min)	153 (1 study) 3 months	⊕⊕⊖⊖ **LOW**[1,2] due to risk of bias, imprecision		The mean change in morning pef (l/min) in the control groups was **10.9 Litres per minute**	The mean change in morning pef (l/min) in the intervention groups was **7.80 higher** (6.37 lower to 21.97 higher)
Change in evening PEF (L/min)	153 (1 study) 3 months	⊕⊕⊖⊖ **LOW**[1,2] due to risk of bias, imprecision		The mean change in evening pef (l/min) in the control groups was **11.1 Litres per minute**	The mean change in evening pef (l/min) in the intervention groups was **12 higher** (3.59 lower to 27.59 higher)
Withdrawal	164 (1 study) 3 months	⊕⊖⊖⊖ **VERY LOW**[1,3] due to risk of bias, imprecision	**RR 1.04** (0.33 to 3.26)	**Moderate** **66 per 1000**	**3 more per 1000** (from 44 fewer to 149 more)

[1] Participants and investigators could not be blind (outcome assessors were blinded)
[2] 95% CI crosses an MID
[3] 95% CI crosses both MIDs

Table 112: Child comparison 3: Tele-health packages versus nothing (usual care) for children with asthma

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Nothing (usual care)	Risk difference with Tele-health packages (95% CI)
Change in quality of life - child Paediatric Asthma Quality of Life Questionnaire. Scale from: 1 to 7.	82 (1 study) 6 months	⊕⊕⊖⊖ **LOW**[1,3] due to risk of bias, imprecision		The mean change in quality of life - child in the control groups was **0.5 units**	The mean change in quality of life - child in the intervention groups was **0.70 higher** (0.29 to 1.11 higher)
Quality of life - caregiver Paediatric Asthma Quality of Life Questionnaire. Scale from: 1 to 7.	181 (2 studies) 6-12 months	⊕⊕⊕⊖ **MODERATE**[1,2] due to risk of bias		The mean quality of life - caregiver in the control groups was **6.31 units**[4]	The mean quality of life - caregiver in the intervention groups was **0.18 higher** (0.10 lower to 0.46 higher)

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects	
UHU hospitalisation	609 (5 studies) 3-12 months	⊕⊖⊖⊖ **VERY LOW**[5,6,7] due to risk of bias, imprecision	**RR 1.43** (0.59 to 3.46)	**Moderate**	
				20 per 1000	**9 more per 1000** (from 8 fewer to 49 more)
UHU ED visit	566 (4 studies) 3-12 months	⊕⊖⊖⊖ **VERY LOW**[5,6,7] due to risk of bias, imprecision	**RR 1** (0.56 to 1.8)	**Moderate**	
				92 per 1000	**0 fewer per 1000** (from 40 fewer to 74 more)
Exacerbations requiring oral steroids	125 (2 studies) 6-12 months	⊕⊖⊖⊖ **VERY LOW**[5,7] due to risk of bias, imprecision	**RR 1.01** (0.8 to 1.27)	**Moderate**	
				719 per 1000	**7 more per 1000** (from 144 fewer to 194 more)
Asthma control Asthma Control Questionnaire. Scale from: 0 to 6.	301 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[1,3] due to risk of bias, imprecision		The mean asthma control in the control groups was **2.21 units**[8]	The mean asthma control in the intervention groups was **0.31 lower** (0.56 to 0.06 lower)
UHU GP visits	99 (1 study) 8 months	⊕⊕⊖⊖ **LOW**[7] due to imprecision	**OR 0.80** (0.30 to 2.13)	**Moderate**	
				157 per 1000	**31 fewer per 1000** (from 110 fewer to 177 more)
Withdrawal	823 (5 studies) 3-12 months	⊕⊖⊖⊖ **VERY LOW**[5,7,9] due to risk of bias, inconsistency, imprecision	**RR 0.86** (0.53 to 1.41)	**Moderate**	
				161 per 1000	**23 fewer per 1000** (from 76 fewer to 66 more)

[1] One or more study did not blind outcome assessors
[2] MID is close to, but does not cross the 0.5 MID
[3] 95% CI crosses one MID
[4] Control score in Halterman 2012. Xu 2010reported change.
[5] Issues across studies with blinding, completeness of outcome data, and allocation concealment
[6] Point estimates were varied but statistical heterogeneity was low - random effects sensitivity did not change imprecision
[7] 95% CI crosses both MIDs
[8] Control group end score obtained from baseline mean (2.32) minus the reported improvement (0.11)
[9] Some inconsistency (I^2= 38%), random effects used

30.3.2 Tele-healthcare with no involvement from a healthcare provider

Three studies were included in the review[17,34,198] these are summarised in Table 113 below. See also the study selection flow chart in Appendix D, study evidence tables in Appendix G, forest plots in Appendix J, GRADE tables in Appendix I and excluded studies list in Appendix K.

In adults age >16 years, evidence comparing tele-healthcare without healthcare professional involvement vs. usual care was available from one study[17] summarised in the clinical evidence summary (Table 114). This study used interactive voice response telephone calls and reported outcomes on QOL and asthma control.

In children age 5-16 years, evidence comparing tele-healthcare without healthcare professional involvement vs usual care was available from two studies[34,198] summarised in the clinical evidence summary (Table 115). One study[34] used a web-based intervention to provide feedback to parents on child's asthma (recommendations regarding controller use and other aspects of asthma care) and reported the outcome use of controller medications. The other study[198] used interaction voice response calls and reported outcomes on exacerbation, UHU, school days lost, parent work days lost and QOL (child and carer). Both studies reported outcomes at ≥6 months.

Table 113: Summary of studies included in the review

Study	Intervention	Comparison	Population	Outcomes
BENDER 2010[17]	2 automated interactive voice response telephone calls separated by one month, with one additional call if they reported recent symptoms of poorly controlled disease or failure to fill a prescription. Calls included content designed to inquire about asthma symptoms, deliver core educational messages, encourage refilling of ICS prescriptions, and increase communication with providers.	No telephone calls	ADULTS 18 to 65 years; physician-diagnosed asthma for which they were prescribed daily ICS treatment	• QOL • Asthma control questionnaire
CHRISTAKIS2012[34]	Web-based intervention: gathers information from parents (day and night time symptoms, quick-reliever use), applies algorithm to determine asthma severity, home care practices (controller use and adherence), functional status, parental beliefs (outcomes expectation and self-efficacy), feedback on child's asthma (recommendations regarding controller use and other aspects of asthma care).	Control parents had similar intervention around reducing media usage among their children.	Parents of children aged 2 to 10 years with asthma (at least 1 clinical encounter – clinic visit, emergency room or inpatient admission – or two prescription refills for bronchodilators in the last year).	• Controller medication use
XU2010[198]	Intervention 1: Interactive Voice Response Intervention 2: Specialist nurse support (see section 30.3.1)	Usual care	CHILDREN 3 to 16 years; asthma; admission to hospital in previous 12 months or presented at least once to emergency department or GP or specialist with acute asthma requiring oral steroid rescue in previous 12 months.	• Exacerbation • UHU • School days lost • Parent work days lost • QOL (child and carer)

Table 114: Adult comparison 4: Telehealthcare without healthcare professional involvement vs usual care

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with no calls	Risk difference with Interactive voice response telephone calls (95% CI)
QOL <6 months AQLQ. Scale from: 0 to 7.	50 (1 study) 10 weeks	⊕⊕⊖⊖ **LOW**[1,2] due to risk of bias, imprecision		The mean qol <6 months in the control groups was **-0.38**	The mean qol <6 months in the intervention groups was **0.23 higher** (0.32 lower to 0.78 higher)
Asthma Control Questionnaire <6 months ACT. Scale from: 5 to 25.	50 (1 study) 10 weeks	⊕⊖⊖⊖ **VERY LOW**[1,3] due to risk of bias, imprecision		The mean asthma control questionnaire <6 months in the control groups was **-1.84**	The mean asthma control questionnaire <6 months in the intervention groups was **0.72 higher** (1.51 lower to 2.95 higher)

[1] Method of randomisation and allocation concealment unclear
[2] Crosses one MID
[3] Crosses two MIDs

Table 115: Child comparison 4: Telehealthcare without healthcare professional involvement vs usual care

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with No calls	Risk difference with Telephone calls (95% CI)
Exacerbations ≥6 months Self report OCS (assumed to be for exacerbation)	79 (1 study) 6 months	⊕⊖⊖⊖ **VERY LOW**[1,2,3,4] due to risk of bias, imprecision	**RR 0.78** (0.48 to 1.26)	**Moderate** **525 per 1000**	**116 fewer per 1000** (from 273 fewer to 136 more)
QOL ≥6 months Pediatric Asthma Quality of Life Questionnaire (carer). Scale from: 0 to 7.	80 (1 study) 6 months	⊕⊖⊖⊖ **VERY LOW**[1,2,3,4] due to risk of bias, imprecision		The mean qol ≥6 months in the control groups was **1.0**	The mean qol ≥6 months in the intervention groups was **0.2 higher** (0.48 lower to 0.88 higher)
QOL ≥6 months Pediatric Asthma Quality of Life Questionnaire (child). Scale from: 0 to 7.	80 (1 study) 6 months	⊕⊕⊖⊖ **LOW**[1,2,3] due to risk of bias		The mean qol ≥6 months in the control groups was **0.5**	The mean qol ≥6 months in the intervention groups was **0.6 higher** (0.16 to 1.04 higher)
UHU ≥6 months ED visit self report	79 (1 study) 6 months	⊕⊖⊖⊖ **VERY LOW**[1,2,3,5] due to risk of bias, imprecision	**RR 1.23** (0.41 to 3.7)	**Moderate** **125 per 1000**	**29 more per 1000** (from 74 fewer to 338 more)
UHU hospitalisation ≥6 months Hospital admission self report	79 (1 study) 6 months	⊕⊖⊖⊖ **VERY LOW**[1,2,3,5] due to risk of bias,	**RR 1.03** (0.28 to 3.82)	**Moderate** **100 per 1000**	**3 more per 1000** (from 72 fewer to 282 more)

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects	
		imprecision			
School days lost ≥6 months Self report (yes/no to any time off school)	77 (1 study) 6 months	⊕⊖⊖⊖ **VERY LOW**[1,2,3,5] due to risk of bias, imprecision	**RR 0.93** (0.62 to 1.4)	**Moderate** **564 per 1000**	**39 fewer per 1000** (from 214 fewer to 226 more)
Parents' work days lost ≥6 months Self report (yes/no to any work days lost)	78 (1 study) 6 months	⊕⊖⊖⊖ **VERY LOW**[1,2,3,5] due to risk of bias, imprecision	**RR 1** (0.53 to 1.87)	**Moderate** **333 per 1000**	**0 fewer per 1000** (from 157 fewer to 290 more)
Controller medication use in patients who should have been on controller medications at baseline ≥6 months i.e. persistent asthma	49 (1 study) 12 months	⊕⊕⊕⊖ **MODERATE**[4] due to imprecision	**RR 2.21** (0.82 to 5.97)	**Moderate** **167 per 1000**	**202 more per 1000** (from 30 fewer to 830 more)
Persistent asthma on controllers at baseline but discontinued at 6 months	100 (1 study) 12 months	⊕⊕⊖⊖ **LOW**[5] due to imprecision	**RR 2.76** (0.73 to 10.42)	**Moderate** **52 per 1000**	**92 more per 1000** (from 14 fewer to 490 more)
Of those who met severity criteria for controllers at baseline, number on them at 12 months	135 (1 study) 12 months	⊕⊕⊕⊖ **MODERATE**[4] due to imprecision	**RR 1.05** (0.81 to 1.37)	**Moderate** **610 per 1000**	**30 more per 1000** (from 116 fewer to 226 more)

[1] Method of randomisation and allocation concealment unclear
[2] Groups not comparable at baseline
[3] Underpowered
[4] Crosses one MID
[5] Crosses two MIDs

30.4 Economic evidence

Published literature

Three economic evaluations were identified with the relevant comparison and have been included in this review.[61,146,195] These are summarised in the economic evidence profiles below (Table 116, Table 117 and Table 118) and the economic evidence tables in Appendix H.

Two economic evaluations relating to this review question were identified but were excluded due to a combination of limited applicability and methodological limitations.[132,134] These are listed in Appendix L, with reasons for exclusion given.

See also the economic article selection flow diagram in Appendix E.

Table 116: Economic evidence profile: tele-healthcare consultations versus face-to-face reviews

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost-effectiveness	Uncertainty
Gruffydd-Jones 2005[61] (UK)	Partially applicable[(a)]	Potentially serious limitations[(b)]	The economic evaluation is a within trial analysis of a 12 month RCT. In the intervention a 6 month face to face asthma review was replaced with telephone calls which screened individuals who would need to come in for further assessment. An individualised asthma plan was also formulated and relayed to the patient. This advised them on what to do if their asthma worsened. In the control arm patients received the usual care option of a face to face to review after 6 months.	Bootstrapped cost difference: −£122.35	Higher health outcomes[(c)] were reported in the intervention arm; however these were not clinically significant.	No formal cost-effectiveness was evaluated however it would appear that t tele-healthcare is the dominant strategy (higher or same health with lower costs)	No uncertainty analysis was conducted

(a) Quality of life not assessed using QALYs which may produce different health outcomes.
(b) Short time horizon of 12 months may not reflect long term health and cost outcomes, adverse events are unlikely to be captured in this time, especially with the small cohort of patients monitored. Only quality, not quantity, of life was assessed and 2 individuals died or their condition was exacerbated in the intervention arm. No discussion was made regarding whether this was due to non-asthma causes or sampling error and whether the intervention was responsible.
(c) Health outcomes measured using the ACQ questionnaire

Table 117: Economic evidence profile: tele-healthcare monitoring versus paper-based self-monitoring

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost-effectiveness	Uncertainty
Ryan 2012[146] (UK)	Partially applicable[(a)]	Potentially serious limitations[(a)]	Economic evaluation based on a one year multicentre RCT conducted in the UK. In the intervention arm tele-healthcare	£70	There was no significant change in	No formal cost-effectiveness evaluated	No sensitivity analysis was conducted

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost-effectiveness	Uncertainty
			was implemented through mobile phone monitoring whereby the patient would enter data (symptoms, drug use and peak flow readings) twice daily using a mobile phone application. The patient receives immediate feedback which prompts action based on a pre-arranged plan. Patients in the control arm were asked to collect the same data on paper.		asthma control or self-efficacy between the 2 interven-tions.	however given there were no changes in health outcomes but a positive cost the tele-healthcare intervention can be seen as dominated.	

(a) Health was not measured using QALYs

(b) Short time horizon of 12 months may not capture long term health and cost outcomes. Also no sensitivity analysis was conducted meaning the results may not be robust to slight changes in certain outcomes intervention costs.

Table 118: Economic evidence profile: tele-healthcare packages versus nothing (usual care)

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost-effectiveness	Uncertainty
Willems 2007[195] (Netherlands)	Partially applicable[(a)]	Potentially serious limitations[(a)]	The economic evaluation based on a 12 month RCT conducted in the Netherlands. In the intervention arm tele-healthcare was implemented using a monitoring device which digitally transferred data to a nurse practitioner who would analyse the data and change asthma medication accordingly.	£421 for children[(b)] £353 for adults[(c)]	0.01 QALYs for children 0.03 for adults	£40,865 per QALY gained for children £10,693 per QALY gained for adults	One sensitivity analysis was conducted by removing monitor costs from the intervention arm, based on the assumption that the price of these devices will fall drastically in the future. This reduced the ICER for adults to £1224 per QALY and for children £10,502 per QALY, therefore making the intervention cost-effective for children by bringing the ICER below the £20,000 threshold.

(a) Study undertaken in the Netherlands therefore costs will be less generalizable to a UK setting.

(b) Short time horizon of 12 months may not capture long term health and cost outcomes.

(c) Children defined as individuals aged 7 to 18 years old

(d) Adults defined as individuals over 18 years old

30.5 Evidence statements

Clinical

ADULTS: Tele-health services versus face-to-face equivalents for adults with asthma

- No evidence was identified for mortality.
- There was a borderline clinically important difference for asthma exacerbations requiring OCS (1 study, N=278, very low quality), UHU hospitalisations (2 studies, N=451, very low quality) and UHU GP visits (2 studies, N=451, low quality).
- There was no clinically important difference forQOL (3 studies, N=960, moderate quality), asthma control questionnaire score (2 studies, N=682, moderate quality) and lung function measured using FEV1 (1 study, N=173, very low quality).

ADULTS: Tele-monitoring versus paper-based monitoring for adults with asthma

- No evidence was identified for mortality.
- There was a borderline clinically important difference for asthma exacerbations requiring OCS (1 study, N=281, low quality), UHU hospitalisations (3 studies, N=386, very low quality) and UHU ED visits (2 studies, N=370, very low quality).
- There was no clinically important difference for QOL (2 studies, N=384, very low quality) and asthma control questionnaire score (2 studies, N=478, very low quality).
- There was a clinically important harm for UHU GP visits (1 study, N=281, moderate quality).
- There was a clinically important benefit for lung function measured using FEV1 or PEF (low quality)

ADULTS: Tele-health packages versus nothing (usual care) for adults with asthma

- No evidence was identified for mortality.
- There was a clinically important benefit for asthma exacerbations requiring OCS (1 study, N=60, very low quality) and UHU hospitalisations (4 studies, N=404, moderate quality).
- There was a borderline clinically important benefit for UHU ED visits (4 studies, N=415, very low quality), UHU GP visits (3 studies, N=295, very low quality) and lung function measured using FEV1 (1 study, N=165, low quality)
- There was no clinically important difference for QOL (3 studies, N=1633, moderate quality), asthma control questionnaire score (1 study, N=556, high quality) and symptom free days and nights (1 study, N=608, moderate quality).

ADULTS: Telehealthcare without healthcare professional involvement vs usual care

- No evidence was identified for mortality, asthma exacerbations or UHU
- There was no clinically important difference for QOL (1 study, N=50, low quality) and asthma control questionnaire score (1 study, N=50, very low quality).

CHILDREN: Tele-health services versus face-to-face equivalents for children with asthma

- No evidence was identified for mortality, asthma exacerbations or asthma control questionnaires.
- There was no clinically important difference for QOL carer and QOL child (1 study, N=120, low quality) and UHU hospitalisation (1 study, N=120, very low quality)
- There was a borderline clinically important difference for ED visits (1 study, N=120, very low quality) and lung function measured using FEV1 (1 study, N=120, low quality).

CHILDREN: Tele-monitoring versus paper-based monitoring for children with asthma

- No evidence was identified for any of the 5 priority outcomes.
- There was no clinically important difference for lung function measured using PEF (1 study, N=153, very low quality).

CHILDREN: Tele-health packages versus nothing (usual care) for children with asthma

- No evidence was identified for mortality.
- There was a clinically important benefit for QOL child (1 study, N=82, low quality).
- There was no clinically important difference for QOL parent (2 studies, N=181, moderate quality), UHU hospitalisations (5 studies, N=609, very low quality), UHU ED visits (4 studies, N=566, very low quality), asthma exacerbations requiring OCS (2 studies, N=125, very low quality) and Asthma control questionnaire score (1 study, N=301, low quality).
- There was a borderline clinically important difference for GP visits (1 study, N=99, low quality).

CHILDREN: Telehealthcare without healthcare professional involvement vs usual care

- No evidence was identified for mortality, asthma control questionnaires
- There was a clinically important benefit for asthma exacerbations requiring OCS (1 study, N=79, very low quality) and QOL child (1 study, N=80, low quality).
- There was no clinically important difference for QOL parent (1 study, N=80, very low quality), UHU (hospitalisation) (1 study, N=79, very low quality) and parent work days lost (1 study, N=78, very low quality).
- There was a borderline clinically important difference for UHU ED visits (1 study, N=79, very low quality) and child school days lost (1 study, N=77, very low quality).
- Evidence suggested that there was a clinical benefit in the number of people taking ICS medication at 12 months, who should have been on medication at baseline.

Economic

- One within trial cost analysis found that tele-healthcare consultations was dominant (produced lower costs and non significant increases in health outcomes) when compared to face-to-face reviews. This analysis was assessed as partially applicable with potentially serious limitations.
- One within trial analysis found that tele-healthcare monitoring was dominated (produced higher costs and no increase in health outcomes) when compared to paper based monitoring. This analysis was assessed as partially applicable with potentially serious limitations.
- One within trial analysis found that tele-healthcare was cost-effective in adults when compared to nothing (ICER: £10,693). However the same analysis found that tele-healthcare versus nothing was not cost-effective in children (ICER: £40,865).

30.6 Recommendations and link to evidence

Recommendations	**No clinical recommendation.**
Research recommendations	**6. What is the long-term (more than 12 months) clinical and cost effectiveness of using tele-healthcare as a means to monitor asthma control in children, young people and adults? Modalities of tele-healthcare can include telephone interview (healthcare professional involvement) and internet or smartphone-based monitoring support (no healthcare professional involvement).**
Relative values of different outcomes	Due to large heterogeneity between telehealthcare interventions for monitoring asthma in the literature, studies were analysed in four separate groups. These comparisons were not pre-specified in the review protocol, but were necessary in

	order to make the analyses clinically meaningful and were constructed prior to extraction and pooling of outcome data. The GC considered the following outcomes as critical for this review: mortality, quality of life, exacerbations, asthma control measured by questionnaires and unscheduled healthcare utilisation. Asthma attacks can lead to both unscheduled healthcare utilisation (emergency department (ED) visits, out-of-hours GP surgery visits or hospitalisation) and the need for a course of oral corticosteroids (OCS). Therefore, the GC considered exacerbations requiring OCS treatment and exacerbations requiring unscheduled healthcare utilisation as separate outcomes. The GC also considered the following important outcomes: lung function (FEV1), symptoms (symptom-free days), dose of regular asthma therapy (ICS dose), rescue medication (SABA dose) and time off school or work. These outcomes were considered as important measures of asthma control for the patient. The GC acknowledged that regular asthma preventer therapy, inhaled corticosteroid (ICS), is often underprescribed / used, and rescue medications (for example short acting beta agonists (SABA)) may be overprescribed / used (for further reference see NRAD[144]). Therefore, the GC reviewed the evidence on these outcomes.

Trade-off between clinical benefits and harms	Four comparisons were considered for both adults and children:
	1) Tele-health services versus face-to-face equivalents;
	2) Tele-monitoring versus paper-based monitoring;
	3) Tele-health packages versus standard or usual care and
	4) Telehealthcare without healthcare professional involvement vs standard or usual care.
	
	<u>ADULTS</u>
	<u>The GC noted the following results:</u>
	For tele-health services versus face-to-face equivalents there was
	• borderline clinically important differences for:
	  ○ asthma exacerbations requiring OCS
	  ○ hospitalisation
	  ○ GP visits
	• no clinically important differences for:
	  ○ QOL
	  ○ asthma control questionnaire score
	  ○ lung function (FEV1)
	
	For tele-monitoring versus paper-based monitoring there was:
	• borderline clinically important differences for:
	  ○ asthma exacerbations requiring OCS
	  ○ hospitalisation
	  ○ ED visits
	• no clinically important differences for:
	  ○ QOL
	  ○ asthma control questionnaire score
	• clinically important benefits for:
	  ○ lung function measured using FEV1 or PEF
	• a clinically important harm for:
	  ○ GP visits

For tele-health packages versus usual care there was:

- clinically important benefits for:
  - asthma exacerbations requiring OCS
  - hospitalisation
- borderline clinically important benefits for:
  - ED visits
  - GP visits
  - Lung function measured using FEV1
- no clinically important differences for:
  - QOL
  - asthma control questionnaire score
  - symptom free days and nights

For telehealthcare without healthcare professional involvement vs usual care there was:

- no clinically important difference for:
  - asthma control questionnaire score
  - QOL

In adults, only telehealthcare packages (compared to standard care) appeared to improve outcomes. However, this intervention cannot be considered as assessing the effect of telehealthcare alone, as people received additional care (telehealthcare was not the only difference between the groups).

The GC acknowledged that monitoring asthma is essential and the consensus of GC opinion was that this should theoretically result in better outcomes. Without further research the additional benefit of telehealthcare is currently unclear.

The GC also noted that the secondary outcomes of importance, FEV1 (in the first three comparisons quoted above) showed a benefit in adults; however, this benefit was not reflected in other outcomes, such as the QOL.

For tele-health services versus face-to-face equivalents the GC noted that there was no difference between the groups. The GC considered this an important finding given that this comparison potentially replaces face-to-face visits with a telehealthcare interaction. The GC also noted the heterogeneity between interventions in this comparison, with some studies involving email and others involving telephone calls. An improvement might be expected using telephone calls but not email, as a telephone intervention could be more interactive and reach people who are already considered to be poor compliers. However, evidence was not available from enough studies to subgroup on the basis of 'type of device' as pre-specified in the protocol.

<u>CHILDREN AND YOUNG PEOPLE</u>

For tele-health services versus face-to-face equivalents there was:

- no clinically important differences for:
  - QOL carer and QOL child
  - hospitalisation
- borderline clinically important benefits for:
  - ED visits
  - lung function measured using FEV1

	For tele-monitoring versus paper-based monitoring no evidence was identified for any of the 5 priority outcomes. There was:
	• no clinically important differences for:
	  ○ lung function measured using PEF
	
	For tele-health packages versus usual care there was:
	• clinically important benefit for:
	• QOL childno clinically important differences for:
	  ○ QOL parent
	  ○ hospitalisation
	  ○ asthma exacerbations requiring OCS
	  ○ asthma control questionnaire score
	  ○ ED visits
	• borderline clinically important benefits for:
	  ○ GP visits
	
	For telehealthcare without healthcare professional involvement vs usual care there was:
	• clinically important benefits for:
	  ○ asthma exacerbations requiring OCS
	  ○ QOL child
	• no clinically important differences for:
	  ○ QOL parent
	  ○ hospitalisation
	  ○ parent work days lost
	• borderline clinically important difference for:
	  ○ ED visits
	  ○ child school days lost
	
	<u>CHILDREN < 5 YEARS</u>
	In children <5 years, no evidence was identified.
	
	Overall, the GC concluded that there was a paucity of telehealthcare evidence for the outcomes of interest to the GC (with no consistency of findings) and the interventions were of limited quality and too heterogeneous to make a recommendation in adults, young people or children for the use of telehealthcare in monitoring asthma. In light of this, the GC agreed to make a future research recommendation.

Economic considerations	Three health economic papers were presented to the GC. All three papers were within-trial analyses that appeared in the clinical review. A study by Gruffydd-Jones 2005 looked at tele-healthcare (THC) which replaced face-to-face reviews with a telephone review and found THC to be cheaper and no less effective. However, the GC noted the considerable heterogeneity that existed over the clinical evidence. Therefore, if the same cost analysis was conducted in another study the results could be very different. The GC also noted the short 12-month time horizon as a serious limitation to the study's results. Therefore, due to considerable uncertainty, the GC did not feel they could make a recommendation concerning the replacement of face-to-face reviews with THC.

	A study by Ryan 2012 compared THC monitoring with paper-based monitoring. The study found that THC monitoring resulted in higher costs for the NHS but no significant differences in patient's quality of life. However, again due to considerable heterogeneity in the clinical review and noting the study's short time horizon the GC did not feel that enough evidence was available to form a recommendation regarding the replacement of paper-based monitoring with THC. Finally, a study by Willems 2007 looked at comparing a THC monitoring intervention to usual care. A formal cost-effectiveness analysis found that THC was cost-effective compared to usual care with an ICER of £10,693 per QALY for adults, but was not cost-effective for children with an ICER of £40,865. The main concern was the study's short time horizon of 12 months; the GC considered that during this period the benefits of the intervention would be highest and, as time went on, the benefits would decrease. The GC considered the economic evidence not enough to make a recommendation on this intervention. The GC also noted that the intervention was very specific and that not enough clinical evidence existed to back up the results to make such a specific recommendation.
Quality of evidence	In adults, the majority of the evidence was of low and very low quality for the critical outcomes. In children (with the exception of the carer QOL outcome in comparison 3 at moderate quality), evidence for all the critical outcomes was of low and very low quality. The evidence was downgraded due to risk of bias, imprecision and inconsistency.
Other considerations	The GC concluded that there was too little evidence and too much heterogeneity between interventions to support or refute the use of telehealthcare for monitoring asthma. There was also heterogeneity in the severity of asthma within the study populations. Severe asthma is excluded from the scope of this guideline. Five studies were noted to potentially include people with severe asthma; however, the GC did not exclude these studies because of lack of clarity (it was unclear whether the population within these studies met ERS/ATS guideline operational definition for severe asthma). The GC agreed that future research in this area is needed, in particular in order to identify the modality of telehealthcare that will be most beneficial. There are some positive indications for telehealthcare, with little evidence of causing clinical harm. However, there was not enough evidence to show a clear clinical benefit and it is important that NHS financial resources are not invested in implementing telehealthcare for monitoring until further evidence is available. Therefore, the GC recommended the use of telehealthcare in a research setting only and made a high-priority research recommendation to investigate the long-term (more than 12 months) clinical and cost-effectiveness of using tele-healthcare as a means to monitor asthma control. Further details on the high-priority research recommendation made can be found in appendix N. The GC also discussed whether a 12-month time horizon may be too short to answer this question and that studies with a longer follow-up than 12 months are essential to observe any long-term benefits or harms of telehealthcare for monitoring asthma. Following completion of the feasibility project in October 2016, results of the updated searches in 2017 are available in appendix R. Two studies were identified which the GC agreed did not suggest a change in the recommendations was warranted.

31 Reference list

1 Aaron SD, Vandemheen KL, Boulet LP, McIvor RA, Fitzgerald JM, Hernandez P et al. Overdiagnosis of asthma in obese and nonobese adults. Canadian Medical Association Journal. 2008; 179(11):1121-1131

2 Abraham CM, Ownby DR, Peterson EL, Wegienka G, Zoratti EM, Williams LK et al. The relationship between seroatopy and symptoms of either allergic rhinitis or asthma. Journal of Allergy and Clinical Immunology. 2007; 119(5):1099-1104

3 Adams RJ, Boath K, Homan S, Campbell DA, Ruffin RE. A randomized trial of peak-flow and symptom-based action plans in adults with moderate-to-severe asthma. Respirology. 2001; 6(4):297-304

4 Al-Showair RAM, Pearson SB, Chrystyn H. The potential of a 2Tone Trainer to help patients use their metered-dose inhalers. Chest. 2007; 131(6):1776-1782

5 Ammari WG, Chrystyn H. Optimizing the inhalation flow and technique through metered dose inhalers of asthmatic adults and children attending a community pharmacy. Journal of Asthma. 2013; 50(5):505-513

6 Anderson SD. Exercise-induced bronchoconstriction in the 21st century. Journal of the American Osteopathic Association. 2011; 111(11 Suppl 7):S3-10

7 Anderson SD, Charlton B, Weiler JM, Nichols S, Spector SL, Pearlman DS et al. Comparison of mannitol and methacholine to predict exercise-induced bronchoconstriction and a clinical diagnosis of asthma. Respiratory Research. 2009; 10:4

8 Asthma UK. Asthma facts and FAQs. 2015. Available from: http://www.asthma.org.uk/asthma-facts-and-statistics [Last accessed: 17 June 2015]

9 Avital A, Godfrey S, Springer C. Exercise, methacholine, and adenosine 5'-monophosphate challenges in children with asthma: relation to severity of the disease. Pediatric Pulmonology. 2000; 30(3):207-214

10 Backer V, Ulrik CS, Wendelboe D, Bach-Mortensen D, Hansen KK, Laursen EM et al. Distribution of serum IgE in children and adolescents aged 7 to 16 years in Copenhagen, in relation to factors of importance. Allergy. 1992; 47(5):484-489

11 Backer V, Nepper-Christensen S, Ulrik CS, von Linstow ML, Porsbjerg C. Factors associated with asthma in young Danish adults. Annals of Allergy, Asthma and Immunology. 2002; 89(2):148-154

12 Baptist AP, Ross JA, Yang Y, Song PX, Clark NM. A randomized controlled trial of a self-regulation intervention for older adults with asthma. Journal of the American Geriatrics Society. 2013; 61(5):747-753

13 Barbanel D, Eldridge S, Griffiths C. Can a self-management programme delivered by a community pharmacist improve asthma control? A randomised trial. Thorax. 2003; 58(10):851-854

14 Basheti IA, Armour CL, Bosnic-Anticevich SZ, Reddel HK. Evaluation of a novel educational strategy, including inhaler-based reminder labels, to improve asthma inhaler technique 1503. Patient Education and Counseling. 2008; 72(1):26-33

15 Basheti IA, Reddel HK, Armour CL, Bosnic-Anticevich SZ. Improved asthma outcomes with a simple inhaler technique intervention by community pharmacists. Journal of Allergy and Clinical Immunology. 2007; 119(6):1537-1538

16 Baur X, Huber H, Degens PO, Allmers H, Ammon J. Relation between occupational asthma case history, bronchial methacholine challenge, and specific challenge test in patients with suspected occupational asthma. American Journal of Industrial Medicine. 1998; 33(2):114-122

17 Bender BG, Apter A, Bogen DK, Dickinson P, Fisher L, Wamboldt FS et al. Test of an interactive voice response intervention to improve adherence to controller medications in adults with asthma. Journal of the American Board of Family Medicine. 2010; 23(2):159-165

18 Berg J, Lindgren P. Economic evaluation of FE(NO) measurement in diagnosis and 1-year management of asthma in Germany. Respiratory Medicine. 2008; 102(2):219-231

19 Berlyne GS, Parameswaran K, Kamada D, Efthimiadis A, Hargreave FE. A comparison of exhaled nitric oxide and induced sputum as markers of airway inflammation. Journal of Allergy and Clinical Immunology. 2000; 106(4):638-644

20 Bhogal SK, Zemek RL, Ducharme F. Written action plans for asthma in children. Cochrane Database of Systematic Reviews. 2006; Issue 3:CD005306. DOI:10.1002/14651858.CD005306.pub2

21 Brand PL, Quanjer PH, Postma DS, Kerstjens HA, Koeter GH, Dekhuijzen PN et al. Interpretation of bronchodilator response in patients with obstructive airways disease. The Dutch Chronic Non-Specific Lung Disease (CNSLD) Study Group. Thorax. 1992; 47(6):429-436

22 British Thoracic Society and Scottish Intercollegiate Guidelines Network. British guideline on the management of asthma: a national clinical guideline, 2014. Available from: https://www.brit-thoracic.org.uk/document-library/clinical-information/asthma/btssign-asthma-guideline-2014/

23 Brouwer AFJ, Visser CAN, Duiverman EJ, Roorda RJ, Brand PLP. Is home spirometry useful in diagnosing asthma in children with nonspecific respiratory symptoms? Pediatric Pulmonology. 2010; 45(4):326-332

24 Buist AS, Vollmer WM, Wilson SR, Frazier EA, Hayward AD. A randomized clinical trial of peak flow versus symptom monitoring in older adults with asthma. American Journal of Respiratory and Critical Care Medicine. 2006; 174(10):1077-1087

25 Burgess SW, Sly PD, Devadason SG. Providing feedback on adherence increases use of preventive medication by asthmatic children. Journal of Asthma. 2010; 47(2):198-201

26 Calhoun WJ, Ameredes BT, King TS, Icitovic N, Bleecker ER, 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

27 Cardinale F, De Benedictis FM, Muggeo V, Giordano P, Loffredo MS, Iacoviello G et al. Exhaled nitric oxide, total serum IgE and allergic sensitization in childhood asthma and allergic rhinitis. Pediatric Allergy and Immunology. 2005; 16(3):236-242

28 Chan DS, Callahan CW, Hatch-Pigott VB, Lawless A, Proffitt HL, Manning NE et al. Internet-based home monitoring and education of children with asthma is comparable to ideal office-based care: results of a 1-year asthma in-home monitoring trial. Pediatrics. 2007; 119(3):569-578

29 Charlton I, Charlton G, Broomfield J, Mullee MA. Evaluation of peak flow and symptoms only self management plans for control of asthma in general practice. BMJ. 1990; 301(6765):1355-1359

30 Chatkin JM, Ansarin K, Silkoff PE, McClean P, Gutierrez C, Zamel N et al. Exhaled nitric oxide as a noninvasive assessment of chronic cough. American Journal of Respiratory and Critical Care Medicine. 1999; 159(6):1810-1813

31 Chatkin JM, Blanco DC, Scaglia N, Wagner MB, Fritscher CC. Impact of a low-cost and simple intervention in enhancing treatment adherence in a Brazilian asthma sample. Journal of Asthma. 2006; 43(4):263-266

32 Chhabra SK. Acute bronchodilator response has limited value in differentiating bronchial asthma from COPD. Journal of Asthma. 2005; 42(5):367-372

33 Choi BW, Yoo KH, Jeong JW, Yoon HJ, Kim SH, Park YM et al. Easy diagnosis of asthma: computer-assisted, symptom-based diagnosis. Journal of Korean Medical Science. 2007; 22(5):832-838

34 Christakis DA, Garrison MM, Lozano P, Meischke H, Zhou C, Zimmerman FJ. Improving parental adherence with asthma treatment guidelines: a randomized controlled trial of an interactive website. Academic Pediatrics. United States 2012; 12(4):302-311

35 Chung KF, Wenzel SE, Brozek JL, Bush A, Castro M, Sterk PJ et al. International ERS/ATS guidelines on definition, evaluation and treatment of severe asthma. European Respiratory Journal. 2014; 43(2):343-373

36 Ciprandi G, Tosca MA, Capasso M. High exhaled nitric oxide levels may predict bronchial reversibility in allergic children with asthma or rhinitis. Journal of Asthma. 2013; 50(1):33-38

37 Cordeiro D, Rudolphus A, Snoey E, Braunstahl GJ. Utility of nitric oxide for the diagnosis of asthma in an allergy clinic population. Allergy and Asthma Proceedings. 2011; 32(2):119-126

38 Cote J, Cartier A, Robichaud P, Boutin H, Malo JL, Rouleau M et al. Influence on asthma morbidity of asthma education programs based on self-management plans following treatment optimization. American Journal of Respiratory and Critical Care Medicine. 1997; 155(5):1509-1514

39 Cowie RL, Revitt SG, Underwood MF, Field SK. The effect of a peak flow-based action plan in the prevention of exacerbations of asthma. Chest. 1997; 112(6):1534-1538

40 Curtis L. Unit costs of health and social care 2013. Canterbury: Personal Social Services Research Unit, University of Kent; 2013. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2013/index.php

41 De Jongste JC, Carraro S, Hop W, Baraldi E. Cutoff values for FENO-guided asthma management. American Journal of Respiratory and Critical Care Medicine. 2009; 180(3):282

42 de Jongste JC, Carraro S, Hop WC, CHARISM Study Group, Baraldi E. Daily telemonitoring of exhaled nitric oxide and symptoms in the treatment of childhood asthma. American Journal of Respiratory and Critical Care Medicine. 2009; 179(2):93-97

43 Deilami GD, Khandashpour M, Paknejad O, Pazooki M. Evaluation of methacholine challenge test results in chronic cough patients referring to clinic of pulmonary disease. Acta Medica Iranica. 2009; 47(3):175-179

44 den Otter JJ, Reijnen GM, van den Bosch WJ, van Schayck CP, Molema J, Van Weel C. Testing bronchial hyper-responsiveness: provocation or peak expiratory flow variability? British Journal of General Practice. 1997; 47(421):487-492

45 Department of Health. NHS reference costs 2011-12. 2012. Available from: https://www.gov.uk/government/publications/nhs-reference-costs-financial-year-2011-to-2012 [Last accessed: 23 October 2013]

46 Department of Health. NHS reference costs 2012-13. 2012. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/261154/nhs_reference_costs_2012-13_acc.pdf [Last accessed: 17 March 2014]

47 Department of Health. NHS Supply Chain Catalogue. 2014. Available from: http://www.supplychain.nhs.uk/ [Last accessed: 21 November 2014]

48 Deschildre A, Beghin L, Salleron J, Iliescu C, Thumerelle C, Santos C et al. Home telemonitoring (forced expiratory volume in 1 s) in children with severe asthma does not reduce exacerbations. European Respiratory Journal. 2012; 39(2):290-296

49 Deykin A, Massaro AF, Drazen JM, Israel E. Exhaled nitric oxide as a diagnostic test for asthma: online versus offline techniques and effect of flow rate. American Journal of Respiratory and Critical Care Medicine. 2002; 165(12):1597-1601

50 Donald KJ, McBurney H, Teichtahl H, Irving L. A pilot study of telephone based asthma management. Australian Family Physician. 2008; 37(3):170-173

51 Drkulec V, Nogalo B, Perica M, Plavec D, Pezer M, Turkalj M. Sensitization profile in differential diagnosis: allergic asthma vs. chronic (nonspecific) cough syndrome. Medical Science Monitor. 2013; 19:409-415

52 Dweik RA, Boggs PB, Erzurum SC, Irvin CG, Leigh MW, Lundberg JO et al. An official ATS clinical practice guideline: interpretation of exhaled nitric oxide levels (FENO) for clinical applications. American Journal of Respiratory and Critical Care Medicine. 2011; 184(5):602-615

53 Eggleston PA. A comparison of the asthmatic response to methacholine and exercise. Journal of Allergy and Clinical Immunology. 1979; 63(2):104-110

54 Ernst P, Spitzer WO, Suissa S, Cockcroft D, Habbick B, Horwitz RI et al. Risk of fatal and near-fatal asthma in relation to inhaled corticosteroid use. JAMA. 1992; 268(24):3462-3464

55 Fortuna AM, Feixas T, Gonzalez M, Casan P. Diagnostic utility of inflammatory biomarkers in asthma: exhaled nitric oxide and induced sputum eosinophil count. Respiratory Medicine. 2007; 101(11):2416-2421

56 Fritsch M, Uxa S, Horak F, 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

57 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

58 Gaig P, Enrique E, Garcia-Ortega P, Olona M, del Mar San Miguel M, Richart C. Asthma, mite sensitization, and sleeping in bunks. Annals of Allergy, Asthma and Immunology. 1999; 82(6):531-533

59 Global Initiative for Asthma. Global strategy for asthma management and prevention, 2014. Available from: www.ginasthma.org

60 Green RH, Brightling CE, 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

61 Gruffydd-Jones K, Hollinghurst S, Ward S, Taylor G. Targeted routine asthma care in general practice using telephone triage. British Journal of General Practice. 2005; 55:918-923

62 Guendelman S, Meade K, Benson M, Chen YQ, Samuels S. Improving asthma outcomes and self-management behaviors of inner-city children: a randomized trial of the Health Buddy interactive device and an asthma diary. Archives of Pediatrics and Adolescent Medicine. 2002; 156(2):114-120

63 Gustafson D, Wise M, Bhattacharya A, Pulvermacher A, Shanovich K, Phillips B et al. The effects of combining web-based eHealth with telephone nurse case management for pediatric asthma control: A randomized controlled trial. Journal of Medical Internet Research. 2012; 14(4):41-59

64 Halterman JS, Sauer J, Fagnano M, Montes G, Fisher S, Tremblay P et al. Working toward a sustainable system of asthma care: Development of the School-Based Preventive Asthma Care Technology (SB-PACT) trial. Journal of Asthma. 2012; 49(4):395-400

65 Halvani A, Tahghighi F, Nadooshan HH. Evaluation of correlation between airway and serum inflammatory markers in asthmatic patients. Lung India. 2012; 29(2):143-146

66 Harnan S, Tappenden P, Essat M, Gomersall T, Minton J, Wong R et al. Measurement of exhaled nitric oxide concentration in asthma - NIOX MINO and NObreath, 2013. Available from: http://guidance.nice.org.uk/DT/13

67 Health and Social Care Information Centre. Health Survey for England - 2010, Respiratory health [NS]. 2011. Available from: http://www.hscic.gov.uk/pubs/hse10report [Last accessed: 16 May 2015]

68 Health and Social Care Information Centre. The Quality and Outcomes Framework (QOF) 2013/14 results. 2015. Available from: http://qof.hscic.gov.uk/ [Last accessed: 16 April 2015]

69 Hedman J, Poussa T, Nieminen MM. A rapid dosimetric methacholine challenge in asthma diagnostics: a clinical study of 230 patients with dyspnoea, wheezing or a cough of unknown cause. Respiratory Medicine. 1998; 92(1):32-39

70 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

71 Honkoop PJ, Loijmans RJ, Termeer EH, Snoeck-Stroband JB, van den Hout WB, Bakker MJ et al. Symptom- and fraction of exhaled nitric oxide-driven strategies for asthma control: A cluster-randomized trial in primary care. Journal of Allergy and Clinical Immunology. 2014; Epublication

72 Honkoop PJ, Loymans RJ, Termeer EH, Snoeck-Stroband JB, Bakker MJ, Assendelft WJ et al. Asthma control cost-utility randomized trial evaluation (ACCURATE): the goals of asthma treatment. BMC Pulmonary Medicine. England 2011; 11:53

73 Hunter CJ, Brightling CE, Woltmann G, Wardlaw AJ, Pavord ID. A comparison of the validity of different diagnostic tests in adults with asthma. Chest. 2002; 121(4):1051-1057

74 Jan RL, Wang JY, Huang MC, Tseng SM, Su HJ, Liu LF. An internet-based interactive telemonitoring system for improving childhood asthma outcomes in Taiwan. Telemedicine Journal and E-Health. 2007; 13(3):257-268

75 Juniper EF, Guyatt GH, Cox FM, Ferrie PJ, King DR. Development and validation of the Mini Asthma Quality of Life Questionnaire. European Respiratory Journal. 1999; 14(1):32-38

76 Juniper EF, Guyatt GH, Feeny DH, Ferrie PJ, Griffith LE, Townsend M. Measuring quality of life in children with asthma. Quality of Life Research. 1996; 5(1):35-46

77 Juniper EF, Guyatt GH, Ferrie PJ, Griffith LE. Measuring quality of life in asthma. American Review of Respiratory Disease. 1993; 147(4):832-838

78 Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. Journal of Clinical Epidemiology. 1994; 47(1):81-87

79 Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. European Respiratory Journal. 1999; 14(4):902-907

80 Kaya Z, Erkan F, Ozkan M, Ozkan S, Kocaman N, Ertekin BA et al. Self-management plans for asthma control and predictors of patient compliance. Journal of Asthma. 2009; 46(3):270-275

81 Kersten ETG, Driessen JMM, van der Berg JD, Thio BJ. Mannitol and exercise challenge tests in asthmatic children. Pediatric Pulmonology. 2009; 44(7):655-661

82 Khakzad MR, Mirsadraee M, Sankian M, Varasteh A, Meshkat M. Is serum or sputum eosinophil cationic protein level adequate for diagnosis of mild asthma? Iranian Journal of Allergy, Asthma and Immunology. 2009; 8(3):155-160

83 Khan MSR, O'Meara M, Stevermuer TL, Henry RL. Randomized controlled trial of asthma education after discharge from an emergency department. Journal of Paediatrics and Child Health. 2004; 40(12):674-677

84 Kim TB, Oh YM, Chang YS, Cho YS, Jang AS, Cho SH et al. The reality of an intermediate type between asthma and COPD in practice. Respiratory Care. 2012; 57(8):1248-1253

85 Klepac-Pulanic T, Macan J, Plavec D, Kanceljak-Macan B. Exercise and allergic diseases. Archives of Industrial Hygiene and Toxicology. 2004; 55(2-3):197-204

86 Koenig SM, Murray JJ, Wolfe J, Andersen L, Yancey S, Prillaman B et al. Does measuring BHR add to guideline derived clinical measures in determining treatment for patients with persistent asthma? Respiratory Medicine. 2008; 102(5):665-673

87 Koskela HO, Hyvarinen L, Brannan JD, Chan HK, Anderson SD. Responsiveness to three bronchial provocation tests in patients with asthma. Chest. 2003; 124(6):2171-2177

88 Kostikas K, Papaioannou AI, Tanou K, Koutsokera A, Papala M, Gourgoulianis KI. Portable exhaled nitric oxide as a screening tool for asthma in young adults during pollen season. Chest. 2008; 133(4):906-913

89 Kotaniemi-Syrjanen A, Reijonen TM, Korhonen K, Korppi M. Wheezing requiring hospitalization in early childhood: predictive factors for asthma in a six-year follow-up. Pediatric Allergy and Immunology. 2002; 13(6):418-425

90 Kowal K, Bodzenta-Lukaszyk A, Zukowski S. Exhaled nitric oxide in evaluation of young adults with chronic cough. Journal of Asthma. 2009; 46(7):692-698

91 Kroegel C, Schuler M, Forster M, Braun R, Grahmann PR. Evidence for eosinophil activation in bronchiectasis unrelated to cystic fibrosis and bronchopulmonary aspergillosis: discrepancy between blood eosinophil counts and serum eosinophil cationic protein levels. Thorax. 1998; 53(6):498-500

92 Labbe A, Aublet-Cuvelier B, Jouaville L, Beaugeon G, Fiani L, Petit I et al. Prospective longitudinal study of urinary eosinophil protein X in children with asthma and chronic cough. Pediatric Pulmonology. 2001; 31(5):354-362

93 Letz KL, Schlie AR, Smits WL. A randomized trial comparing peak expiratory flow versus symptom self-management plans for children with persistent asthma. Pediatric Asthma, Allergy and Immunology. 2004; 17(3):177-190

94 Lin CC, Wu JL, Huang WC, Lin CY. A bronchial response comparison of exercise and methacholine in asthmatic subjects. Journal of Asthma. 1991; 28(1):31-40

95 LindenSmith J, Morrison D, Deveau C, Hernandez P. Overdiagnosis of asthma in the community. Canadian Respiratory Journal. 2004; 11(2):111-116

96 Linneberg A, Husemoen LLN, Nielsen NH, Madsen F, Frolund L, Johansen N. Screening for allergic respiratory disease in the general population with the ADVIA Centaur Allergy Screen Assay. Allergy. 2006; 61(3):344-348

97 Lipworth BJ, Short PM, Williamson PA, Clearie KL, Fardon TC, Jackson CM. A randomized primary care trial of steroid titration against mannitol in persistent asthma: STAMINA trial. Chest. 2012; 141(3):607-615

98 Liu AH, Zeiger R, Sorkness C, Mahr T, Ostrom N, Burgess S et al. Development and cross-sectional validation of the Childhood Asthma Control Test. Journal of Allergy and Clinical Immunology. 2007; 119(4):817-825

99 Liu WT, Huang CD, Wang CH, Lee KY, Lin SM, Kuo HP. A mobile telephone-based interactive self-care system improves asthma control. European Respiratory Journal. 2011; 37(2):310-317

100 Lopez-Vina A, del Castillo-Arevalo E. Influence of peak expiratory flow monitoring on an asthma self-management education programme. Respiratory Medicine. 2000; 94(8):760-766

101 Louhelainen N, Rytila P, Obase Y, Makela M, Haahtela T, Kinnula VL et al. The value of sputum 8-isoprostane in detecting oxidative stress in mild asthma. Journal of Asthma. 2008; 45(2):149-154

102 Main C, Shepherd J, Anderson R, Rogers G, Thompson-Coon J, Liu Z et al. Systematic review and economic analysis of the comparative effectiveness of different inhaled corticosteroids and their

usage with long-acting beta2 agonists for the treatment of chronic asthma in children under the age of 12 years. Health Technology Assessment. 2008; 12(20)

103 Malo JL, Ghezzo H, L'Archeveque J, Lagier F, Perrin B, Cartier A. Is the clinical history a satisfactory means of diagnosing occupational asthma? American Review of Respiratory Disease. 1991; 143(3):528-532

104 Marklund B, Tunsater A, Bengtsson C. How often is the diagnosis bronchial asthma correct? Family Practice. 1999; 16(2):112-116

105 May KL. Allergy to Artemisia vulgaris in the region of Warsaw. Allergologia Et Immunopathologia. 1990; 18(1):57-60

106 McSharry CP, McKay IC, Chaudhuri R, Livingston E, Fraser I, Thomson NC. Short and long-term effects of cigarette smoking independently influence exhaled nitric oxide concentration in asthma. Journal of Allergy and Clinical Immunology. 2005; 116(1):88-93

107 Mehuys E, Van Bortel L, De Bolle L, Van Tongelen I, Annemans L, Remon JP et al. Effectiveness of pharmacist intervention for asthma control improvement. European Respiratory Journal. 2008; 31(4):790-799

108 Metso T, Kilpio K, Bjorksten F, Kiviranta K, Haahtela T. Detection and treatment of early asthma. Allergy. 2000; 55(5):505-509

109 Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A et al. Standardisation of spirometry. European Respiratory Journal. 2005; 26(2):319-338

110 Miraglia Del Giudice M, Pedulla M, Piacentini GL, Capristo C, Brunese FP, Decimo F et al. Atopy and house dust mite sensitization as risk factors for asthma in children. Allergy. 2002; 57(2):169-172

111 Nadif R, Matran R, Maccario J, Bechet M, Le Moual N, Scheinmann P et al. Passive and active smoking and exhaled nitric oxide levels according to asthma and atopy in adults. Annals of Allergy, Asthma and Immunology. 2010; 104(5):385-393

112 Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, Marcus P et al. Development of the asthma control test: a survey for assessing asthma control. Journal of Allergy and Clinical Immunology. 2004; 113(1):59-65

113 National Institute for Health and Clinical Excellence. Social value judgements: principles for the development of NICE guidance. 2nd edition. London: National Institute for Health and Clinical Excellence; 2008. Available from: https://www.nice.org.uk/media/default/about/what-we-do/research-and-development/social-value-judgements-principles-for-the-development-of-nice-guidance.pdf

114 National Institute for Health and Clinical Excellence. The guidelines manual. London: National Institute for Health and Clinical Excellence; 2012. Available from: http://publications.nice.org.uk/the-guidelines-manual-pmg6/

115 National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal 2013. 2nd edition. London: National Institute for Health and Clinical Excellence; 2013. Available from: http://publications.nice.org.uk/pmg9

116 Nieminen MM. Unimodal distribution of bronchial hyperresponsiveness to methacholine in asthmatic patients. Chest. 1992; 102(5):1537-1543

117 Nordlund B, Konradsen JR, Kull I, Borres MP, Onell A, Hedlin G et al. IgE antibodies to animal-derived lipocalin, kallikrein and secretoglobin are markers of bronchial inflammation in severe childhood asthma. Allergy. 2012; 67(5):661-669

118 Nuijsink M, Hop WCJ, Sterk PJ, Duiverman EJ, De Jongste JC. Long-term asthma treatment guided by airway hyperresponsiveness in children: a randomised controlled trial. European Respiratory Journal. 2007; 30(3):457-466

119 Onyirimba F, Apter A, Reisine S, Litt M, McCusker C, Connors M et al. Direct clinician-to-patient feedback discussion of inhaled steroid use: its effect on adherence. Annals of Allergy, Asthma and Immunology. 2003; 90(4):411-415

120 Organisation for Economic Co-operation and Development (OECD). Purchasing power parities (PPP). 2013. Available from: http://www.oecd.org/std/ppp [Last accessed: 27 November 2013]

121 Ostojic V, Cvoriscec B, Ostojic SB, Reznikoff D, Stipic-Markovic A, Tudjman Z. Improving asthma control through telemedicine: A study of short-message service. Telemedicine Journal and E-Health. 2005; 11(1):28-35

122 Otsuki M, Eakin MN, Rand CS, Butz AM, Hsu VD, Zuckerman IH et al. Adherence feedback to improve asthma outcomes among inner-city children: a randomized trial. Pediatrics. 2009; 124(6):1513-1521

123 Palange P, Ward SA, Carlsen KH, Casaburi R, Gallagher CG, Gosselink R et al. Recommendations on the use of exercise testing in clinical practice. European Respiratory Journal. 2007; 29(1):185-209

124 Parsons JP, Hallstrand TS, Mastronarde JG, Kaminsky DA, Rundell KW, Hull JH et al. An official American Thoracic Society clinical practice guideline: exercise-induced bronchoconstriction. American Journal of Respiratory and Critical Care Medicine. 2013; 187(9):1016-1027

125 Peirsman EJ, Carvelli TJ, Hage PY, Hanssens LS, Pattyn L, Raes MM et al. Exhaled nitric oxide in childhood allergic asthma management a randomised controlled trial. Pediatric Pulmonology. 2013;

126 Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R et al. Interpretative strategies for lung function tests. European Respiratory Journal. 2005; 26(5):948-968

127 Petsky HL, Li AM, Au CT, Kynaston JA, Turner C, Chang AB. Management based on exhaled nitric oxide levels adjusted for atopy reduces asthma exacerbations in children: A dual centre randomized controlled trial. Pediatric Pulmonology. 2014;

128 Petsky HL, Cates CJ, Li A, Kynaston JA, Turner C, Chang AB. Tailored interventions based on exhaled nitric oxide versus clinical symptoms for asthma in children and adults. Cochrane Database of Systematic Reviews. 2009; Issue 4:CD006340. DOI:10.1002/14651858.CD006340.pub3

129 Piippo-Savolainen E, Remes S, Korppi M. Does blood eosinophilia in wheezing infants predict later asthma? A prospective 18-20-year follow-up. Allergy and Asthma Proceedings. 2007; 28(2):163-169

130 Pijnenburg MW, Bakker EM, Hop WC, De Jongste JC. Titrating inhaled steroids on exhaled nitric oxide improves FEV1 in allergic asthmatic children. American Thoracic Society 2005 International Conference; May 20-25; San Diego, California. 2005;C47

131 Pike K, Selby A, Price S, Warner J, Connett G, Legg J et al. Exhaled nitric oxide monitoring does not reduce exacerbation frequency or inhaled corticosteroid dose in paediatric asthma: a randomised controlled trial. Clinical Respiratory Journal. 2012;

132 Pinnock H, Adlem L, Gaskin S, Harris J, Snellgrove C, Sheikh A. Accessibility, clinical effectiveness, and practice costs of providing a telephone option for routine asthma reviews: phase IV controlled implementation study. British Journal of General Practice. United Kingdom 2007; 57:714-722

133 Pinnock H, Bawden R, Proctor S, Wolfe S, Scullion J, Price D et al. Accessibility, acceptability, and effectiveness in primary care of routine telephone review of asthma: pragmatic, randomised controlled trial. BMJ. 2003; 326(7387):477-479

134 Pinnock H, McKenzie L, Price D, Sheikh A. Cost-effectiveness of telephone or surgery asthma reviews: economic analysis of a randomised controlled trial. British Journal of General Practice. 2005; 55:119-124:119-124

135 Pino JM, Garcia-Rio F, Prados C, Alvarez-Sala R, Diaz S, Villasante C et al. Value of the peak expiratory flow in bronchodynamic tests. Allergologia Et Immunopathologia. 1996; 24(2):54-57

136 Plaschke P, Janson C, Norrman E, Björnsson E, Ellbjär S, Järvholm B. Association between atopic sensitization and asthma and bronchial hyperresponsiveness in swedish adults: pets, and not mites, are the most important allergens. Journal of Allergy and Clinical Immunology. 1999; 104(1):58-65

137 Popovic-Grle S, Mehulic M, Pavicic F, Babic I, Beg-Zec Z. Clinical validation of bronchial hyperresponsiveness, allergy tests and lung function in the diagnosis of asthma in persons with dyspnea. Collegium Antropologicum. 2002; 26 Suppl:119-127

138 Powell H, Murphy VE, Taylor DR, Hensley MJ, 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

139 Prabhakaran L, Chee J, Chua KC, Mun WW. The Use of Text Messaging to Improve Asthma Control: A Study of Short Message Service (SMS). Respirology. 2009; 14(Suppl 3):A217

140 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

141 Quadrelli SA, Roncoroni AJ, Montiel GC. Evaluation of bronchodilator response in patients with airway obstruction. Respiratory Medicine. 1999; 93(9):630-636

142 Rasmussen LM, Phanareth K, Nolte H, Backer V. Can internet-based management improve asthma control? A long term randomised clinical study of 300 asthmatics. European Respiratory Journal. 2005; 26(Suppl.49)

143 Rikkers-Mutsaerts ERVM, Winters AE, Bakker MJ, van Stel HF, van der Meer V, De Jongste JC et al. Internet-based self-management compared with usual care in adolescents with asthma: a randomized controlled trial. Pediatric Pulmonology. 2012; 47(12):1170-1179

144 Royal College of Physicians of London. Why asthma still kills: The national review of asthma deaths (NRAD) confidential enquiry report May 2014. London. Royal College of Physicians, 2014. Available from: https://www.rcplondon.ac.uk/sites/default/files/why-asthma-still-kills-full-report.pdf

145 Rust G, Zhang S, Reynolds J. Inhaled corticosteroid adherence and emergency department utilization among Medicaid-enrolled children with asthma. Journal of Asthma. 2013; 50(7):769-775

146 Ryan D, Price D, Musgrave SD, Malhotra S, Lee AJ, Ayansina D et al. Clinical and cost effectiveness of mobile phone supported self monitoring of asthma: multicentre randomised controlled trial. BMJ. 2012; 344:e1756

147 Rytila P, Metso T, Heikkinen K, Saarelainen P, Helenius IJ, Haahtela T. Airway inflammation in patients with symptoms suggesting asthma but with normal lung function. European Respiratory Journal. 2000; 16(5):824-830

148 Santanello NC, Zhang J, Seidenberg B, Reiss TF, Barber BL. What are minimal important changes for asthma measures in a clinical trial? European Respiratory Journal. 1999; 14(1):23-27

149 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

150 Schatz M, Kosinski M, Yarlas AS, Hanlon J, Watson ME, Jhingran P. The minimally important difference of the Asthma Control Test. Journal of Allergy and Clinical Immunology. 2009; 124(4):719-723

151 Schleich FN, 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

152 Schneider A, Ay M, Faderl B, Linde K, Wagenpfeil S. Diagnostic accuracy of clinical symptoms in obstructive airway diseases varied within different health care sectors. Journal of Clinical Epidemiology. 2012; 65(8):846-854

153 Schneider A, Gindner L, Tilemann L, Schermer T, Dinant GJ, Meyer FJ et al. Diagnostic accuracy of spirometry in primary care. BMC Pulmonary Medicine. 2009; 9:31

154 Seid M, D'Amico EJ, Varni JW, Munafo JK, Britto MT, Kercsmar CM et al. The In Vivo adherence intervention for at risk adolescents with asthma: Report of a randomized pilot study. Journal of Pediatric Psychology. United States 2012; 37(4):390-403

155 Shaw D, Green R, Berry M, Mellor S, Hargadon B, Shelley M et al. A cross-sectional study of patterns of airway dysfunction, symptoms and morbidity in primary care asthma. Primary Care Respiratory Journal. 2012; 21(3):283-287

156 Shaw DE, Berry MA, Thomas M, Green RH, Brightling CE, Wardlaw AJ et al. The use of exhaled nitric oxide to guide asthma management: a randomized controlled trial. American Journal of Respiratory and Critical Care Medicine. 2007; 176(3):231-237

157 Shields MD, Brown V, Stevenson EC, Fitch PS, Schock BC, Turner G et al. Serum eosinophilic cationic protein and blood eosinophil counts for the prediction of the presence of airways inflammation in children with wheezing. Clinical and Experimental Allergy. 1999; 29(10):1382-1389

158 Shimoda T, Obase Y, Kishikawa R, Iwanaga T, Miyatake A, Kasayama S. The fractional exhaled nitric oxide and serum high sensitivity C-reactive protein levels in cough variant asthma and typical bronchial asthma. Allergology International. 2013; 62(2):251-257

159 Shome GP, Starnes III JD, Shearer M, Kennedy R, Way A, Arif A et al. Exhaled nitric oxide in asthma: Variability, relation to asthma severity, and peripheral blood lymphocyte cytokine expression. Journal of Asthma. 2006; 43(2):95-99

160 Silvestri M, Oddera S, Rossi GA, Crimi P. Sensitization to airborne allergens in children with respiratory symptoms. Annals of Allergy, Asthma and Immunology. 1996; 76(3):239-244

161 Silvestri M, Sabatini F, Spallarossa D, Fregonese L, Battistini E, Biraghi MG et al. Exhaled nitric oxide levels in non-allergic and allergic mono- or polysensitised children with asthma. Thorax. 2001; 56(11):857-862

162 Silvestri M, Sabatini F, Sale R, Defilippi AC, Fregonese L, Battistini E et al. Correlations between exhaled nitric oxide levels, blood eosinophilia, and airway obstruction reversibility in childhood asthma are detectable only in atopic individuals. Pediatric Pulmonology. 2003; 35(5):358-363

163 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

164 Smith AD, Cowan JO, Brassett KP, Herbison GP, Taylor DR. Use of exhaled nitric oxide measurements to guide treatment in chronic asthma. New England Journal of Medicine. 2005; 352(21):2163-2173

165 Smith AD, Cowan JO, 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 and Critical Care Medicine. 2004; 169(4):473-478

166 Sont JK, Willems LN, Bel EH, van Krieken JH, Vandenbroucke JP, Sterk PJ. Clinical control and histopathologic outcome of asthma when using airway hyperresponsiveness as an additional guide to long-term treatment. The AMPUL Study Group. American Journal of Respiratory and Critical Care Medicine. 1999; 159(4 Pt 1):1043-1051

167 Soriano JB, Anto JM, Sunyer J, Tobias A, Kogevinas M, Almar E et al. Risk of asthma in the general Spanish population attributable to specific immunoresponse. Spanish Group of the European Community Respiratory Health Survey. International Journal of Epidemiology. 1999; 28(4):728-734

168 Soriano JB, Anto JM, Sunyer J, Tobias A, Kogevinas M, Almar E et al. Risk of asthma in the general Spanish population attributable to specific immunoresponse. International Journal of Epidemiology. 1999; 28(4):728-734

169 Starren ES, Roberts NJ, Tahir M, O'Byrne L, Haffenden R, Patel IS et al. A centralised respiratory diagnostic service for primary care: a 4-year audit. Primary Care Respiratory Journal. 2012; 21(2):180-186

170 Suissa S, Ernst P, Benayoun S, Baltzan M, Cai B. Low-dose inhaled corticosteroids and the prevention of death from asthma. New England Journal of Medicine. 2000; 343(5):332-336

171 Suissa S, Ernst P, Kezouh A. Regular use of inhaled corticosteroids and the long term prevention of hospitalisation for asthma. Thorax. 2002; 57(10):880-884

172 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. Journal of Allergy and Clinical Immunology: In Practice. 2013; 1(6):639-648

173 Szefler SJ, Mitchell H, Sorkness CA, Gergen PJ, O'Connor GT, Morgan WJ 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

174 Thiadens HA, De Bock GH, Dekker FW, Huysman JA, Van Houwelingen JC, Springer MP et al. Value of measuring diurnal peak flow variability in the recognition of asthma: a study in general practice. European Respiratory Journal. 1998; 12(4):842-847

175 Tilemann L, Gindner L, Meyer F, Szecsenyi J, Schneider A. Differences in local and systemic inflammatory markers in patients with obstructive airways disease. Primary Care Respiratory Journal. 2011; 20(4):407-414

176 Tomasiak-Lozowska MM, Zietkowski Z, Przeslaw K, Tomasiak M, Skiepko R, Bodzenta-Lukaszyk A. Inflammatory markers and acid-base equilibrium in exhaled breath condensate of stable and unstable asthma patients. International Archives of Allergy and Immunology. 2012; 159(2):121-129

177 Tomita K, Sano H, Chiba Y, Sato R, Sano A, Nishiyama O et al. A scoring algorithm for predicting the presence of adult asthma: a prospective derivation study. Primary Care Respiratory Journal. 2013; 22(1):51-58

178 Tschopp JM, Sistek D, Schindler C, Leuenberger P, Perruchoud AP, Wuthrich B et al. Current allergic asthma and rhinitis: diagnostic efficiency of three commonly used atopic markers (IgE, skin prick tests, and Phadiatop). Results from 8329 randomized adults from the SAPALDIA Study. Swiss Study on Air Pollution and Lung Diseases in Adults. Allergy. 1998; 53(6):608-613

179 Tuchinda M, Habananada S, Vareenil J, Srimaruta N, Piromrat K. Asthma in Thai children: a study of 2000 cases. Annals of Allergy. 1987; 59(3):207-211

180 Turner MO, Taylor D, Bennett R, Fitzgerald JM. A randomized trial comparing peak expiratory flow and symptom self-management plans for patients with asthma attending a primary care clinic. American Journal of Respiratory and Critical Care Medicine. 1998; 157(2):540-546

181 Ulrik CS, Postma DS, Backer V. Recognition of asthma in adolescents and young adults: which objective measure is best? Journal of Asthma. 2005; 42(7):549-554

182 University of Bristol. QUADAS-2. 2011. Available from: http://www.bris.ac.uk/quadas/quadas-2/ [Last accessed: 18 February 2014]

183 van der Mark LB, van Wonderen KE, Mohrs J, van Aalderen WMC, Ter Riet G, Bindels PJE. Predicting asthma in preschool children at high risk presenting in primary care: development of a clinical asthma prediction score. Primary Care Respiratory Journal. 2014; 23(1):52-59

184 van der Meer V, Bakker MJ, van den Hout WB, Rabe KF, Sterk PJ, Kievit J et al. Internet-based self-management plus education compared with usual care in asthma: a randomized trial. Annals of Internal Medicine. 2009; 151(2):110-120

185 Vandenplas O, Binard-Van Cangh F, Brumagne A, Caroyer JM, Thimpont J, Sohy C et al. Occupational asthma in symptomatic workers exposed to natural rubber latex: evaluation of diagnostic procedures. Journal of Allergy and Clinical Immunology. 2001; 107(3):542-547

186 Vandenplas O, Ghezzo H, Munoz X, Moscato G, Perfetti L, Lemiere C et al. What are the questionnaire items most useful in identifying subjects with occupational asthma? European Respiratory Journal. 2005; 26(6):1056-1063

187 Verini M, Consilvio NP, Di PS, 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; 2010:691425

188 Verleden GM, Dupont LJ, Verpeut AC, Demedts MG. The effect of cigarette smoking on exhaled nitric oxide in mild steroid-naive asthmatics. Chest. 1999; 116(1):59-64

189 Vila-Indurain B, Munoz-Lopez F, Martin-Mateos M. Evaluation of blood eosinophilia and the eosinophil cationic protein (ECP) in the serum of asthmatic children with varying degree of severity. Allergologia Et Immunopathologia. 1999; 27(6):304-308

190 Vollmer WM, Kirshner M, Peters D, Drane A, Stibolt T, Hickey T et al. Use and impact of an automated telephone outreach system for asthma in a managed care setting. American Journal of Managed Care. 2006; 12(12):725-733

191 Voutilainen M, Malmberg LP, Vasankari T, Haahtela T. Exhaled nitric oxide indicates poorly athlete's asthma. Clinical Respiratory Journal. 2013; 7(4):347-353

192 Wensley D, Silverman M. Peak flow monitoring for guided self-management in childhood asthma: a randomized controlled trial. American Journal of Respiratory and Critical Care Medicine. 2004; 170(6):606-612

193 Wever-Hess J, Kouwenberg JM, Duiverman EJ, Hermans J, Wever AM. Prognostic characteristics of asthma diagnosis in early childhood in clinical practice. Acta Paediatrica. 1999; 88(8):827-834

194 Willems DC, Joore MA, Hendriks JJ, van Duurling RA, Wouters EF, Severens JL. Process evaluation of a nurse-led telemonitoring programme for patients with asthma. Journal of Telemedicine and Telecare. 2007; 13(6):310-317

195 Willems DC, Joore MA, Hendriks JJ, Wouters EF, Severens JL. Cost-effectiveness of a nurse-led telemonitoring intervention based on peak expiratory flow measurements in asthmatics: results of a randomised controlled trial. Cost Effectiveness and Resource Allocation. 2007; 5:10

196 Williams LK, Peterson EL, Wells K, Campbell J, Wang M, Chowdhry VK et al. A cluster-randomized trial to provide clinicians inhaled corticosteroid adherence information for their patients with asthma. Journal of Allergy and Clinical Immunology. 2010; 126(2):225-4

197 Woo SI, Lee JH, Kim H, Kang JW, Sun YH, Hahn YS. Utility of fractional exhaled nitric oxide (F(E)NO) measurements in diagnosing asthma. Respiratory Medicine. 2012; 106(8):1103-1109

198 Xu C, Jackson M, Scuffham PA, Wootton R, Simpson P, Whitty J et al. A randomized controlled trial of an interactive voice response telephone system and specialist nurse support for childhood asthma management. Journal of Asthma. 2010; 47(7):768-773

199 Yoos HL, Kitzman H, McMullen A, Henderson C, Sidora K. Symptom monitoring in childhood asthma: a randomized clinical trial comparing peak expiratory flow rate with symptom monitoring. Annals of Allergy, Asthma and Immunology. 2002; 88(3):283-291

200 Young HN, Havican SN, Griesbach S, Thorpe JM, Chewning BA, Sorkness CA. Patient and phaRmacist telephonic encounters (PARTE) in an underserved rural patient population with asthma: results of a pilot study. Telemedicine Journal and E-Health. 2012; 18(6):427-433

201 Zietkowski Z, Bodzenta-Lukaszyk A, Tomasiak MM, Skiepko R, Szmitkowski M. Comparison of exhaled nitric oxide measurement with conventional tests in steroid-naive asthma patients. Journal of Investigational Allergology and Clinical Immunology. 2006; 16(4):239-246

32 Acronyms and abbreviations

<	Less than
>	More than
≤	Less than or equal to
≥	More than or equal to
A&E	Accident and emergency
ACQ	Asthma Control Questionnaire
ACT	Asthma Control Test
AHR	Airway hyper-reactivity
AQLQ	Asthma Quality of Life Questionnaire
ATS	American Thoracic Society
BCT	Bronchial challenge test
BDR	Bronchodilator reversibility
BHR	Bronchial hyper-reactivity
BTS	British Thoracic Society
CACT	Children's Asthma Control Test
CS	Corticosteroid
Dx	Diagnosis
ED	Emergency department
ERS	European Respiratory Society
FeNO	Fractional exhaled nitric oxide
FEV_1	Forced expiratory volume in one second
FVC	Forced vital capacity
GC	Guideline development group
GINA	Global Initiative for Asthma
ICS	Inhaled corticosteroids
IgE	Immunoglobulin E
LABA	Long-acting beta agonist
MCT	Methacholine challenge test
N/A	Not applicable
NAEPP	National asthma education and prevention program
NGC	National Guideline Centre
NICE	National Institute for Health and Care Excellence
NIH	National Institutes of Health
OAD	Obstructive airways disease
OCS	Oral corticosteroids
pAQLQ	Paediatric Asthma Quality of Life Questionnaire
PACQ	Paediatric Asthma Control Questionnaire
PBE	Peripheral blood eosinophils
PEF	Peak expiratory flow
PEFv	Peak expiratory flow variability
ppb	Parts per billion
RCP 3 Questions	Royal College of Physicians 3 Questions

RCP NRAD	Royal College of Physicians National Review of Asthma Deaths
SABA	Short-acting beta agonist
SIGN	Scottish Intercollegiate Guidelines Network
sn	Sensitivity
sp	Specificity
SIC	Specific inhalation challenge
SPT	Skin prick test
THC	Tele-healthcare
Tx	Treatment
UC	Usual care
UHU	Unscheduled healthcare utilisation

33 Glossary

33.1 Guideline-specific terms

Adherence (to treatment)	The extent to which a patient's action matches the agreed recommendations.
Airway hyper-reactivity	See 'bronchial hyper-reactivity'.
Asthma	A common long-term incurable condition of unknown cause that affects people of all ages whereby the small tubes in the lungs (bronchi) become inflamed when the person encounters something that irritates their lungs (asthma triggers) causing the airways to become narrower making it difficult to breathe and can induce coughing, wheezing and tightness in the chest. Asthma is usually associated with an expiratory polyphonic wheeze. Severity of symptoms varies from person to person; and even in the same person at different times of the day or year. Worsening of symptoms can occur gradually or suddenly (known as an 'asthma attack' or 'asthma exacerbation').
Asthma attack	A worsening of asthma symptoms requiring the use of systemic corticosteroids to prevent a serious outcome.
Asthma exacerbation	See 'asthma attack'.
Atopic disorders	Allergic conditions including allergic rhinitis (hay fever), atopic dermatitis (eczema), allergic asthma and other specific and non-specific allergic conditions such as food allergies.
Bronchial challenge test	A test to measure airway reactivity after inhalation of a non-specific drug.
Bronchial hyper-reactivity	A measure of how easily bronchospasm can be induced in the airways.
Bronchodilator	A drug that widens the airways making it easier to breathe.
Bronchodilator response	See 'bronchodilator reversibility'.
Bronchodilator reversibility	A measure of the ability to reverse an obstruction in the airways using drugs that widen the airways (bronchodilators).
Emergency department	Hospital department that assesses and treats patients with serious or life-threatening injuries or illnesses.
Eosinophilia	A higher than normal number of the type of white blood cell eosinophils circulating in the blood.
Exercise	Any physical activity requiring effort or exertion of the body at a greater intensity than that of a normal resting state.
FeNO test	A test that measures the amount of nitric oxide (NO) present upon exhalation, generally expressed in parts per billion.
FEV_1	The amount of air you can blow out in one second (forced expiratory volume in one second).
Forced vital capacity	The amount of air which can be forcibly exhaled from the lungs after taking the deepest breath possible.
Histamine	An organic chemical compound which is released by cells in the body as part of a local immune response to certain allergic stimuli causing an inflammatory response and the constriction of smooth muscle.
IgE test	A blood test that measures the amount of IgE antibody circulating in the blood.
Inhaler	A portable device for administering an inhaled drug.
Mannitol	An osmotic diuretic which leads to constriction of the airways.
Methacholine	A synthetic compound that causes constriction of the airways.
Objective test	A test designed to exclude as far as possible the subjective element on the

	part of the person taking, the person administering and the person assessing, the test.
Occupational asthma	Work-related asthma attributable to a particular exposure in the workplace and not due to stimuli encountered outside the workplace.
Peak expiratory flow rate	A measure of the maximum speed of expiration, generally expressed in litres per minute.
Peak expiratory flow variability	A measure of how much the maximum speed of expiration varies in a person over time.
Peripheral blood eosinophil count	A blood test that measures the number of the type of white blood cell eosinophils circulating in the blood.
Questionnaire	A written set of questions on a particular topic designed for the purpose of gathering specific information from a respondent.
Sensitivity (degree of)	• Low: 0-50% • Moderate: 50-75% • High: 75-100% See also 'Sensitivity' in the list of general terms below.
Skin prick test	A test that measures the allergic response of an individual to certain specific allergens when a very small amount of the specific allergen is introduced into the skin (usually the inner forearm).
Specificity (degree of)	• Low: 0-50% • Moderate: 50-75% • High: 75-100% See also 'Specificity' in the list of general terms below.
Spirometry	A test that measures how a person exhales volumes of air as a function of time.
Tele-healthcare	Information and communication technologies used by patients and healthcare professionals to deliver healthcare, health promotion or to carry out research where the participants are not in the same place. Examples include telephone interviews with a healthcare professional, internet and smartphone-based monitoring support.
Wheeze	A continuous, coarse, whistling sound produced in the airways during breathing (inspiration or expiration) due to a narrowing or obstruction in a part of the respiratory tree. Can be polyphonic (multiple pitches and tones heard over a variable area of the lung) or monophonic (a single pitch and tonal quality heard over an isolated area of the lung).

33.2 General terms

Abstract	Summary of a study, which may be published alone or as an introduction to a full scientific paper.
Algorithm (in guidelines)	A flow chart of the clinical decision pathway described in the guideline, where decision points are represented with boxes, linked with arrows.
Allocation concealment	The process used to prevent advance knowledge of group assignment in an RCT. The allocation process should be impervious to any influence by the individual making the allocation, by being administered by someone who is not responsible for recruiting participants.
Applicability	How well the results of a study or NICE evidence review can answer a clinical question or be applied to the population being considered.
Arm (of a clinical study)	Subsection of individuals within a study who receive one particular intervention, for example placebo arm.
Association	Statistical relationship between 2 or more events, characteristics or other

	variables. The relationship may or may not be causal.
Baseline	The initial set of measurements at the beginning of a study (after run-in period where applicable), with which subsequent results are compared.
Bias	Influences on a study that can make the results look better or worse than they really are. (Bias can even make it look as if a treatment works when it does not.) Bias can occur by chance, deliberately or as a result of systematic errors in the design and execution of a study. It can also occur at different stages in the research process, for example, during the collection, analysis, interpretation, publication or review of research data. For examples see selection bias, performance bias, information bias, confounding factor, and publication bias.
Blinding	A way to prevent researchers, doctors and patients in a clinical trial from knowing which study group each patient is in so they cannot influence the results. The best way to do this is by sorting patients into study groups randomly. The purpose of 'blinding' or 'masking' is to protect against bias. A single-blinded study is one in which patients do not know which study group they are in (for example whether they are taking the experimental drug or a placebo). A double-blinded study is one in which neither patients nor the researchers/doctors know which study group the patients are in. A triple blind study is one in which neither the patients, clinicians or the people carrying out the statistical analysis know which treatment patients received.
Carer (caregiver)	Someone who looks after family, partners or friends in need of help because they are ill, frail or have a disability.
Case–control study	A study to find out the cause(s) of a disease or condition. This is done by comparing a group of patients who have the disease or condition (cases) with a group of people who do not have it (controls) but who are otherwise as similar as possible (in characteristics thought to be unrelated to the causes of the disease or condition). This means the researcher can look for aspects of their lives that differ to see if they may cause the condition. For example, a group of people with lung cancer might be compared with a group of people the same age that do not have lung cancer. The researcher could compare how long both groups had been exposed to tobacco smoke. Such studies are retrospective because they look back in time from the outcome to the possible causes of a disease or condition.
Case series	Report of a number of cases of a given disease, usually covering the course of the disease and the response to treatment. There is no comparison (control) group of patients.
Clinical efficacy	The extent to which an intervention is active when studied under controlled research conditions.
Clinical effectiveness	How well a specific test or treatment works when used in the 'real world' (for example, when used by a doctor with a patient at home), rather than in a carefully controlled clinical trial. Trials that assess clinical effectiveness are sometimes called management trials. Clinical effectiveness is not the same as efficacy.
Clinician	A healthcare professional who provides patient care. For example, a doctor, nurse or physiotherapist.
Cochrane Review	The Cochrane Library consists of a regularly updated collection of evidence-based medicine databases including the Cochrane Database of Systematic Reviews (reviews of randomised controlled trials prepared by the Cochrane Collaboration).
Cohort study	A study with 2 or more groups of people – cohorts – with similar characteristics. One group receives a treatment, is exposed to a risk factor or has a particular symptom and the other group does not. The study

	follows their progress over time and records what happens. See also observational study.
Comorbidity	A disease or condition that someone has in addition to the health problem being studied or treated.
Comparability	Similarity of the groups in characteristics likely to affect the study results (such as health status or age).
Confidence interval (CI)	There is always some uncertainty in research. This is because a small group of patients is studied to predict the effects of a treatment on the wider population. The confidence interval is a way of expressing how certain we are about the findings from a study, using statistics. It gives a range of results that is likely to include the 'true' value for the population. The CI is usually stated as '95% CI', which means that the range of values has a 95 in a 100 chance of including the 'true' value. For example, a study may state that 'based on our sample findings, we are 95% certain that the 'true' population blood pressure is not higher than 150 and not lower than 110'. In such a case the 95% CI would be 110 to 150. A wide confidence interval indicates a lack of certainty about the true effect of the test or treatment - often because a small group of patients has been studied. A narrow confidence interval indicates a more precise estimate (for example, if a large number of patients have been studied).
Confounding factor	Something that influences a study and can result in misleading findings if it is not understood or appropriately dealt with. For example, a study of heart disease may look at a group of people that exercises regularly and a group that does not exercise. If the ages of the people in the 2 groups are different, then any difference in heart disease rates between the 2 groups could be because of age rather than exercise. Therefore age is a confounding factor.
Consensus methods	Techniques used to reach agreement on a particular issue. Consensus methods may be used to develop NICE guidance if there is not enough good quality research evidence to give a clear answer to a question. Formal consensus methods include Delphi and nominal group techniques.
Control group	A group of people in a study who do not receive the treatment or test being studied. Instead, they may receive the standard treatment (sometimes called 'usual care') or a dummy treatment (placebo). The results for the control group are compared with those for a group receiving the treatment being tested. The aim is to check for any differences. Ideally, the people in the control group should be as similar as possible to those in the treatment group, to make it as easy as possible to detect any effects due to the treatment.
Cost–benefit analysis (CBA)	Cost-benefit analysis is one of the tools used to carry out an economic evaluation. The costs and benefits are measured using the same monetary units (for example, pounds sterling) to see whether the benefits exceed the costs.
Cost–consequences analysis (CCA)	Cost-consequence analysis is one of the tools used to carry out an economic evaluation. This compares the costs (such as treatment and hospital care) and the consequences (such as health outcomes) of a test or treatment with a suitable alternative. Unlike cost-benefit analysis or cost-effectiveness analysis, it does not attempt to summarise outcomes in a single measure (like the quality-adjusted life year) or in financial terms. Instead, outcomes are shown in their natural units (some of which may be monetary) and it is left to decision-makers to determine whether, overall, the treatment is worth carrying out.
Cost-effectiveness analysis (CEA)	Cost-effectiveness analysis is one of the tools used to carry out an economic evaluation. The benefits are expressed in non-monetary terms related to

	health, such as symptom-free days, heart attacks avoided, deaths avoided or life years gained (that is, the number of years by which life is extended as a result of the intervention).
Cost-effectiveness model	An explicit mathematical framework, which is used to represent clinical decision problems and incorporate evidence from a variety of sources in order to estimate the costs and health outcomes.
Cost–utility analysis (CUA)	Cost-utility analysis is one of the tools used to carry out an economic evaluation. The benefits are assessed in terms of both quality and duration of life, and expressed as quality-adjusted life years (QALYs). See also utility.
Decision analysis	An explicit quantitative approach to decision-making under uncertainty, based on evidence from research. This evidence is translated into probabilities, and then into diagrams or decision trees which direct the clinician through a succession of possible scenarios, actions and outcomes.
Discounting	Costs and perhaps benefits incurred today have a higher value than costs and benefits occurring in the future. Discounting health benefits reflects individual preference for benefits to be experienced in the present rather than the future. Discounting costs reflects individual preference for costs to be experienced in the future rather than the present.
Dominance	A health economics term. When comparing tests or treatments, an option that is both less effective and costs more is said to be 'dominated' by the alternative.
Drop-out	A participant who withdraws from a trial before the end.
Economic evaluation	An economic evaluation is used to assess the cost-effectiveness of healthcare interventions (that is, to compare the costs and benefits of a healthcare intervention to assess whether it is worth doing). The aim of an economic evaluation is to maximise the level of benefits - health effects - relative to the resources available. It should be used to inform and support the decision-making process; it is not supposed to replace the judgement of healthcare professionals. There are several types of economic evaluation: cost-benefit analysis, cost-consequence analysis, cost-effectiveness analysis, cost-minimisation analysis and cost-utility analysis. They use similar methods to define and evaluate costs, but differ in the way they estimate the benefits of a particular drug, programme or intervention.
Effect (as in effect measure, treatment effect, estimate of effect, effect size)	A measure that shows the magnitude of the outcome in one group compared with that in a control group. For example, if the absolute risk reduction is shown to be 5% and it is the outcome of interest, the effect size is 5%. The effect size is usually tested, using statistics, to find out how likely it is that the effect is a result of the treatment and has not just happened by chance (that is, to see if it is statistically significant).
Effectiveness	How beneficial a test or treatment is under usual or everyday conditions, compared with doing nothing or opting for another type of care.
Efficacy	How beneficial a test, treatment or public health intervention is under ideal conditions (for example, in a laboratory), compared with doing nothing or opting for another type of care.
Epidemiological study	The study of a disease within a population, defining its incidence and prevalence and examining the roles of external influences (for example, infection, diet) and interventions.
EQ-5D (EuroQol 5 dimensions)	A standardised instrument used to measure health-related quality of life. It provides a single index value for health status.
Evidence	Information on which a decision or guidance is based. Evidence is obtained from a range of sources including randomised controlled trials,

	observational studies, expert opinion (of clinical professionals or patients).
Exclusion criteria (literature review)	Explicit standards used to decide which studies should be excluded from consideration as potential sources of evidence.
Exclusion criteria (clinical study)	Criteria that define who is not eligible to participate in a clinical study.
Extended dominance	If Option A is both more clinically effective than Option B and has a lower cost per unit of effect, when both are compared with a do-nothing alternative then Option A is said to have extended dominance over Option B. Option A is therefore more cost-effective and should be preferred, other things remaining equal.
Extrapolation	An assumption that the results of studies of a specific population will also hold true for another population with similar characteristics.
Follow-up	Observation over a period of time of an individual, group or initially defined population whose appropriate characteristics have been assessed in order to observe changes in health status or health-related variables.
Generalisability	The extent to which the results of a study hold true for groups that did not participate in the research. See also external validity.
Gold standard	A method, procedure or measurement that is widely accepted as being the best available to test for or treat a disease.
GRADE, GRADE profile	A system developed by the GRADE Working Group to address the shortcomings of present grading systems in healthcare. The GRADE system uses a common, sensible and transparent approach to grading the quality of evidence. The results of applying the GRADE system to clinical trial data are displayed in a table known as a GRADE profile.
Harms	Adverse effects of an intervention.
Health economics	Study or analysis of the cost of using and distributing healthcare resources.
Health-related quality of life (HRQoL)	A measure of the effects of an illness to see how it affects someone's day-to-day life.
Heterogeneity or Lack of homogeneity	The term is used in meta-analyses and systematic reviews to describe when the results of a test or treatment (or estimates of its effect) differ significantly in different studies. Such differences may occur as a result of differences in the populations studied, the outcome measures used or because of different definitions of the variables involved. It is the opposite of homogeneity.
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of effect.
Inclusion criteria (literature review)	Explicit criteria used to decide which studies should be considered as potential sources of evidence.
Incremental analysis	The analysis of additional costs and additional clinical outcomes with different interventions.
Incremental cost	The extra cost linked to using one test or treatment rather than another. Or the additional cost of doing a test or providing a treatment more frequently.
Incremental cost-effectiveness ratio (ICER)	The difference in the mean costs in the population of interest divided by the differences in the mean outcomes in the population of interest for one treatment compared with another.
Incremental net benefit (INB)	The value (usually in monetary terms) of an intervention net of its cost compared with a comparator intervention. The INB can be calculated for a given cost-effectiveness (willingness to pay) threshold. If the threshold is £20,000 per QALY gained then the INB is calculated as: (£20,000 x QALYs gained) – Incremental cost.
Indirectness	The available evidence is different to the review question being addressed,

	in terms of PICO (population, intervention, comparison and outcome).
Intention-to-treat analysis (ITT)	An assessment of the people taking part in a clinical trial, based on the group they were initially (and randomly) allocated to. This is regardless of whether or not they dropped out, fully complied with the treatment or switched to an alternative treatment. Intention-to-treat analyses are often used to assess clinical effectiveness because they mirror actual practice: that is, not everyone complies with treatment and the treatment people receive may be changed according to how they respond to it.
Intervention	In medical terms this could be a drug treatment, surgical procedure, diagnostic or psychological therapy. Examples of public health interventions could include action to help someone to be physically active or to eat a more healthy diet.
Length of stay	The total number of days a participant stays in hospital.
Licence	See 'Product licence'.
Life years gained	Mean average years of life gained per person as a result of the intervention compared with an alternative intervention.
Likelihood ratio	The likelihood ratio combines information about the sensitivity and specificity. It tells you how much a positive or negative result changes the likelihood that a patient would have the disease. The likelihood ratio of a positive test result (LR+) is sensitivity divided by (1 minus specificity).
Long-term care	Residential care in a home that may include skilled nursing care and help with everyday activities. This includes nursing homes and residential homes.
Markov model	A method for estimating long-term costs and effects for recurrent or chronic conditions, based on health states and the probability of transition between them within a given time period (cycle).
Meta-analysis	A method often used in systematic reviews. Results from several studies of the same test or treatment are combined to estimate the overall effect of the treatment.
Negative predictive value (NPV)	In screening or diagnostic tests: A measure of the usefulness of a screening or diagnostic test. It is the proportion of those with a negative test result who do not have the disease, and can be interpreted as the probability that a negative test result is correct. It is calculated as follows: NPV = TN/(TN + FN)
Observational study	Individuals or groups are observed or certain factors are measured. No attempt is made to affect the outcome. For example, an observational study of a disease or treatment would allow 'nature' or usual medical care to take its course. Changes or differences in one characteristic (for example, whether or not people received a specific treatment or intervention) are studied without intervening. There is a greater risk of selection bias than in experimental studies.
Odds ratio	Odds are a way to represent how likely it is that something will happen (the probability). An odds ratio compares the probability of something in one group with the probability of the same thing in another. An odds ratio of 1 between 2 groups would show that the probability of the event (for example a person developing a disease, or a treatment working) is the same for both. An odds ratio greater than 1 means the event is more likely in the first group. An odds ratio less than 1 means that the event is less likely in the first group. Sometimes probability can be compared across more than 2 groups - in this case, one of the groups is chosen as the 'reference category', and the odds ratio is calculated for each group compared with the reference category. For example, to compare the risk of dying from lung cancer for non-smokers, occasional smokers and regular smokers, non-smokers could be used as the reference category. Odds ratios would be worked out for occasional

	smokers compared with non-smokers and for regular smokers compared with non-smokers. See also confidence interval, relative risk, risk ratio.
Opportunity cost	The loss of other health care programmes displaced by investment in or introduction of another intervention. This may be best measured by the health benefits that could have been achieved had the money been spent on the next best alternative healthcare intervention.
Outcome	The impact that a test, treatment, policy, programme or other intervention has on a person, group or population. Outcomes from interventions to improve the public's health could include changes in knowledge and behaviour related to health, societal changes (for example, a reduction in crime rates) and a change in people's health and wellbeing or health status. In clinical terms, outcomes could include the number of patients who fully recover from an illness or the number of hospital admissions, and an improvement or deterioration in someone's health, functional ability, symptoms or situation. Researchers should decide what outcomes to measure before a study begins.
P value	The p value is a statistical measure that indicates whether or not an effect is statistically significant. For example, if a study comparing 2 treatments found that one seems more effective than the other, the p value is the probability of obtaining these results by chance. By convention, if the p value is below 0.05 (that is, there is less than a 5% probability that the results occurred by chance) it is considered that there probably is a real difference between treatments. If the p value is 0.001 or less (less than a 1% probability that the results occurred by chance), the result is seen as highly significant. If the p value shows that there is likely to be a difference between treatments, the confidence interval describes how big the difference in effect might be.
Placebo	A fake (or dummy) treatment given to participants in the control group of a clinical trial. It is indistinguishable from the actual treatment (which is given to participants in the experimental group). The aim is to determine what effect the experimental treatment has had - over and above any placebo effect caused because someone has received (or thinks they have received) care or attention.
Polypharmacy	The use or prescription of multiple medications.
Positive predictive value (PPV)	In screening or diagnostic tests: A measure of the usefulness of a screening or diagnostic test. It is the proportion of those with a positive test result who have the disease, and can be interpreted as the probability that a positive test result is correct. It is calculated as follows: PPV = TP/(TP + FP)
Power (statistical)	The ability to demonstrate an association when one exists. Power is related to sample size; the larger the sample size, the greater the power and the lower the risk that a possible association could be missed.
Preoperative	The period before surgery commences.
Primary care	Healthcare delivered outside hospitals. It includes a range of services provided by GPs, nurses, health visitors, midwives and other healthcare professionals and allied health professionals such as dentists, pharmacists and opticians.
Primary outcome	The outcome of greatest importance, usually the one in a study that the power calculation is based on.
Product licence	An authorisation from the MHRA to market a medicinal product.
Prognosis	A probable course or outcome of a disease. Prognostic factors are patient or disease characteristics that influence the course. Good prognosis is associated with low rate of undesirable outcomes; poor prognosis is associated with a high rate of undesirable outcomes.

Prospective study	A research study in which the health or other characteristic of participants is monitored (or 'followed up') for a period of time, with events recorded as they happen. This contrasts with retrospective studies.
Publication bias	Publication bias occurs when researchers publish the results of studies showing that a treatment works well and don't publish those showing it did not have any effect. If this happens, analysis of the published results will not give an accurate idea of how well the treatment works. This type of bias can be assessed by a funnel plot.
Quality of life	See 'Health-related quality of life'.
Quality-adjusted life year (QALY)	A measure of the state of health of a person or group in which the benefits, in terms of length of life, are adjusted to reflect the quality of life. One QALY is equal to 1 year of life in perfect health. QALYS are calculated by estimating the years of life remaining for a patient following a particular treatment or intervention and weighting each year with a quality of life score (on a scale of 0 to 1). It is often measured in terms of the person's ability to perform the activities of daily life, freedom from pain and mental disturbance.
Randomisation	Assigning participants in a research study to different groups without taking any similarities or differences between them into account. For example, it could involve using a random numbers table or a computer-generated random sequence. It means that each individual (or each group in the case of cluster randomisation) has the same chance of receiving each intervention.
Randomised controlled trial (RCT)	A study in which a number of similar people are randomly assigned to 2 (or more) groups to test a specific drug or treatment. One group (the experimental group) receives the treatment being tested, the other (the comparison or control group) receives an alternative treatment, a dummy treatment (placebo) or no treatment at all. The groups are followed up to see how effective the experimental treatment was. Outcomes are measured at specific times and any difference in response between the groups is assessed statistically. This method is also used to reduce bias.
RCT	See 'Randomised controlled trial'.
Receiver operated characteristic (ROC) curve	A graphical method of assessing the accuracy of a diagnostic test. Sensitivity is plotted against 1 minus specificity. A perfect test will have a positive, vertical linear slope starting at the origin. A good test will be somewhere close to this ideal.
Reference standard	The test that is considered to be the best available method to establish the presence or absence of the outcome – this may not be the one that is routinely used in practice.
Relative risk (RR)	The ratio of the risk of disease or death among those exposed to certain conditions compared with the risk for those who are not exposed to the same conditions (for example, the risk of people who smoke getting lung cancer compared with the risk for people who do not smoke). If both groups face the same level of risk, the relative risk is 1. If the first group had a relative risk of 2, subjects in that group would be twice as likely to have the event happen. A relative risk of less than one means the outcome is less likely in the first group. Relative risk is sometimes referred to as risk ratio.
Resource implication	The likely impact in terms of finance, workforce or other NHS resources.
Retrospective study	A research study that focuses on the past and present. The study examines past exposure to suspected risk factors for the disease or condition. Unlike prospective studies, it does not cover events that occur after the study group is selected.
Review question	In guideline development, this term refers to the questions about treatment

	and care that are formulated to guide the development of evidence-based recommendations.
Secondary outcome	An outcome used to evaluate additional effects of the intervention deemed a priori as being less important than the primary outcomes.
Selection bias	Selection bias occurs if: a) The characteristics of the people selected for a study differ from the wider population from which they have been drawn, or b) There are differences between groups of participants in a study in terms of how likely they are to get better.
Sensitivity	How well a test detects the thing it is testing for. If a diagnostic test for a disease has high sensitivity, it is likely to pick up all cases of the disease in people who have it (that is, give a 'true positive' result). But if a test is too sensitive it will sometimes also give a positive result in people who don't have the disease (that is, give a 'false positive'). For example, if a test were developed to detect if a woman is 6 months pregnant, a very sensitive test would detect everyone who was 6 months pregnant, but would probably also include those who are 5 and 7 months pregnant. If the same test were more specific (sometimes referred to as having higher specificity), it would detect only those who are 6 months pregnant, and someone who was 5 months pregnant would get a negative result (a 'true negative'). But it would probably also miss some people who were 6 months pregnant (that is, give a 'false negative'). Breast screening is a 'real-life' example. The number of women who are recalled for a second breast screening test is relatively high because the test is very sensitive. If it were made more specific, people who don't have the disease would be less likely to be called back for a second test but more women who have the disease would be missed.
Sensitivity analysis	A means of representing uncertainty in the results of economic evaluations. Uncertainty may arise from missing data, imprecise estimates or methodological controversy. Sensitivity analysis also allows for exploring the generalisability of results to other settings. The analysis is repeated using different assumptions to examine the effect on the results. One-way simple sensitivity analysis (univariate analysis): each parameter is varied individually in order to isolate the consequences of each parameter on the results of the study. Multi-way simple sensitivity analysis (scenario analysis): 2 or more parameters are varied at the same time and the overall effect on the results is evaluated. Threshold sensitivity analysis: the critical value of parameters above or below which the conclusions of the study will change are identified. Probabilistic sensitivity analysis: probability distributions are assigned to the uncertain parameters and are incorporated into evaluation models based on decision analytical techniques (for example, Monte Carlo simulation).
Significance (statistical)	A result is deemed statistically significant if the probability of the result occurring by chance is less than 1 in 20 ($p<0.05$).
Specificity	The proportion of true negatives that are correctly identified as such. For example in diagnostic testing the specificity is the proportion of non-cases correctly diagnosed as non-cases. See related term 'Sensitivity'. In terms of literature searching a highly specific search is generally narrow and aimed at picking up the key papers in a field and avoiding a wide range of papers.
Stakeholder	An organisation with an interest in a topic that NICE is developing a clinical

	guideline or piece of public health guidance on. Organisations that register as stakeholders can comment on the draft scope and the draft guidance. Stakeholders may be: • manufacturers of drugs or equipment • national patient and carer organisations • NHS organisations • organisations representing healthcare professionals.
Systematic review	A review in which evidence from scientific studies has been identified, appraised and synthesised in a methodical way according to predetermined criteria. It may include a meta-analysis.
Time horizon	The time span over which costs and health outcomes are considered in a decision analysis or economic evaluation.
Treatment allocation	Assigning a participant to a particular arm of a trial.
Univariate	Analysis which separately explores each variable in a data set.
Utility	In health economics, a 'utility' is the measure of the preference or value that an individual or society places upon a particular health state. It is generally a number between 0 (representing death) and 1 (perfect health). The most widely used measure of benefit in cost-utility analysis is the quality-adjusted life year, but other measures include disability-adjusted life years (DALYs) and healthy year equivalents (HYEs).

NB The NICE abbreviations and glossary can be found on the NICE website.