National Institute for Health and Care Excellence

Preoperative tests (update)

Routine preoperative tests for elective surgery

Clinical guideline NG45

Methods, evidence and recommendations

April 2016

Developed by the National Guideline Centre,
hosted by the Royal College of Physicians

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, where appropriate, their guardian or carer.

Copyright
National Institute for Health and Care Excellence 2016

Contents

Guideline Development Group members

Name	Role
Ian Smith (Chair)	Senior Lecturer in Anaesthesia, Keele University; Consultant Anaesthetist, University Hospitals of North Midlands
Liz Clark	Patient member
Martin Crook	Consultant in Clinical Biochemistry and Metabolic Medicine, Guy's and St Thomas' Hospital Foundation Trust and Lewisham and Greenwich NHS Trust; Visiting Professor University of Greenwich and Trinity Dublin
Jugdeep Dhesi	Consultant Physician in Geriatrics and General Internal Medicine & Clinical Lead Proactive Care for Older People undergoing Surgery (POPS), Guy's and St Thomas' Hospital Foundation Trust
Luke Howard	Consultant Respiratory Physician, Hammersmith Hospital, Imperial College Healthcare NHS Trust
Nicholas Levy	Consultant Anaesthetist, West Suffolk Hospital
Audrai Morgan	Pre Assessment Lead Nurse. BMI The Clementine Churchill Hospital
Hasan Mukhtar	Consultant General and Colorectal Surgeon & Honorary Senior Lecturer, Whittington Health
Burra Murthy	Consultant Anaesthetist, Royal Liverpool and Broadgreen University Hospital; Honorary Professor of Anaesthesia, Liverpool John Moores University
Amanda Perrill	Ward Sister, Royal Berkshire NHS Foundation Trust
Marco Scarci	Consultant Thoracic Surgeon, Papworth Hospital
Ann Seymour	Patient member

NCGC technical team members

Name	Role
Katie Broomfield	Document Editor/Process Assistant (from February 2015)
Eleanor Carter	Specialist trainee advisor
Saskia Cheyne	Research Fellow (until April 2015)
Dalia Dawoud	Health Economist (from May to September 2014)
Katharina Dworzynski	Senior Research Fellow (September 2014 to February 2015)
Lindsay Dytham	Document Editor/Process Assistant (until February 2015)
Elisabetta Fenu	Health economic lead (until November 2014)
Alex Haines	Senior Health Economist (from September 2014)
Yasmina Hedhli	Project Manager (until June 2015)
Frank O'Neill	Senior Research Fellow (from February 2015)
Elizabeth Pearton	Information Scientist (until March 2015)
Natalie Pink	Project Manager (from June 2015)
Emma Pottinger	Research Fellow (from June 2015)
Gill Ritchie	Associate Director and Guideline Lead
Elizabeth Rouncefield	Information Scientist (until April 2015)
Peter Rouse	Health Economist (from February 2015)
Eleanor Samarasekera	Research Fellow (from April 2015)
Grammati Sarri	Senior Research Fellow (until October 2014)

Name	Role
Ruth Wong	Information Scientist (from June 2015)

Co-optees

Name	Role
Abdallah Al-Mohammad	Consultant Cardiologist and Honorary Senior Clinical Lecturer, Sheffield Teaching Hospitals NHS Foundation Trust
Eshan Senanayake	General Practitioner

Acknowledgements

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

Sara Carrillo de Albornoz, Jill Cobb, Tamara Diaz, Jessica Glen, Lina Gulhane, Katie Jones, Kate Kelley, Bethany King, Mark Perry, Joshua Ruegger, Carlos Sharpin.

1 Guideline summary

1.1 Full list of recommendations

Communication

1. When offering tests before surgery, give people information in line with recommendations (including those on consent and capacity) made in the NICE guideline on patient experience in adult NHS services.
2. Ensure that the results of any preoperative tests undertaken in primary care are included when referring people for surgical consultation.

Considering existing medicines

3. Take into account any medicines people are taking when considering whether to offer any preoperative test.

Resting ECG

4. Resting ECG

	Surgery grade		
ASA grade	**Minor**	**Intermediate**	**Major or complex**
ASA 1	**Do not routinely offer**	**Do not routinely offer**	**Consider for people aged over 65 if no ECG results available from past 12 months**
ASA 2	**Do not routinely offer**	**Consider for people with cardiovascular, renal or diabetes comorbidities**	**Offer**
ASA3 or ASA4	**Consider if no ECG results available from past 12 months**	**Offer**	**Offer**

Resting echocardiography

5. Do not routinely offer resting echocardiography before surgery.
6. Consider resting echocardiography if the person has:
   - a heart murmur and any cardiac symptom (including breathlessness, pre-syncope, syncope or chest pain) or
   - signs or symptoms of heart failure.

   Before ordering the resting echocardiogram, carry out a resting electrocardiogram (ECG) and discuss the findings with an anaesthetist.

Chest X-ray

7. Do not routinely offer chest X-rays before surgery.

Lung function tests and arterial blood gas analysis

8. Do not routinely offer lung function tests or arterial blood gas analysis before surgery.
9. Consider seeking advice from a senior anaesthetist as soon as possible after assessment for people who:
   - are ASA grade 3 or 4 due to known or suspected respiratory disease and
   - are having intermediate or major or complex surgery.

Full blood count test

10. Full blood count test

	Surgery grade		
ASA grade	**Minor**	**Intermediate**	**Major or complex**
ASA 1	**Do not routinely offer**	**Do not routinely offer**	**Offer**
ASA2	**Do not routinely offer**	**Do not routinely offer**	**Offer**
ASA3 or ASA4	**Do not routinely offer**	**Consider for people with cardiovascular or renal disease if any symptoms not recently investigated**	**Offer**

Kidney function tests

11. Kidney function tests

	Surgery grade		
ASA grade	**Minor**	**Intermediate**	**Major or complex**
ASA 1	**Do not routinely offer**	**Do not routinely offer**	**Consider in people at risk of AKI[a]**
ASA2	**Do not routinely offer**	**Consider in people at risk of AKI[b]**	**Offer**
ASA3 or ASA4	**Consider in people at risk of AKI[c]**	**Offer**	**Offer**

Haemostasis tests

12. Do not routinely offer haemostasis tests before surgery.

13. Consider haemostasis tests in people with chronic liver disease having intermediate or major or complex surgery.

    - If people taking anticoagulants need modification of their treatment regimen, make an individualised plan in line with local guidance.
    - If clotting status needs to be tested before surgery (depending on local guidance) use point-of-care testing.[d]

Glycated haemoglobin (HbA1c) test in people with diagnosed diabetes

14. People with diabetes who are being referred for surgical consultation from primary care should have their most recent HbA1c test results included in their referral information.

15. Offer HbA1c testing to people with diabetes having surgery if they have not been tested in the last 3 months.

Glycated haemoglobin (HbA1c) test in people without diagnosed diabetes

16. Do not routinely offer HbA1c testing before surgery to people without diagnosed diabetes.

Sickle cell disease or sickle celltrait tests

17. Do not routinely offer testing for sickle cell disease or sickle cell trait before surgery.

a See recommendation 1.1.8 in NICE CG169: Acute kidney injury: prevention, detection and management (2013)
b See recommendation 1.1.8 in NICE CG169: Acute kidney injury: prevention, detection and management (2013)
c See recommendation 1.1.8 in NICE CG169: Acute kidney injury: prevention, detection and management (2013)
d Note that currently the effects of direct oral anticoagulants (DOACs) cannot be measured by routine testing.

18. Ask the person having surgery if they or any member of their family have sickle cell disease.
19. If the person is known to have sickle cell disease and has their disease managed by a specialist sickle cell service, liaise with this team before surgery.

Urinalysis

20. Do not routinely offer urine dipstick tests before surgery.
21. Consider microscopy and culture of midstream urine sample before surgery if the presence of a urinary tract infection would influence the decision to operate.

Pregnancy testing

22. On the day of surgery, sensitively ask all women of childbearing potential whether there is any possibility they could be pregnant.
23. Make sure women who could possibly be pregnant are aware of the risks of the anaesthetic and the procedure to the fetus.
24. Document all discussions with women about whether or not to carry out a pregnancy test.
25. Carry out a pregnancy test with the woman's consent if there is any doubt about whether she could be pregnant.
26. Develop locally agreed protocols for checking pregnancy status before surgery.
27. Make sure protocols are documented and audited, and in line with statutory and professional guidance.

1.2 Key research recommendations

Polysomnography

1. Does preoperative screening of people who are at risk of obstructive sleep apnoea (OSA) with polysomnography identify those at higher risk of postoperative complications?
2. Does treating OSA perioperatively improve outcomes?

Glycated haemoglobin (HbA1c) test

3. Does optimisation of HbA1c in people with poorly controlled diabetes improve surgical outcomes?

1.3 How this clinical guideline was updated

This guideline is a complete update of NICE clinical guideline 3: 'Preoperative tests: The use of routine preoperative tests for elective surgery'.[71] This update replaces the 2003 guideline, which will be stood down.

The update was commissioned to include and update the results of the 2012 Heath Technology Assessment 'What is the value of routinely testing full blood count, urea and electrolytes, and lung function tests before elective surgery in patients with no apparent clinical indication and in

subgroups of patients with common comorbidities: a systematic review of the clinical and cost-effective literature'.[24]

In the areas where new evidence was identified as part of the NICE review update, full searches were undertaken and systematic reviews conducted. No systematic reviews were undertaken for areas where the NICE review update found no new evidence. Formal consensus methods, in the form of a modified Delphi survey, were used in addition to the updated evidence reviews to support the development of recommendations, and for those areas where no evidence review was to be conducted.

Clinical tests not in the original guideline that were included in the update are:

- Cardiopulmonary exercise test (CPET)
- Resting echocardiography
- Polysomnography
- Glycated haemoglobin (HbA1c)

Random blood glucose was included in the original guideline, but removed from the update and replaced by the glycated haemoglobin (HbA1c) test. This reflects the change in current practice.

Children, pregnant women and people having cardiothoracic or neurosurgery were excluded from the scope of the update.

All recommendations in the original guideline have been replaced by the recommendations in this update.

- The original guideline made recommendations for specific age bands, whereas the update considers ASA grade only (with an assumption that age is indirectly reflected within the ASA grade)
- The original guideline made individual recommendations for single comorbidities, whereas the update combines all comorbidities within recommendations
- The update does not use surgery grades as in the original guideline, but instead uses the terms 'minor', 'intermediate' and 'major or complex' to refer to increasing complexity of surgical procedures
- Where tests are recommended for consideration, the update provides clarification on the populations or circumstances in which doing the test would have value.

2 Introduction

In 2003, NICE first issued guidance on the use of routine preoperative tests for elective surgery (NICE CG3).[71] The guideline evaluated the practice of routinely performing preoperative diagnostic tests for elective surgery in healthy and comorbid populations.

Much of the evidence in the original guideline was inconclusive and a formal consensus survey about the appropriateness of preoperative testing was conducted to inform the recommendations made by the Guideline Development Group (GDG). Since the guideline was issued there has been a reduction in the ordering of routine tests for young, healthy patients undergoing minor surgery,[24] however there remains a concern that unnecessary tests continue to be requested.

Excessive testing can cause significant anxiety in patients, delays in treatment and unnecessary, costly, and possibly harmful treatments.[9,59] Moreover, even genuinely abnormal results often do not result in any significant change in perioperative management in relatively healthy patients.[69,85]

In 2012/13 the NHS in England completed 10.6 million operations compared with 6.61 million in 2002/03: an increase of 60%. Therefore even a small percentage of unnecessary testing can affect large numbers of patients.

Preoperative tests provide a benefit where they yield additional information that cannot be obtained from a patient history and physical examination alone, and also where they:

- help to assess the risk to the patient and inform discussions about the risks and benefits of surgery
- allow the patient's clinical management to be altered, if necessary, in order to reduce possible harm or increase the benefit of surgery
- help to predict postoperative complications
- establish a baseline measurement for later reference where potentially abnormal postoperative test results cannot be adequately interpreted in isolation.[48]

Since the original NICE guideline was issued in 2003, preoperative assessment has changed radically. In the past preoperative tests were requested by junior medical staff in anticipation of, and readiness for, an assessment by an anaesthetist shortly before surgery. Currently most patients are seen well in advance of surgery in a preoperative assessment clinic, where a structured history and targeted examination are performed by experienced nursing staff according to protocols developed by anaesthetists.[2,18] Early preoperative assessment by nurses can determine the patient's functional status, which remains a major determinant of perioperative risk,[19] and has been shown to reduce the number of investigations which are requested.[58]

In light of these developments, it is clear that an update to the guideline is required. However a review of newly published evidence highlighted the paucity of high quality studies evaluating the benefit of routine preoperative testing in adults undergoing elective non-cardiac surgery. For this reason a modified Delphi consensus survey was undertaken to re-evaluate the usage of routine preoperative tests amongst clinicians, which helped the GDG update and revise the recommendations made in 2003.

A number of other developments have occurred since 2003 and are reflected in the scope of this update. Random blood glucose has been largely abandoned in the detection and optimisation of diabetes mellitus and replaced by glycated haemoglobin (HbA1c).[1] Several new preoperative tests are increasingly used in patients undergoing elective surgery (for example non-invasive cardiac stress tests, cardiopulmonary exercise testing, polysomnography). It is hoped that these tests may provide more information on the best form of perioperative management and assist the prediction of postoperative complications in certain higher risk patients.

In a final change from the 2003 guideline, children, and patients undergoing cardiothoracic procedures or neurosurgery, are populations not covered by this update because their management is highly specialised and specialist guidance exists elsewhere.

3 Development of the guideline

3.1 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 healthcare. 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 patient and health professional.

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 Clinical Guideline Centre (NCGC).
- The NCGC establishes a Guideline Development Group.
- 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 NCGC 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.

3.2 Remit

NICE received the remit for this guideline from the Department of Health. They commissioned the National Clinical Guideline Centre (NCGC) to produce the guideline.

This guideline is a complete update of NICE clinical guideline 3: 'Preoperative tests: The use of routine preoperative tests for elective surgery'.[71] This update replaces the 2003 guideline, which will be stood down.

3.3 Who developed this guideline?

A multidisciplinary Guideline Development Group (GDG) comprising health professionals and lay member, supported by health service researchers from the National Clinical Guideline Centre (NCGC), developed this guideline (see the list of GDG members and the acknowledgements).

The National Institute for Health and Care Excellence (NICE) funds the NCGC and thus supported the development of this guideline. The GDG was convened by the NCGC and chaired by Ian Smith in accordance with guidance from NICE.

The group met every 6 weeks during the development of the guideline, except for a period of 3 months while the Delphi survey was conducted. At the start of the guideline development process all GDG members declared interests including consultancies, fee-paid work, share-holdings, fellowships and support from the healthcare industry. At all subsequent GDG 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 NCGC 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 GDG.

3.3.1 What this guideline covers

The population covered by this guideline is adults and young people (over 16 years old) who are ASA grade 1 to 4 (see Table 1); who may have one or more of the following comorbidities: cardiovascular, respiratory, renal, diabetes or obesity; and who are having minor, intermediate, or major or complex elective surgery.

3.3.1.1 ASA grades

The American Society of Anesthesiologists (ASA) Physical Status Classification System is a simple scale describing fitness to undergo an anaesthetic. The ASA states that it does not endorse any elaboration of these definitions. However, anaesthetists in the UK often qualify (or interpret) these grades as relating to functional capacity – that is, comorbidity that does not (ASA grade 2) or does (ASA grade 3) limit a person's activity. Table 1 sets out the ASA grades used in this guideline (adapted from the ASA website).[8]

Table 1: ASA Physical Status Classification System

ASA grade	Explanation
ASA grade 1	A normal healthy person
ASA grade 2	A person with mild systemic disease
ASA grade 3	A person with severe systemic disease
ASA grade 4	A person with severe systemic disease that is a constant threat to life

3.3.1.2 Surgery grades

An operation represents a physiological stress. The magnitude of the physiological stress increases with the 'invasiveness' of the procedure. There is no widely accepted and validated system for

classifying the stressfulness of operative procedures, so this guideline adopted a simple scale, illustrated with examples.

Table 2: Surgery grades and examples

Minor	Intermediate	Major or complex
• excising skin lesion • draining breast abscess	• primary repair of inguinal hernia • excising varicose veins in the leg • tonsillectomy or adenotonsillectomy • knee arthroscopy	• total abdominal hysterectomy • endoscopic resection of prostate • lumbar discectomy • thyroidectomy • total joint replacement • lung operations • colonic resection • radical neck dissection

3.3.1.3 Tests

This guideline covers the following routine preoperative clinical tests:

- Full blood count (haemoglobin, white blood cell count and platelet count)
- Kidney function (urea, estimated glomerular filtration rate and electrolyte tests)
- Lung function and arterial blood gas analysis
- Resting electrocardiography (ECG)
- Cardiopulmonary exercise test (CPET)
- Resting echocardiography
- Polysomnography (to detect obstructive sleep apnoea [OSA])
- Glycated haemoglobin (HbA1c)
- Haemostasis tests
- Chest X-ray
- Urinalysis
- Pregnancy testing
- Sickle cell disease/trait tests.

For further details please refer to the scope in Appendix A and the review questions in Section 4.1.

3.3.2 What this guideline does not cover

Populations not covered by this guideline are:

- Children and young people (0 to 16 years)
- Pregnant women
- Adults who are ASA2 or above, with comorbidities other than cardiovascular, respiratory, renal, diabetes or obesity
- People having cardiothoracic or neurosurgery.

This guideline does not cover the following clinical areas:

- Random blood glucose tests

- Computed tomography scan of the thorax
- Haemoglobin electrophoresis
- Blood cross-matching
- Screening tests for methicillin-resistant Staphylococcus aureus (MRSA), Clostridium difficile (C.Diff), vancomycin-resistant enterococci (VRE), carbapenem-resistant Enterobacteriacaea (CRE), carbapenem-resistant Klebsiella pneumoniae (CRKP) and other hospital-acquired 'superbug'infections
- Preoperative clinical assessment (including taking a medical history, physical examination and advice on the assessment and wider clinical management of people's conditions before surgery or during follow-up) and the optimal setting for preoperative testing.

3.3.3 Relationships between the guideline and other NICE guidance

Related NICE clinical guidelines: Patient experience in adult NHS services: improving the experience of care for people using adult NHS services. NICE clinical guideline 138[72]

3.4 General principles

3.4.1 Communication

3.4.1.1 Patient consent

Recommendation 1.5.3 from the 2003 guideline highlighted the issue of consent to undergo preoperative tests, referring to the 'Good Practice in Consent' guidance from the Department of Health in 2001.[26] Specifically, the 2003 guideline indicated that:

'Patients should have access to sufficient information about risks, benefits and alternatives to be able to make an informed decision about whether to consent.'

The scope of the update did not include patient consent, but the GDG did discuss the principles of the consent process in relation to preoperative testing. In particular, the GDG noted the recommendations in NICE clinical guideline 138: Patient experience in adult NHS services[72] on obtaining consent, and wished to cross-refer to Section 1.4: Continuity of care and relationships, and Section 1.5: Information and shared decision-making. These sections provide guidance on discussing risks and benefits of investigations with the patient, and providing timely information in an accessible format that enables the patient to make informed decisions. The GDG specifically highlighted the following recommendations from CG138 in relation to preoperative testing:

> **1.2.12** Obtain and document informed consent from the patient, in accordance with:
>
> - in England, Department of Health policy and guidance
> - in Wales, advice from the Welsh Government.
>
> **1.2.13** Assess the patient's capacity to make each decision using the principles in the Mental Capacity Act (2005).[68]

The following recommendation was made for this guideline:

1. **When offering tests before surgery, give people information in line with recommendations (including those on consent and capacity) made in the NICE guideline on patient experience in adult NHS services.**

Guidance on consent for young people aged 16–17 is available from the Reference Guide to Consent for Examination or Treatment (second edition).[27]

3.4.1.2 Communicating test results

The GDG made an overarching recommendation to be considered across all preoperative tests in this guideline:

2. **Ensure that the results of any preoperative tests undertaken in primary care are included when referring people for surgical consultation.**

This assists with the preoperative management of the patient and prevents unnecessary duplication of tests. See the 'recommendations and link to evidence' section for kidney function tests for details (section 12.10).

3.4.2 Considering existing medicines

Underpinning this guideline is a general principle that before surgery, a clinical assessment should take place that will inform decisions as to which preoperative tests are necessary. In particular, a patient's existing drug therapy should be taken into account when considering a preoperative test, as some medication may alter the results, such that perioperative management of the patient would need to change. The GDG therefore made an overarching recommendation to this effect:

3. **Take into account any medicines people are taking when considering whether to offer any preoperative test.**

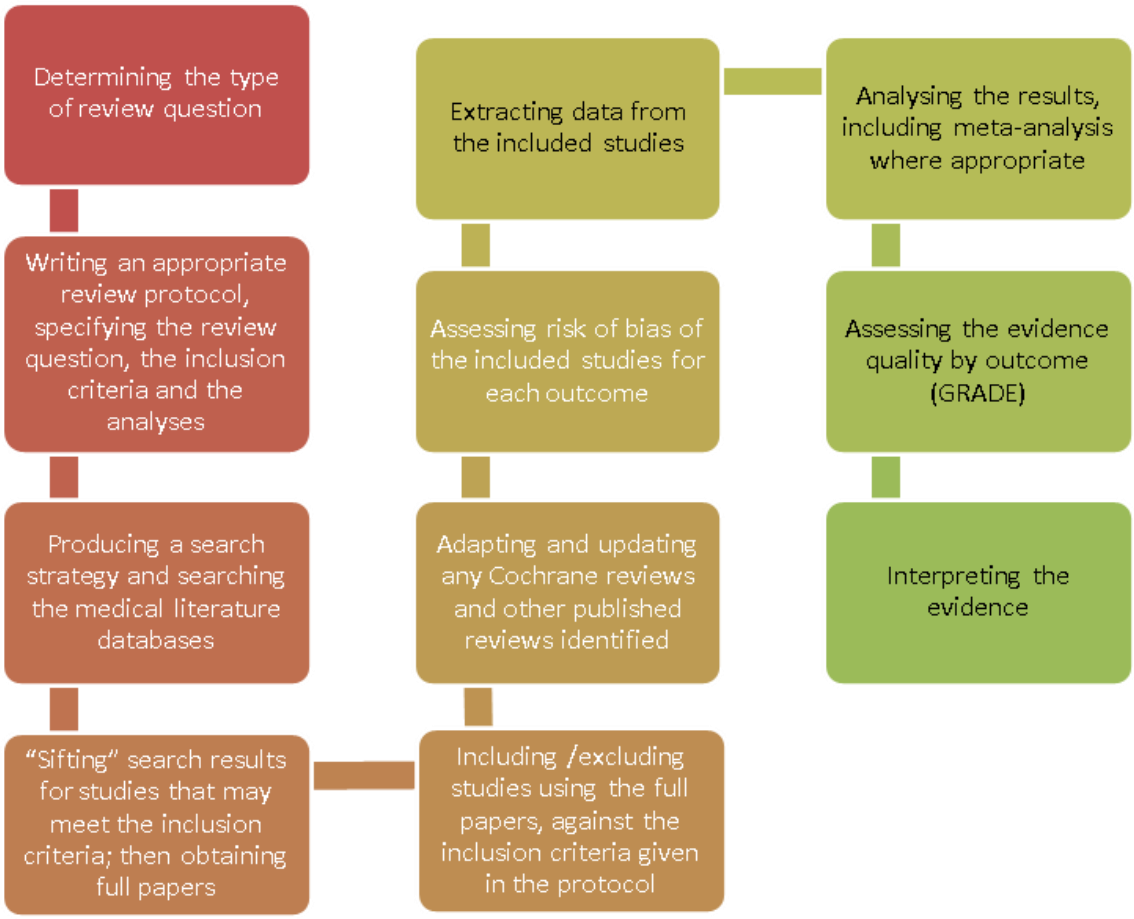
Clinical judgement should be used in all cases to consider individual circumstances and make decisions that are appropriate to the individual patient.

4 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 and 2014 versions.[77,78]

Sections 4.1 to 4.6 describe the process of reviewing clinical evidence (summarised in Figure 1) and section 4.7 the process of reviewing the cost-effectiveness evidence.

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



4.1 Developing the review questions and outcomes

Review questions were developed in a PICO framework (patient, intervention, comparison and outcome) for intervention reviews 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 GDG. The review questions were drafted by the NCGC technical team and refined and validated by the GDG. The questions were based on the key clinical areas identified in the scope (Appendix A).

A total of 17 review questions were identified.

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

Table 3: Review questions

Chapter	Type of review	Review questions	Outcomes
5	Intervention	What is the clinical- and cost-effectiveness of using resting electrocardiography (ECG) as a preoperative test in improving patient outcomes in adults and young people undergoing non-cardiac elective surgery?	**Critical:** • All-cause mortality • Health related quality of life **Important:** • Complications related to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney failure, infection) • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing • Intensive care unit (ICU) admission • Composite outcomes such as major adverse cardiovascular events (MACE) that incorporate cardiac deaths and non-fatal cardiac events • Optimisation of medical therapy
5	Prognostic	Does resting ECG predict prognosis (patient outcomes after surgery) in adults and young people undergoing non-cardiac elective surgery?	**Critical:** • All-cause mortality **Important:** • Complications relating to surgery or anaesthesia • Length of hospital stay • Hospital readmission • Adverse events caused by testing • Health-related quality of life • Intensive care unit (ICU) admission
6	Intervention	What is the usefulness of resting echocardiography as a preoperative test in altering perioperative management for adults and young people with mild to severe comorbidities undergoing major or complex elective surgery?	**Critical:** • Change in healthcare management (for example cancellation of surgery or treating ischaemia, valvular disease or heart failure on the basis of the results of the tests) **Important:** • All-cause mortality • Complications related to surgery or anaesthesia • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing (time of testing).

Chapter	Type of review	Review questions	Outcomes
			• Health-related quality of life • Intensive care unit (ICU) admission • Composite outcomes such as major adverse cardiovascular events (MACE) that incorporate cardiac deaths and non-fatal cardiac events • Optimisation of medical therapy
7	Intervention	What is the clinical- and cost-effectiveness of using cardiopulmonary exercise test (CPET) as a preoperative test in improving patient outcomes in adults and young people with mild to severe comorbidities undergoing major or complex non-cardiac elective surgery?	**Critical:** • All-cause mortality • Health-related quality of life **Important:** • Complications related to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney failure, infection) • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing • Intensive care unit (ICU) admission
7	Prognostic	Does cardiopulmonary exercise testing (CPET) predict prognosis (patient outcomes after surgery) in adults and young people with mild to severe comorbidities undergoing major or complex non-cardiac elective surgery?	**Critical:** • All-cause mortality **Important:** • Complications relating to surgery or anaesthesia • Length of hospital stay • Hospital readmission • Adverse events caused by testing • Health-related quality of life • Intensive care unit (ICU) admission
9	Intervention	What is the clinical and cost-effectiveness of using polysomnography as a preoperative test (to detect obstructive sleep apnoea) in improving patient outcomes in adults and young people with obesity undergoing major or complex elective non-cardiac surgery?	**Critical:** • All-cause mortality **Important:** • Complications related to surgery or anaesthesia • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing • Health-related quality of life • Intensive care unit (ICU) admission • Optimisation of therapy • Change in management
9	Prognostic	Does polysomnography predict prognosis (patient outcomes after surgery) in adults and young people with obesity undergoing major or complex elective non-	**Critical:** • All-cause mortality **Important:** • Complications relating to surgery or anaesthesia

Chapter	Type of review	Review questions	Outcomes
		cardiac surgery?	• Length of hospital stay (post-operation) • Hospital readmission • Adverse events after surgery (wound infection) • Health-related quality of life • Intensive care unit (ICU) admission
10	Intervention	What is the usefulness of lung function tests in predicting outcome or altering perioperative management for adults and young people undergoing any type of elective surgery?	**Critical:** • All-cause mortality **Important:** • Change in healthcare management (for example cancellation of surgery) • Complications relating to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney injury, infection) • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing • Health-related quality of life • Intensive care unit (ICU) admission
10	Prognostic	Do lung function tests (also including blood gas analysis) predict prognosis (patient outcomes after surgery) in adults and young people ASA 1–4 undergoing any type of elective non-cardiac surgery?	**Critical:** • All-cause mortality **Important:** • Complications relating to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney injury, infection) • Change in healthcare management (for example cancellation of surgery) • Length of hospital stay • Hospital readmission • Adverse events caused by testing (time of testing) • Health-related quality of life • Intensive care unit (ICU) admission
11	Intervention	What is the usefulness of full blood count (haemoglobin, white blood cell count and platelet count) in predicting outcome or altering perioperative management for adults and young people undergoing any type of elective non-cardiac surgery?	**Critical:** • All-cause mortality **Important:** • Change in healthcare management (for example cancellation of surgery) • Complications relating to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney injury, infection) • Length of hospital stay after an

Chapter	Type of review	Review questions	Outcomes
			operation • Hospital readmission • Adverse events caused by testing • Health-related quality of life • Intensive care unit (ICU) admission
11	Prognostic	Do full blood count tests (haemoglobin, white blood cell count and platelet count) predict prognosis (patient outcomes after surgery) in adults and young people ASA 1–4 undergoing any type of elective non-cardiac surgery?	**Critical:** • All-cause mortality **Important:** • Complications relating to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney injury, infection) • Change in healthcare management (for example cancellation of surgery) • Length of hospital stay • Hospital readmission • Adverse events caused by testing (time of testing) • Health-related quality of life • Intensive care unit (ICU) admission
12	Intervention	What is the usefulness of kidney function tests (urea, estimated glomerular filtration rate and electrolyte tests) in predicting outcome or altering perioperative management for adults and young people undergoing any type of elective non-cardiac surgery?	**Critical:** • All-cause mortality Important: • Change in healthcare management (for example cancellation of surgery) • Complications relating to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney injury, infection) • Length hospital stay after an operation • Hospital readmission • Adverse events caused by testing • Health-related quality of life • Intensive care unit (ICU) admission
12	Prognostic	Do kidney function tests (urea, estimated glomerular filtration rate and electrolyte tests) predict prognosis (patient outcomes after surgery) in adults and young people ASA 1–4 undergoing any type of elective non-cardiac surgery?	**Critical:** • All-cause mortality Important: • Complications relating to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney injury, infection) • Change in healthcare management (for example cancellation of surgery) • Length of hospital stay • Hospital readmission • Adverse events caused by testing

Chapter	Type of review	Review questions	Outcomes
			(time of testing) • Health-related quality of life • Intensive care unit (ICU) admission
14.1	Intervention	What is the clinical- and cost-effectiveness of using HbA1c (glycated haemoglobin) as a preoperative test in improving patient outcomes in adults and young people with diabetes and mild to severe comorbidities undergoing elective non-cardiac surgery?	**Critical:** • All-cause mortality • Health-related quality of life **Important:** • Complications related to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney failure, infection) • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing • Intensive care unit (ICU) admission
14.1	Prognostic	Does HbA1c (glycated haemoglobin) predict prognosis (patient outcomes after surgery) of adults and young people with diabetes (all types) and mild to severe comorbidities undergoing major or complex elective non-cardiac surgery?	**Critical:** • All-cause mortality **Important:** • Complications relating to surgery or anaesthesia • Length of hospital stay (post-operation) • Hospital readmission • Adverse events after surgery (wound infection) • Health-related quality of life • Intensive care unit (ICU) admission
14.2	Intervention	What is the clinical- and cost-effectiveness of using HbA1c (glycated haemoglobin) as a preoperative test in improving patient outcomes in adults and young people with mild to severe comorbidities undergoing elective non-cardiac surgery?	**Critical:** • All-cause mortality • Health-related quality of life **Important:** • Complications related to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney failure, infection) • Length of hospital stay after an operation • Hospital readmission • Intensive care unit (ICU) admission
14.2	Prognostic	Does HbA1c (glycated haemoglobin) predict prognosis (patient outcomes after surgery) of people with mild to severe comorbidities undergoing major or complex elective non-cardiac surgery?	**Critical:** • All-cause mortality • **Important:** • Complications relating to surgery or anaesthesia • Length of hospital stay (post-operation) • Hospital readmission

Chapter	Type of review	Review questions	Outcomes
			• Adverse events after surgery (wound infection) • Health-related quality of life • Intensive care unit (ICU) admission

4.1.1 Issues with guideline development

The GDG noted that the guideline population was large and cross-cutting, and that the recommendations only provide guidance for routine preoperative assessment. Specific clinical conditions were not considered and all recommendations must be interpreted with appropriate clinical experience. The GDG also anticipated a lack of high quality clinical evidence (RCTs or sufficiently large cohort studies) to inform the recommendations, and therefore decided to use prognostic data and formal consensus methods to identify areas of agreement on which to base recommendations.

4.1.2 Hierarchy of evidence

In the absence of high quality evidence the GDG developed a pragmatic process by which to make recommendations:

Order of preference for study designs:

Systematic reviews of randomised controlled trials that meet our PICOs

Randomised controlled trials (RCTs)

Where no RCTs are available, we will consider:

Abstracts of RCTs

Where no RCTs or abstracts of RCTs are available:

Non-randomised trials: prospective or retrospective cohort studies

Non-blinded, single and double-blinded trials will be included

Where no randomised or non-randomised evidence are available (when applicable):
Prognostic evidence

A formal consensus method and informal consensus and clinical experience of the GDG were used to inform all recommendations. The discussions are documented in the 'recommendations and link to evidence' section in each chapter.

4.2 Health technology assessment (HTA) update

We were commissioned to update the 2012 HTA of preoperative tests for full blood count (haemoglobin, white blood cell count and platelet count), kidney function tests (urea, estimated glomerular filtration rate and electrolyte tests) and lung function tests (also including blood gas analysis).[24] This included an update of evidence for adult patients who are ASA grade 1 or 2 (with cardiovascular, renal and respiratory diseases) undergoing minor or intermediate surgery from May 2009.

In addition to updating the 2012 HTA, we widened the patient population to include ASA grades 1–4, including people who are obese who were previously excluded, and all surgery types as specified in the scope of this guideline (Appendix A).

4.3 Delphi consensus survey

In the absence of a strong evidence base and clear guidance for clinical practice, a formal Delphi consensus technique was used to provide the GDG with a basis for decision-making.

The NCGC technical team, in partnership with the GDG, developed a survey based on the current recommendations made in the 2003 guideline, in order to reassess the consensus view amongst the relevant health professionals since 2003. The survey was a modified Delphi survey, which used an anonymous, multi-round, consensus-building technique. The results of the survey along with any economic evidence were considered by the GDG when drafting consensus-based recommendations for the updated guideline. The full methodology and results of the Delphi consensus survey are contained in Appendix L.

4.4 Summary of methodological approach for each preoperative test

Table 4 lists the areas covered by the guideline and the methodological approach taken for each test.

Table 4: Methodological approach to guideline areas

Preoperative test	Methods	Rationale
Resting electrocardiography (ECG)	• In the original 2003 guideline • Systematic review conducted • Included in modified Delphi survey	New evidence included in the update review was not sufficient to enable the GDG to revise the 2003 recommendations and therefore resting ECG was included in the consensus survey to re-evaluate clinical opinion.
Resting echocardiography	• New topic • Systematic review conducted • Not included in modified Delphi survey	Resting echocardiography was identified as a new topic for inclusion during the NICE update review. A systematic review was conducted and recommendations made.
Cardiopulmonary exercise test (CPET)	• New topic • Systematic review conducted • Not included in modified Delphi survey	CPET was identified as a new topic for inclusion during the NICE update review. A systematic review was conducted but the evidence was not sufficient to enable the GDG to make new recommendations until further research is available As this was a new area not included in the original guideline, it was not included in the consensus survey.
Chest X-ray	• In original 2003 guideline • No new evidence identified • Included in modified Delphi survey	No new evidence was identified during the NICE update review, but the opinion from stakeholders during the scoping of the update was that clinical opinion and practice had changed since the original guideline published in 2003. The test was included in the consensus survey to re-evaluate

Preoperative test	Methods	Rationale
		current opinion.
Polysomnography	• New topic • Systematic review conducted • Not included in modified Delphi survey	Polysomnography was identified as a new topic for inclusion during the NICE update review. A systematic review was conducted but the evidence was not sufficient to enable the GDG to make new recommendations until further research is available As this was a new area not included in the original guideline it was not included in the consensus survey.
Lung function tests and blood gas analysis	• In the original 2003 guideline • update of 2012 HTA systematic review conducted • Included in modified Delphi survey	New evidence included in the update review was not sufficient to enable the GDG to revise the 2003 recommendations and therefore lung function tests were included in the consensus survey to re-evaluate clinical opinion.
Full blood count test (haemoglobin, white blood cell count and platelet count)	• In the original 2003 guideline • Update of 2012 HTA systematic review conducted • Included in modified Delphi survey	New evidence included in the update review was not sufficient to enable the GDG to revise the 2003 recommendations and therefore full blood count was included in the consensus survey to re-evaluate clinical opinion.
Kidney function (urea, estimated glomerular filtration rate and electrolyte tests)	• In the original 2003 guideline • Update of 2012 HTA systematic review conducted • Included in modified Delphi survey	New evidence included in the update review was not sufficient to enable the GDG to revise the 2003 recommendations and therefore kidney function tests were included in the consensus survey to re-evaluate clinical opinion.
Haemostasis tests	• In original 2003 guideline • No new evidence identified • Included in modified Delphi survey	No new evidence was identified during the NICE update review, but the opinion from stakeholders during the scoping of the update was that clinical opinion and practice had changed since the original guideline published in 2003. The test was included in the consensus survey to re-evaluate current opinion.
Glycated haemoglobin (HbA1c) test	• Random blood glucose tests withdrawn from original 2003 guideline • Systematic review conducted for HbA1c testing • Included in modified Delphi survey	Random blood glucose tests have been replaced by HbA1c since the publication of the original guideline in 2003. A new review of HbA1c was undertaken however the evidence was not sufficient to enable the GDG to make recommendations and therefore was included in the consensus survey to evaluate clinical opinion.

Preoperative test	Methods	Rationale
Sickle cell disease/trait test	• In original 2003 guideline • No new evidence identified • Included in modified Delphi survey	No new evidence was identified during the NICE update review, but the opinion from stakeholders during the scoping of the update was that clinical opinion and practice had changed since the original guideline published in 2003. The test was included in the consensus survey to re-evaluate current opinion.
Urinalysis	• In original 2003 guideline • No new evidence identified • Included in modified Delphi survey	No new evidence was identified during the NICE update review, but the opinion from stakeholders during the scoping of the update was that clinical opinion and practice had changed since the original guideline published in 2003. The test was included in the consensus survey to re-evaluate current opinion.
Pregnancy testing	• In original 2003 guideline • No new evidence identified • Included in modified Delphi survey	Pregnancy testing was included in the consensus survey conducted for the original 2003 guideline. The test was included in this consensus survey to re-evaluate current opinion.

4.5 Searching for evidence

4.5.1 Clinical literature search

Systematic literature searches were undertaken to identify all published clinical evidence relevant to the review questions. Searches were undertaken according to the parameters stipulated within the NICE guidelines manual.[77,78] 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 8 July 2015. Papers published after this date were not considered.

Search strategies were quality assured by cross-checking reference lists of highly relevant papers, analysing search strategies in other systematic reviews, and asking GDG members to highlight any additional studies. Any additional papers identified were ordered by the technical team and assessed for inclusion in the full guideline. The review questions, the study types applied, the databases searched and the years covered can be found in Appendix G.

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, from organisations relevant to the topic. Searching for 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)
- NHS Evidence Search (www.evidence.nhs.uk)
- TRIP database (https://www.tripdatabase.com)
- BMJ Clinical Evidence (http://clinicalevidence.bmj.com)
- DUETS (https://www.library.nhs.uk/duets/)
- CRD (http://www.york.ac.uk/crd/)
- PROSPERO (http://www.crd.york.ac.uk/PROSPERO/)

4.5.2 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 elective surgery 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. This was supplemented by additional searches that looked for economic papers specifically relating to cardiopulmonary exercise testing on Medline, Embase, the NHS Economic Evaluations Database, the Health Technology Assessment database and the Health Economic Evaluation Database, as it became apparent that some papers in this area were not being identified through the first search. 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 G. All searches were updated on 8 July 2015. Papers published after this date were not considered.

4.6 Evidence of effectiveness

The tasks of the research fellow are listed below. The research fellow:

- Identified potentially relevant studies for each review question from the relevant search results by reviewing titles and abstracts, and deciding which should be ordered as full papers. Full papers were then obtained.
- Reviewed full papers against pre-specified inclusion/exclusion criteria to identify studies that addressed the review question in the appropriate population, and reported on outcomes of interest (see Appendix C for clinical review protocols).
- Critically appraised relevant studies using the appropriate study design checklists as specified in The Guidelines Manual.[77] Available from: https://www.nice.org.uk/article/PMG6/chapter/1Introduction
- Critically appraised relevant studies with a prognostic study design using the NCGC checklist.
- Extracted key information about interventional study methods and results using Evibase, NCGC purpose-built software. Evibase produces summary evidence tables, with critical appraisal ratings. Key information about non-interventional study methods and results were manually extracted onto standard evidence tables and critically appraised separately (see Appendix H for the clinical evidence tables).
- Generated summaries of the evidence by outcome. Outcome data was combined, analysed and reported according to study design:

  - Randomised data were meta-analysed where appropriate and reported in GRADE profiles
  - Observational data were presented as a range of values in GRADE profiles
  - Prognostic data were meta-analysed where appropriate and reported in GRADE profiles.
- A sample of a minimum of 20% of the abstract lists of the first three sifts by new reviewers were double sifted by a senior research fellow. As no papers were missed by any reviewers, no further double sifting was carried out. All of the evidence reviews were quality assured by a senior research fellow. This included checking:
  - papers were included or excluded appropriately
  - a sample of the data extractions
  - correct methods were used to synthesise data
  - a sample of the risk of bias assessments.

4.6.1 Inclusion and exclusion criteria

The inclusion and exclusion of studies was based on the clinical 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 GDG was consulted about any uncertainty regarding inclusion or exclusion.

The key population inclusion criteria were:

- Adults and young people (over 16 years of age) classified as patients who are ASA grade 1–4 undergoing elective surgery
- The guideline covers selected comorbidities: cardiovascular, respiratory, renal, obesity, diabetes

The key population exclusion criteria were:

- Children and young people (0–16 years old)
- Cardiovascular and neurological surgery
- Other comorbidities
- Pregnant women

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

Conference abstracts were not automatically excluded from the review but were initially assessed against the inclusion criteria and 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 relevant conference abstracts were identified for this guideline. Literature reviews, posters, letters, editorials, comment articles, unpublished studies and studies not in English were excluded.

The clinical review protocols are presented in Appendix C.

4.6.2 Type of studies

Randomised trials, non-randomised trials, and observational studies (including prognostic studies) were included in the evidence reviews as appropriate (See section 4.1.2).

Randomised controlled trials (RCTs) were not included within this guideline as no studies meeting the PICO were found. Non-randomised studies were found for the resting echocardiography, CPET and polysomnography evidence reviews but none were presented as a combined meta-analysis. The GDG considered the quality of evidence and made recommendations on observational data where appropriate (for example resting echocardiography).

Where comparative studies were absent or not considered to be of sufficient quality, prognostic reviews with RCTs, pooled analysis of patient level data, and retrospective cohort or prospective cohort studies were included. Case-control studies were excluded because of their high risk of recall bias. Recommendations for resting electrocardiography (ECG), HbA1c and all tests included in the HTA 2012 report were informed by prognostic reviews.

4.6.3 Methods of combining clinical studies

4.6.3.1 Data synthesis for intervention reviews

Where possible, meta-analyses were conducted to combine the data from the studies for each of the outcomes in the review question using RevMan5[4] software.

All analyses were stratified by surgery grade, ASA grade (grades 1–4), and comorbidity (cardiovascular, respiratory, renal, obesity, diabetes) where reported, which meant that strata were not combined and analysed together.

Analysis of different types of data

Dichotomous outcomes

Fixed-effects (Mantel-Haenszel) techniques (using an inverse variance method for pooling) were used to calculate risk ratios (relative risk) for the binary outcomes, which included:

- 30-day mortality
- Delay in surgery
- Hospital readmission

The absolute risk difference was also calculated using GRADEpro[42] software, using the median event rate in the control arm of the pooled results.

For binary variables where there were zero events in either arm, Peto odds ratios, rather than risk ratios, were calculated. Peto odds ratios are more appropriate for data with a low number of events.

Continuous outcomes

The continuous outcomes were analysed using an inverse variance method for pooling weighted mean differences. These outcomes included:

- Heath-related quality of life
- Length of stay

Where the studies within a single meta-analysis had different scales of measurement, standardised mean differences were used (providing all studies reported either change from baseline or final values rather than a mixture of the two), where each different measure in each study was 'normalised' to the standard deviation value pooled between the intervention and comparator groups in that same study.

The means and standard deviations of continuous outcomes are required for meta-analysis. However, in cases where standard deviations were not reported, the standard error was calculated if the p values or 95% confidence intervals were reported, and meta-analysis was undertaken with the mean and standard error using the generic inverse variance method in Cochrane Review Manager (RevMan5)[4] software. Where p values were reported as "less than", a conservative approach was undertaken. For example, if a p value was reported as "p ≤0.001", the calculations for standard deviations were 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 (version 5.1.0, updated March 2011) were applied.

Generic inverse variance

If a study reported only the summary statistic and 95% confidence intervals the generic-inverse variance method was used to enter data into RevMan5[4]. If the control event rate was reported this was used to generate the absolute risk difference in GRADEpro[42]. If multivariate analysis was used to derive the summary statistic but no adjusted control event rate was reported no absolute risk difference was calculated.

Heterogeneity

Statistical heterogeneity was assessed for each meta-analysis estimate by considering the chi-squared test for significance at $p<0.1$, or an I-squared inconsistency statistic of >50%, as indicating significant heterogeneity as well as the distribution of effects. Where significant heterogeneity was present, a priori subgrouping of studies (determined by the GDG in each protocol) was carried out.

If the subgroup analysis resolved heterogeneity within all of the derived subgroups, then each of the derived subgroups were adopted as separate outcomes (providing at least one study remained in each subgroup). For example instead of the single outcome of 'missed diagnosis', this would be separated into two outcomes. Assessments of potential differences in effect between subgroups were based on the chi-squared tests for heterogeneity statistics between subgroups. Any subgroup differences were interpreted with caution, as separating the groups breaks the study randomisation and as such, are subject to uncontrolled confounding.

For some questions additional subgrouping was applied, and this is documented in the individual clinical review protocols (Appendix C). These additional subgrouping strategies were applied independently, so subunits of subgroups were not created, unlike the situation with strata. Other subgrouping strategies were only used if the age category subgroup was unable to explain heterogeneity: then, these further subgrouping strategies were applied in order of priority. Again, once a subgrouping strategy was found to explain heterogeneity from all derived subgroups, further subgrouping strategies were not used.

If all pre-defined strategies of subgrouping were unable to explain statistical heterogeneity within each derived subgroup, then a random effects (DerSimonian and Laird) model was employed to the entire group of studies in the meta-analysis. A random-effects model assumes a distribution of populations, rather than a single population. This leads to a widening of the confidence intervals around the overall estimate, thus providing a more realistic interpretation of the true distribution of effects across more than one population. If, however, the GDG considered the heterogeneity was so large that meta-analysis was inappropriate, then the results were described narratively.

4.6.3.2 Data synthesis for prognostic factor reviews

A variety of prognostic effect measures were extracted from papers, depending on the type of outcome.

For binary outcomes, odds ratios (ORs), risk ratios (RRs) or hazard ratios (HRs) (with their 95% confidence intervals) for the independent effect of each prognostic factors on the outcome were extracted. Beta coefficients for dichotomous outcomes were normally converted to an OR by taking the anti-natural logarithm of the beta coefficient (as beta coefficient = ln OR).

For continuous outcomes the beta coefficients (or standardised beta coefficients) with their 95% confidence intervals for the independent effect of each prognostic factor were extracted.

RCTs, pooled analysis of patient level data, and prospective or retrospective cohort studies were included.

All non-RCT studies were required to have considered all key confounders previously identified by the GDG at the protocol stage for that outcome. For a key confounder to be regarded as having been adequately considered, it would have to have been included in the multivariable analysis (although in a step-wise model it would not necessarily have to be present in the final model) or would have to have been shown to be matched across risk factor or outcome groups at baseline. Moreover, the GDG identified several additional confounders for each prognostic protocol: these were desired and studies were not necessarily excluded if these were not adequately considered in final analysis. Univariate analysis was excluded from the final guideline.

If more than one study covered the same combination of population, risk factor and outcome then meta-analysis was used to pool results. Meta-analysis was carried out using the generic inverse variance function on Review Manager using fixed effects. Heterogeneity was assessed using the same criteria as for intervention studies, with an I^2 of 50–74% representing serious inconsistency and an I^2 of >75% representing very serious inconsistency. If serious or very serious heterogeneity existed, then subgrouping strategies were based on pre-specified subgrouping criteria as for interventional reviews. If subgrouping failed to explain heterogeneity, then the random effects model was used.

Where evidence was not meta-analysed because studies differed in population, outcome or risk factors then no alternative pooling strategies were carried out, on the basis that such pooling would have little meaning. Results from single studies were presented.

Studies of lower risk of bias were preferred, taking into account the analysis and the study design. In particular, prospective cohort studies were preferred if they reported multivariable analyses, which adjusted for the key confounders identified by the GDG at the protocol stage for that outcome.

4.6.4 Appraising the quality of evidence by outcomes

4.6.4.1 Interventional studies

The evidence for outcomes from the included 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 (GRADEpro[42]) developed by the GRADE working group was used to assess the quality of each outcome, taking into account individual study quality and the meta-analysis results.

Each outcome was first examined for each of the quality elements listed and defined in Table 5.

Table 5: Description of quality elements in GRADE for intervention studies

Quality element	Description
Risk of bias	Limitations in the study design and implementation may bias the estimates of the treatment effect. Major limitations in studies decrease the confidence in the estimate of the effect. Examples of such limitations are selection bias (often due to poor allocation concealment), performance and detection bias (often due to a lack of blinding of the patient, healthcare professional and assessor) and attrition bias (due to missing data causing systematic bias in the analysis).
Indirectness	Indirectness refers to differences in study population, intervention, comparator and outcomes between the available evidence and the review question.
Inconsistency	Inconsistency refers to an unexplained heterogeneity of effect estimates between studies in the same meta-analysis.
Imprecision	Results are imprecise when studies include relatively few patients and few events (or highly variable measures) and thus have wide confidence intervals around the estimate of the effect relative to clinically important thresholds. 95% confidence intervals denote the possible range of locations of the true population effect at a 95% probability, and so

Quality element	Description
	wide confidence intervals may denote a result that is consistent with conflicting interpretations (for example a result may be consistent with both clinical benefit AND clinical harm) and thus be imprecise.
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. A closely related phenomenon is where some papers fail to report an outcome that is inconclusive, thus leading to an over-estimate of the effectiveness of that outcome.
Other issues	Sometimes randomisation may not adequately lead to group equivalence of confounders, and if so this may lead to bias, which should be taken into account. Potential conflicts of interest, often caused by excessive pharmaceutical company involvement in the publication of a study, should also be noted.

Details of how the four main quality elements (risk of bias, indirectness, inconsistency and imprecision) were appraised for each outcome are given below. Publication or other bias was only taken into consideration in the quality assessment if it was apparent.

Risk of bias

The main domains of bias for RCTs are listed in Table 6. Each outcome had its risk of bias assessed within each paper first. For each paper, if there were no risks of bias in any domain, the risk of bias was given a rating of 0. If there was risk of bias in just one domain, the risk of bias was given a 'serious' rating of -1, but if there was risk of bias in two or more domains the risk of bias was given a 'very serious' rating of -2. A weighted average score was then calculated across all studies contributing to the outcome, by taking into account the weighting of studies according to study precision. For example if the most precise studies tended to each have a score of -1 for that outcome, the overall score for that outcome would tend towards -1.

Table 6: Principle domains of bias in randomised controlled trials

Limitation	Explanation
Selection bias – sequence generation and allocation concealment	If those enrolling patients are aware of the group to which the next enrolled patient will be allocated, either because of a non-random sequence that is predictable, or because a truly random sequence was not concealed from the researcher, this may translate into systematic selection bias. This may occur if the researcher chooses not to recruit a participant into that specific group because of 1) knowledge of that participant's likely prognostic characteristics and 2) a desire for one group to do better than the other.
Performance and detection bias - lack of patient and health care professional blinding	Patients, caregivers, those adjudicating and/or recording outcomes, and data analysts should not be aware of the arm to which patients are allocated. Knowledge of group can influence 1) the experience of the placebo effect, 2) performance in outcome measures, 3) the level of care and attention received, and 4) the methods of measurement or analysis, all of which can contribute to systematic bias.
Attrition bias	Attrition bias results from loss of data beyond a certain level (a differential of 10% between groups) which is not accounted for. Loss of data can occur when participants are compulsorily withdrawn from a group by the researchers (for example when a per-protocol approach is used) or when participants do not attend assessment sessions. If the missing data are likely to be different from the data of those remaining in the groups, and there is a differential rate of such missing data from groups, systematic attrition bias may result.
Selective outcome reporting	Reporting of some outcomes and not others on the basis of the results can also lead to bias, as this may distort the overall impression of efficacy.
Other limitations	For example:

Limitation	Explanation
	• Stopping early for benefit observed in randomised trials, in particular in the absence of adequate stopping rules • Use of unvalidated patient-reported outcomes • Lack of washout periods to avoid carry-over effects in cross-over trials Recruitment bias in cluster randomised trials

Indirectness

Indirectness refers to the extent to which the populations, interventions, comparisons and outcome measures are dissimilar 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. As for risk of bias, each outcome had its indirectness assessed within each paper first. For each paper, if there were no sources of indirectness, indirectness was given a rating of 0. If there was indirectness in just one source (for example in terms of population), indirectness was given a 'serious' rating of -1, but if there was indirectness in two or more sources (for example, in terms of population and treatment) the indirectness was given a 'very serious' rating of -2. A weighted average score was then calculated across all studies contributing to the outcome, by taking into account study precision. For example if the most precise studies tended to have an indirectness score of -1 each for that outcome, the overall score for that outcome would probably tend towards -1.

Inconsistency

Inconsistency refers to an unexplained heterogeneity of results for an outcome across different studies. When estimates of the treatment effect across studies differ widely, this suggests true differences in underlying treatment effect, which may be due to differences in populations, settings or doses. When heterogeneity existed within an outcome (Chi square $p<0.1$ or I^2 inconsistency statistic of >50%), but no plausible explanation could be found, the quality of evidence for that outcome was downgraded. Inconsistency for that outcome was given a 'serious' score of -1 if the I^2 was 50–74, and a 'very serious' score of -2 if the I^2 was 75 or more.

If inconsistency could be explained based on pre-specified subgroup analysis (that is, each subgroup had an I^2 <50), the GDG took this into account and considered whether to make separate recommendations on new outcomes based on the subgroups defined by the assumed explanatory factors. In such a situation the quality of evidence was not downgraded for those emergent outcomes.

Since the inconsistency score was based on the meta-analysis results, the score represented the whole outcome and so weighted averaging across studies was not necessary.

Imprecision

The criteria applied for imprecision were based on the confidence intervals for the pooled estimate of effect, and the minimal important differences (MID) for the outcome. The MIDs are the threshold for appreciable benefits and harms, separated by a zone either side of the line of no effect where there is assumed to be no clinically important effect. If either of the 95% confidence intervals of the overall estimate of effect crossed one of the MID lines, imprecision was regarded as serious and a 'serious' score of -1 was given. This was because the overall result, as represented by the span of the confidence intervals, was consistent with two interpretations as defined by the MID (for example, no clinically important effect and either clinical benefit or harm). If both MID lines were crossed by either or both of the confidence intervals then imprecision was regarded as very serious and a 'very serious' score of -2 was given. This was because the overall result was consistent with three

interpretations defined by the MID (no clinically important effect and clinical benefit and clinical harm). This is illustrated in Figure 2. As for inconsistency, since the imprecision score was based on the meta-analysis results, the score represented the whole outcome and so weighted averaging across studies was not necessary.

The position of the MID lines is ideally determined by values as reported in the literature. "Anchor-based" methods aim to establish clinically meaningful changes in a continuous outcome variable by relating or "anchoring" them to patient-centred measures of clinical effectiveness that could be regarded as gold standards with a high level of face validity. For example, the minimum amount of change in an outcome necessary to make a patient decide that they felt their quality of life had "significantly improved" might define the MID for that outcome. MIDs in the literature may also be based on expert clinician or consensus opinion concerning the minimum amount of change in a variable deemed to affect quality of life or health. For binary variables, any MIDs reported in the literature will inevitably be based on expert consensus. As such, MIDs relate to all-or-nothing population effects rather than measurable effects on an individual, and so are not amenable to patient-centred "anchor" methods.

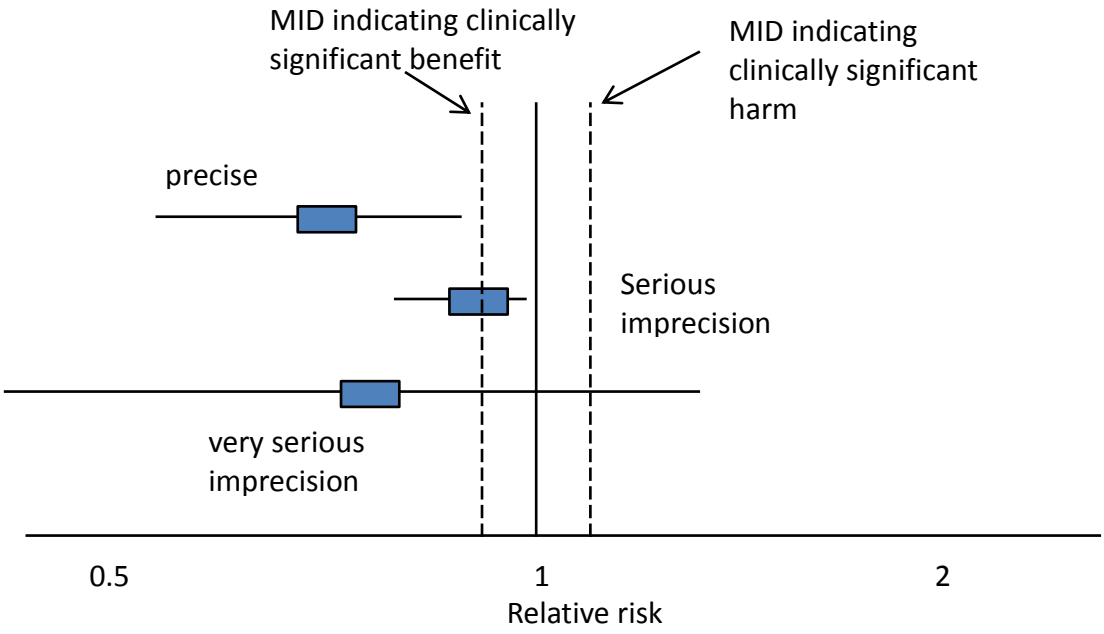
In the absence of literature values, the alternative approach to deciding on MID levels is the "default" method, as follows:

- For categorical outcomes the MIDs are taken as RRs of 0.75 and 1.25. For 'positive' outcomes such as 'patient satisfaction', the RR of 0.75 is taken as the line denoting the boundary between no clinically important effect and a clinically significant harm, whilst the RR of 1.25 is taken as the line denoting the boundary between no clinically important effect and a clinically significant benefit. For 'negative' outcomes such as 'bleeding', the opposite occurs, so the RR of 0.75 is taken as the line denoting the boundary between no clinically important effect and a clinically significant benefit, whilst the RR of 1.25 is taken as the line denoting the boundary between no clinically important effect and a clinically significant harm.
- For continuous outcome variables the MID is taken as half the median baseline standard deviation of that variable across all studies in the meta-analysis. Hence the MID denoting the minimum clinically significant benefit will be a positive for a "positive" outcome (for example, a quality of life measure where a higher score denotes better health), and negative for a "negative" outcome (for example, a visual analogue scale [VAS] pain score). Clinically significant harms will be the converse of these. If baseline values are unavailable then half the median comparator group standard deviation of that variable will be taken as the MID.
- If standardised mean differences have been used, then the MID will be set at the absolute value of +0.5. This follows because standardised mean differences are mean differences normalised to the pooled standard deviation of the two groups, and are thus effectively expressed in units of "numbers of standard deviation". The 0.5 MID value in this context therefore indicates half a standard deviation: the same definition of MID as used for non-standardised mean differences.

The default MID value was subject to amendment after discussion with the GDG. If the GDG decided that the MID level should be altered, after consideration of absolute as well as relative effects, this was allowed, provided that any such decision was not influenced by any bias towards making stronger or weaker recommendations for specific outcomes.

For this guideline, no appropriate MIDs for continuous or dichotomous outcomes were found in the literature, and so the default method was used.

Figure 2: Illustration of precise and imprecise outcomes based on the confidence interval of dichotomous outcomes in a forest plot. Note that all three results would be pooled estimates and would not, in practice, be placed on the same forest plot



Overall grading of the quality of clinical evidence

Once an outcome had been appraised for the main quality elements, as above, an overall quality grade was calculated for that outcome. The scores from each of the main quality elements (0, -1 or -2) were summed to give a score that could be anything from 0 (the best possible) to -8 (the worst possible). However, scores were capped at -3. This final score was then applied to the starting grade that had originally been applied to the outcome by default, based on study design. For example, all RCTs started as HIGH and the overall quality became MODERATE, LOW or VERY LOW if the overall score was -1, -2 or -3 points respectively. The significance of these overall ratings is explained in Table 7. The reasons or criteria used for downgrading were specified in the footnotes of the GRADE tables.

On the other hand, observational interventional studies started at LOW, and so a score of -1 would be enough to take the grade to the lowest level of VERY LOW. Observational studies could, however, be upgraded if there was: a large magnitude of effect, a dose-response gradient, and if all plausible confounding would reduce a demonstrated effect.

Table 7: 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

4.6.4.2 Prognostic studies

The quality of evidence for prognostic studies was evaluated according to the criteria given in Table 6. If data were meta-analysed the quality for pooled studies was presented. If the data was not pooled then a quality rating was presented for each study. A modified GRADE methodology was used for prognostic studies, considering risk of bias, indirectness, inconsistency and imprecision.

Risk of bias

The quality of evidence for prognostic studies was evaluated according to the criteria given in Table 8.

Table 8: Description of quality elements for prospective studies

Domain	Risk of bias for prognostic risk factor studies	Response and score
Selection bias	Was there a lack of reported attempts made to achieve some group comparability between the risk factor and non-risk factor groups? (ignore if 2 or more risk factors considered)	Consider if this was moderate, high or very high risk of bias if answer was 'yes'.
	Was there a lack of consideration of any of the key confounders, or was this unclear? Note that if the study can show that a particular confounder was not at risk of causing bias (for example by being well-matched at baseline between groups) then this confounder does not have to have been adjusted for in a multivariate analysis.	EXCLUDE
	Was there a lack of consideration of non-key plausible confounders, or was this unclear? Note that if the study can show that a particular confounder was not at risk of causing bias (for example by being well-matched at baseline between groups) then this confounder does not have to have been adjusted for in a multivariate analysis.	Consider if this was moderate, high or very high risk of bias if answer was 'yes'.
	If the outcome is categorical: were there <10 events per variable included in the multivariable analysis? If the outcome is continuous: were there <10 people per variable included in the multivariable analysis?	Consider if this was moderate, high or very high risk of bias if answer was 'yes' to either
	Was it very clear that one group was more likely to have had more outcomes occurring at baseline than another group?	Consider if this was moderate, high or very high risk of bias if answer was 'yes'.
Detection bias	Was there a lack of assessor blinding AND the outcome was not completely objective?	Consider if this was moderate, high or very high risk of bias if answer was 'yes'.
	Were the risk factors measured in a way that would systematically favour either group?	Consider if this was moderate, high or very high risk of bias if answer was 'yes'.
	Were the outcomes measured in a way that would systematically favour either group?	Consider if this was moderate, high or very high risk of bias if answer was 'yes'.
	If there were multiple raters, was there lack of adjustment for systematic inter-rater measurement errors, or was inter-rater reliability unreported?	Consider if this was moderate, high or very high risk of bias if answer was 'yes'.
	Was there an excessively short follow up, such that there was not enough time for outcomes to occur?	Consider if this was moderate, high or very high risk of bias if answer was 'yes'.

Domain	Risk of bias for prognostic risk factor studies	Response and score
Attrition bias	Was there >10% group differential attrition (for reasons related to outcome) and there was no appropriate imputation? (if one risk factor) or Was there >10% overall attrition (for reasons related to outcome) and there was no appropriate imputation? (if >1 risk factor).	Consider if this was moderate, high or very high risk of bias if answer was 'yes'. Consider if this was moderate, high or very high risk of bias if answer was 'yes'.
	For each domain make a judgement of risk of bias (for example very high if there are two moderate boxes and a high box) Sum these domain risks to form an overall rating of risk of bias (for example no risk, serious risk or very serious risk)	

The risk of bias rating was assigned per study for each outcome. When studies were pooled the overall risk of bias for all studies covering a specific risk factor/outcome was determined by a weighted mean of the ratings across the studies (with no risk = 0; serious risk = -1 and very serious risk = -2). The weighting depended on the weighting used in the meta-analysis, as for intervention reviews. Where a meta-analysis had not been conducted a simple average was used.

Indirectness

See section 4.3.4

Inconsistency

See section 4.3.4

Imprecision

In meta-analysed outcomes, or for non-pooled outcomes, the position of the confidence intervals in relation to the null line determined the existence of imprecision. If the confidence intervals did not cross the null line, then no serious imprecision was recorded. If the confidence intervals crossed the null line, then serious imprecision was recorded.

Overall grading of the quality of clinical evidence

Quality rating started at HIGH for prospective studies, and each major limitation (see Table 6) brought the rating down by one increment to a minimum grade of LOW, as explained for interventional studies (See section 4.3.4.1).

4.6.5 Assessing clinical importance

The GDG 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.

Quality rating started at HIGH for prospective studies, and each major limitation (see Table 6) brought the rating down by one increment to a minimum grade of LOW, as explained for interventional studies.

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

4.6.6 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
- 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).

4.7 Evidence of cost-effectiveness

The GDG 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.[78] 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.

4.7.1 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 NICE guidelines manual.[77,78]
- Generated summaries of the evidence in NICE economic evidence profiles (included in the relevant chapter for each review question) – see below for details.

4.7.1.1 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–consequences 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. Studies published before 1999 and studies from non-OECD countries or the USA were also excluded, on the basis that the applicability of such studies to the present UK NHS context is likely to be too low for them to be helpful for decision-making.

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 Table 9 below and the economic evaluation checklist (Appendix G of the NICE guidelines manual 2012[77]) and the health economics review protocol in Appendix D.

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 GDG to inform the possible economic implications of the recommendations.

4.7.1.2 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 NICE guidelines manual.[77] 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 9 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.[80]

Table 9: 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 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

Item	Description
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 NICE guidelines manual (2012)[77]

4.7.2 Undertaking new health economic analysis

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

Given the lack of clinical evidence in other review areas, the GDG identified polysomnography as the highest priority area for original economic modelling. Initial scoping searches did not identify any studies assessing the cost-effectiveness of polysomnography in the preoperative setting. Given the cost of polysomnography and the high prevalence of obstructive sleep apnoea, the GDG felt that polysomnography could be assigned high priority for original economic modelling, subject to the results of the clinical review.

However due to insufficient clinical and cost-effective evidence, original economic analysis was deemed unfeasible. In the absence of original economic analysis, unit costs of preoperative testing for polysomnography were presented to inform recommendations (see section 9.6).

4.7.3 Cost-effectiveness criteria

NICE's report 'Social value judgements: principles for the development of NICE guidance' sets out the principles that GDGs should consider when judging whether an intervention offers good value for money.[76] 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 GDG 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'.[76]

If a study reported the cost per life year gained but not QALYs, the cost per QALY gained was estimated by multiplying by an appropriate utility estimate to aid interpretation. The estimated cost per QALY gained is reported in the economic evidence profile with a footnote detailing the life years gained and the utility value used. When QALYs or life years gained are not used in the analysis, results are difficult to interpret unless one strategy dominates the others with respect to every relevant health outcome and cost.

4.7.4 In the absence of economic evidence

When no relevant published studies were found, and new economic analysis was not prioritised, the GDG 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 effectiveness review.

The UK NHS costs reported in the guideline are those that were presented to the GDG and were correct at the time recommendations were drafted. They may have changed subsequently before the time of publication. However, we have no reason to believe they have changed substantially.

4.8 Developing recommendations

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

- Evidence tables of the clinical and economic evidence reviewed from the literature. All evidence tables are in Appendix H
- Summaries of clinical and economic evidence and quality (as presented in Chapters 5–17)
- Forest plots (Appendix J)
- A description of the methods of the cost-effectiveness analysis undertaken for the guideline (Appendix M)
- Results of the modified Delphi consensus survey (Appendix L)

Recommendations were drafted on the basis of the GDG's 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 GDG 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 GDG's values and preferences), and the confidence the GDG had in the evidence (evidence quality). Secondly, whether the net benefit justified any differences in costs was assessed.

When clinical and economic evidence was of poor quality, conflicting or absent, the GDG 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 GDG meeting. The results of the modified Delphi consensus survey were considered alongside any evidence where available, and informed the GDG in their decision making when drafting recommendations. The GDG 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 4.8.1 below).

The GDG considered the 'strength' of recommendations. This takes into account the quality of the evidence but is conceptually different. Some recommendations are 'strong' in that the GDG believes that the vast majority of healthcare and other professionals and patients would choose a particular intervention if they considered the evidence in the same way that the GDG has. This is generally the case if the benefits clearly outweigh the harms for most people and the intervention is likely to be cost-effective. However, there is often a closer balance between benefits and harms, and some patients would not choose an intervention whereas others would. This may happen, for example, if some patients are particularly averse to some side effects and others are not. In these circumstances the recommendation is generally weaker, although it may be possible to make stronger recommendations about specific groups of patients.

The GDG focused on the following factors in agreeing the wording of the recommendations:

- 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 weaker 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 (see Section 9.3 in the NICE guidelines manual[77]).

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

For certain areas (for example resting ECG) the GDG used a traffic light system to show the degree of consensus reached by the GDG (see Table 10 below)

Table 10: Traffic light system

Offer	Test recommended
Do not routinely offer	Test not recommended on a routine basis
Consider	The value of carrying out a preoperative test is not known and may depend on specific patient characteristics

The ASA Physical Status Classification System was used when developing recommendations to describe fitness to undergo an anaesthetic. See section 3.3.1.1 for an overview of the ASA grades used.

4.8.1 Research recommendations

When areas were identified for which good evidence was lacking, the GDG 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.

4.8.2 Validation process

This guideline is subject to a 6-week public consultation and feedback as part of the quality assurance and peer review of the document. All comments received from registered stakeholders are responded to in turn and posted on the NICE website.

4.8.3 Updating the guideline

Following publication, and in accordance with the NICE guidelines manual, NICE will undertake a review of whether the evidence base has progressed significantly to alter the guideline recommendations and warrant an update.

4.8.4 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 Clinical 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.

4.8.5 Funding

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

5 Resting electrocardiography

5.1 Introduction

Resting 12-lead electrocardiography (ECG) is a non-invasive test that can detect abnormalities including arrhythmias, evidence of coronary heart disease, left ventricular hypertrophy and bundle branch blocks. In the preoperative setting, resting ECG is used to assess known cardiovascular diseases, to detect previously undiagnosed cardiovascular diseases, and to provide a baseline standard against which to measure changes in the postoperative period. When conducted by a suitably trained individual, the resting ECG is simple to perform and interpret, with the only described complication being minor allergy to the ECG electrodes resulting in self-limiting skin reddening. However, there is uncertainty regarding the prognostic significance of different resting ECG abnormalities in the perioperative setting, especially in asymptomatic patients.

See section 4.4 for a summary of the methodological approach taken for this preoperative test.

5.2 Review question (intervention): What is the clinical- and cost-effectiveness of using resting electrocardiography (ECG) as a preoperative test in improving patient outcomes in adults and young people undergoing non-cardiac elective surgery?

Table 11: PICO characteristics of review question

Population	All adults and young people (ASA grade 1 or above) undergoing non-cardiac surgery. Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory and renal diseases, obesity, diabetes
Intervention	Preoperative resting ECG
Comparison(s)	No preoperative resting ECG
Outcomes	**Critical:** • All-cause mortality • Health-related quality of life **Important:** • Complications related to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney failure, infection) • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing • Intensive care unit (ICU) admission • Composite outcomes such as major adverse cardiovascular events (MACE) that incorporate cardiac deaths and non-fatal cardiac events • Optimisation of medical therapy

5.3 Clinical evidence

No relevant clinical studies comparing preoperative resting ECG testing with no preoperative resting ECG testing were identified.

5.4 Review question (prognostic): Does resting ECG predict prognosis (patient outcomes after surgery) in adults and young people undergoing non-cardiac elective surgery?

Table 12: Characteristics of the review question

Population	All adults and young people (ASA grade 1 or above) undergoing non-cardiac elective surgery. Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory and renal diseases, obesity, diabetes
Prognostic test	Resting ECG
Key confounding factors	• Age • Comorbidities
Outcomes (30 days post-surgery)	**Critical** • All-cause mortality **Important** • Complications relating to surgery or anaesthesia • Length of hospital stay • Hospital readmission • Adverse events caused by testing • Health-related quality of life • Intensive care unit (ICU) admission

5.5 Clinical evidence

Seven studies were included in the review.[14,38,60-62,65,89] Evidence from these are summarised in the clinical evidence profile below. See also the clinical study selection flow chart in Appendix E, forest plots in Appendix J, GRADE tables in Appendix I, clinical evidence tables in Appendix H and excluded clinical studies list in Appendix K.

All studies conducted a multivariable analysis, but different variables were analysed in the studies (see Table 13).

Table 13: Summary of studies included in the review

Study	Population	Analysis	Prognostic variable(s)	Confounders (list	Outcomes	Limitations
Biteker 2012[14]	Single prospective cohort n=660	Multivariate logistic regression	Resting ECG	Age, gender, comorbidity, pharmacological treatment, QRS duration, clinical risk indicators	Perioperative cardiovascular event	Short follow-up period. High risk surgery not included in analysis. Only patients with a preoperative cardiovascular work up were included. High risk of bias.
Fritsch 2012[38]	Single centre prospective cohort n=1363	Multivariate forward likelihood ratio	Resting ECG	Gender, age, invasiveness of procedure, comorbidity, preoperative tests	Cardiac, cerebro-vascular, respiratory and bleeding complications.	Assessor blinding not clear. Length of follow up not standardised. High risk of bias.
Koike 1999[60]	Single centre prospective cohort n=114	Multivariate Cox proportional hazard analysis	Resting ECG	Age, gender, type of fracture, preoperative interval, intercurrent illness, type of housing, Goldman's cardiac risk index, preoperative dependence, mental function, anaemia, blood urea, ECG abnormality, malignancy, malnutrition.	One year mortality	Assessor blinding not clear. Inter-rater reliability unknown. High risk of bias.
Kyo 1993[61]	Single centre retro-spective cohort	Multivariate Cox proportional hazard	Resting ECG	Gender, age, prefracture activities of daily living	Survival rate	High risk of bias.

Study	Population	Analysis	Prognostic variable(s)	Confounders (list	Outcomes	Limitations
	n=427	analysis		(ADL), ECG, electro-encephalography (EEG), Hasewaga's score, haemoglobin, total protein, type of fracture.		Assessor blinding not clear.
Liu 2002[65]	Single centre prospective cohort n=513	Multivariate stepwise logistic regression	Resting ECG	Confounding variables not clearly described	Postoperative cardiac complications	High risk of bias. Downgraded due to lack of clarity regarding confounding variables. Potential for systematic error as ECG findings were considered in unison with general health.
Landesberg 2007[62]	Single centre retro-spective cohort n=624	Multivariate Cox proportional hazard analysis	Resting ECG	Age, diabetes mellitus, cerebrovascular disease, ischaemic heart disease congestive heart failure, kidney disease, preoperative creatinine , calcium blockers, hypolipidaemic agents.	Long-term survival	High risk of bias. Retrospective. No external validation.
Vanklei 2007[89]	Multicentre prospective cohort n=2967	Multivariate logistic regression	Resting ECG	Age, gender, high risk surgery, ischaemic heart disease, right bundle branch block, left bundle branch block.	Postoperative myocardial infarction Death during admission	Multiple raters not adjusted for. Length of follow up not standardised. Retrospective. High risk of bias.

Table 14: Clinical evidence summary: Non-cardiac, non-vascular surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Normal ECG versus prolonged QTc interval for predicting perioperative cardiovascular events (adjusted OR)	1	Adjusted OR [95% CI]: OR 1.04 [1.03–1.06]	No serious imprecision	LOW[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*

Table 15: Clinical evidence summary: Elective surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Normal ECG versus abnormal ECG for predicting cardiac, cerebrovascular, respiratory and bleeding complications (adjusted ORs)	1	Adjusted OR [95% CI]: 2.81 [1.36–5.82]	No serious imprecision	LOW[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*

Table 16: Clinical evidence summary: Hip fracture surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Normal ECG versus abnormal ECG for predicting one year mortality (adjusted RRs)	1	Adjusted RR: 1.54 [0.95–2.50]	Serious	LOW[ab]
Normal ECG versus abnormal ECG for predicting survival rate (adjusted HRs)	1	Adjusted HR: 2.66 [1.54–4.59]	No serious imprecision	MODERATE[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

Table 17: Clinical evidence summary: Major vascular surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Normal ECG versus ST segment depression for predicting long-term survival (adjusted HRs)	1	Adjusted HR [95% CI]: 1.94 [1.48–2.54]	No serious imprecision	MODERATE[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*

Table 18: Clinical evidence summary: Non-cardiac surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Normal ECG versus abnormal ECG for predicting postoperative cardiac complications (adjusted ORs)	1	Adjusted OR: 0.63 [0.28–1.42]	Serious imprecision	VERY LOW[ab]
Normal ECG versus left bundle branch block for predicting postoperative myocardial infarction (adjusted ORs)	1	Adjusted OR [95% CI]: 3.10 [1.0–9.61]	Serious	VERY LOW[ab]
Normal ECG versus left bundle branch block for predicting death during admission (adjusted ORs)	1	Adjusted OR [95% CI]: 3.50[1.30–9.42]	No serious imprecision	LOW[a]
Normal ECG versus right bundle branch block for predicting postoperative myocardial infarction (adjusted ORs)	1	Adjusted OR [95% CI]: 2.10 [1.00–4.41]	Serious	VERY LOW[ab]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

5.6 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

Unit costs

Unit costs were provided for consideration alongside the Delphi survey results. Please see Appendix M for details. These are reported in Table 19 and Table 20 below.

Table 19: Electrocardiography (cost using resources used)

Equipment/staff	Quantity (units/minutes)	Cost	Source
ECG machine[(a)]	1	£2.00	GDG opinion
Disposables[(b)]	1	£1.00	NHS supply chain catalogue[3]
Nurse time	10	£5.66	PSSRU 13/14[23]
	Total per patient	£8.66	

(a) The machine has been estimated to cost £2,000 and if we assume it is used 1,000 times before it is replaced, the marginal cost of using this machine equates to £2.00 per patient.
(b) Disposables incorporated with ECG testing currently estimated at £1.00 per patient. This includes resting ECG disposable electrodes at £0.03 per electrode and other equipment such as gels and cables.
(c) Nurse time based on day ward nurse costing £34 per hour.

Table 20: Electrocardiography (cost identified in the NHS reference costs)

Equipment/staff	Quantity (units/minutes)	Cost	Source
ECG[(a)]	1	£37	NHS reference costs 2010-11[28]
	Total per patient	£37	

The GDG noted that the cost reported in the NHS Reference Cost was likely to be obtained from a specialist setting and that in reality the cost would likely be lower in a preoperative setting. Therefore this cost can be seen as a maximum.

5.7 Evidence statements

5.7.1 Clinical

For forest plots, see Section J.1 in Appendix J.

5.7.1.1 Intervention review

No relevant studies were identified.

5.7.1.2 Prognostic review

Non-cardiac, non-vascular surgery

One prospective cohort found a prolonged QT interval to be an independent predictor of perioperative cardiovascular events (OR=1.043 [1.028 to 1.058]) in multivariate analysis [Very low quality].

Elective surgery

One prospective cohort found an abnormal ECG to be an independent predictor of postoperative cardiac, cerebrovascular, respiratory and bleeding complications (OR=2.814 [1.36 to 5.82]) in multivariate analysis [Very low quality].

Hip fracture surgery

One prospective cohort found an abnormal ECG to be an independent predictor of mortality at one year (RR=1.54 [0.95 to 2.49]) in multivariate analysis [Very low quality].

One retrospective cohort found an abnormal ECG to be an independent predictor of survival (HR=2.66 [1.54 to 4.59]) in multivariate analysis [Very low quality].

Major vascular surgery

One retrospective cohort found ST segment depression to be an independent predictor of long-term survival (HR=1.94 [1.48 to 2.54]) in multivariate analysis [Low quality].

Non-cardiac surgery

Two prospective cohort studies were included. One found an abnormal ECG to be an independent predictor of postoperative cardiac complications (OR=0.63 [0.28 to 1.42]) in multivariate analysis [Very low quality].

The second found left bundle branch block to be an independent predictor of both postoperative myocardial infarction and death during admission (OR=3.1 [1.0 to 9.61]) and (OR=3.5 [1.3 to 9.42]) respectively, in multivariate analysis [Very low quality].

The same study also found right bundle branch block to be an independent predictor of postoperative myocardial infarction (OR=2.1 [1.0 to 4.41]) in multivariate analysis [Very low quality].

5.7.2 Economic

No relevant economic evaluations were identified.

5.8 Delphi survey results

Preoperative resting electrocardiography (ECG) was included in the modified Delphi consensus survey.

The survey participants were asked if resting electrocardiography (ECG) should be used as a routine preoperative test for patients undergoing elective surgery. Participants rated their response from strongly disagree to strongly agree using a nine-point Likert scale. Each question was considered to have reached consensus if more than 70% of responses were in a single category (0–3 strongly disagree, 4–6 unclear, 7–9 strongly agree). Please see Appendix L for full details on the survey method and results.

5.8.1 Delphi statements where consensus was reached

Table 21: Patients: ASA1

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	86.22 strongly disagree (round 1)

Table 22: Patients: ASA2 with cardiovascular comorbidity

Surgery grade	Results % (round in which consensus was achieved)
Intermediate surgery	79.38 strongly agree (round 1)
Major or complex surgery	83.04 strongly agree (round 1)

Table 23: Patients: ASA3 or ASA4 with cardiovascular comorbidity

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	75.58 strongly agree (round 1)
Intermediate surgery	81.07 strongly agree (round 1)
Major or complex surgery	84.79 strongly agree (round 1)

Table 24: Patients: ASA2 with diabetes, respiratory, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Intermediate surgery	76.19 strongly agree (round 2)
Major or complex surgery	79.65 strongly agree (round 1)

Table 25: Patients: ASA3 or ASA4 with diabetes, respiratory, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	76.05 strongly agree (round 3)
Intermediate surgery	77.77 strongly agree (round 1)
Major or complex surgery	83.82 strongly agree (round 1)

5.8.2 Delphi statements where consensus was not reached

Table 26: Patients: ASA1

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Intermediate surgery	1	63.37	16.86
Major or complex surgery		36.00	21.71
Intermediate surgery	2	64.68	14.64
Major or complex surgery		28.28	50.59
Intermediate surgery	3	66.67	9.72
Major or complex surgery		36.11	36.10

Table 27: Patients: ASA2 with cardiovascular comorbidity

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Minor surgery	1	22.73	24.43
	2	12.95	47.21
	3	17.81	64.39

Table 28: Patients: ASA2 with diabetes, respiratory, renal or obesity comorbidities

Surgery grade		Results %	
	Round of Delphi	**Strongly disagree**	**Strongly agree**
Minor surgery	1	26.70	24.43
	2	20.23	48.81
	3	22.91	44.45

5.9 Recommendations and link to evidence

Recommendations

4. Resting ECG

	Surgery grade		
ASA grade	**Minor**	**Intermediate**	**Major or complex**
ASA 1	**Do not routinely offer**	**Do not routinely offer**	**Consider for people aged over 65 if no ECG results available from past 12 months**
ASA 2	**Do not routinely offer**	**Consider for people with cardiovascular, renal or diabetes comorbidities**	**Offer**
ASA3 or ASA4	**Consider if no ECG results available from past 12 months**	**Offer**	**Offer**

Relative values of different outcomes	The GDG identified all-cause mortality as a critical outcome. The following outcomes were identified as important: complications relating to surgery or anaesthesia, length of hospital stay, admission to intensive care or readmission to hospital, adverse events caused by testing and health-related quality of life.
Trade-off between clinical benefits and harms	There is a low risk associated with undergoing a resting ECG and the GDG considered the test to be relatively easy to perform with the correct training. Electrocardiography can identify cardiac abnormalities including arrhythmias, and sometimes indicates abnormalities of the right or left ventricular structure (hypertrophy or strain pattern), or its blood supply (ischaemia). Identification of these abnormalities preoperatively can lead to further investigations and treatments which may reduce postoperative complications. The GDG felt that

	obtaining a preoperative resting ECG was of value as it served as a baseline comparator in the event of a patient being investigated for a suspected perioperative adverse cardiac event. However, there were some concerns among the GDG that performing a preoperative resting ECG in asymptomatic patients could lead to detection of abnormalities with unknown clinical significance. This could potentially lead to further investigations and treatment that might have risks and could delay surgery unnecessarily.
Economic evidence	No economic evaluations were identified for this review question.
Quality of evidence	The ideal evidence would have been testing as an intervention rather than a prognostic factor. However, no such evidence was identified. The prognostic evidence identified was of mainly low or very low quality. While the evidence suggested that negative resting ECG results were predictive of adverse events, this was problematic as it was unclear whether test results led to management changes prior to surgery, or whether the physician was aware of the test results prior to surgery. For this reason, the GDG agreed to add this review question to the Delphi survey.
Other considerations	Some GDG members felt that resting ECG is generally over-used in practice. In current practice, the test is done routinely in all patients over 65 and in younger patients with known cardiovascular disease, and this is reflected in the 2003 guideline. The GDG agreed that the current recommendations should be re-evaluated within the Delphi survey to assess if the consensus view amongst health professionals has changed. It was noted that the European Society of Anaesthesiology (ESA) guideline[25] currently recommends resting ECG for patients undergoing intermediate or high risk surgery with risk factors for ischaemic heart disease, heart failure, stroke, diabetes or renal dysfunction. The GDG also noted an RCT[21] (Chung 2009) which was a battery of tests (full blood count, electrolytes, blood glucose, creatinine, ECG and chest X-ray). As this study presented the outcomes based on a battery of tests versus no testing, it did not match our protocol. However, the study did not conflict with the included evidence. The paper concluded that there was no significant difference in perioperative adverse events between the testing and no testing groups, and that there was no association between abnormal test results and perioperative adverse events.
Delphi	**Minor surgery** The results of the Delphi survey suggested that clinicians would consider resting ECG in minor surgery. The GDG queried the value of resting ECG as a routine preoperative test for minor elective surgery as the results of the test are unlikely to affect management. For ASA1 and ASA2 patients, the GDG felt that resting ECG is often used to provide reassurance to clinicians without having any clinical benefit. While the GDG accepted that it was tempting for clinicians to use a preoperative resting ECG as a baseline measurement, this would be largely unhelpful for minor surgery as patients are rarely monitored post-surgery. However for ASA3 or ASA4 patients, there may be some circumstances when ECG would be considered, such as uncontrolled atrial fibrillation. The Delphi survey showed that clinicians would complete a preoperative resting ECG for ASA3 or ASA4 patients with cardiovascular, diabetes, respiratory, renal or obesity comorbidities. Moreover, these results were consistent with the current guidance from the ESA. The GDG discussed that it was quite common to detect abnormalities, especially in older patients and patients with renal, cardiovascular and diabetes comorbidities.

	However in most cases this information would not alter perioperative management. The GDG also commented that resting ECG may not be the most effective means of diagnosing abnormalities preoperatively. **Intermediate surgery** The GDG noted that the ESA guideline recommends resting ECG as a preoperative test for intermediate surgery.[25] The Delphi survey results showed that clinicians would consider using resting ECG before intermediate surgery in ASA1 patients. However, the GDG felt that any abnormalities arising from this were not likely to change perioperative management. The patient members of the GDG also noted that an abnormal result may cause unnecessary stress. The Delphi survey also showed that clinicians would perform a resting ECG in all patients ASA2 and above undergoing intermediate surgery. The GDG commented that this was an inappropriate use of resources as the results of a resting ECG would be unlikely to alter the management of these patients. The GDG felt that it would be appropriate to consider preoperative resting ECG in ASA2 patients with cardiovascular, renal or diabetes comorbidities, as a baseline measurement. This was deemed to be useful information in the instance that a patient had a cardiac event during or after surgery. The GDG agreed that resting ECG should be obtained for ASA3 or ASA4 patients as this information would be helpful when considering the perioperative care of these patients. It was also noted that all patients would have resting ECG monitoring immediately before and during surgery as a final safety net. The GDG stated that if a patient has more than one comorbidity, the dominant condition should be given greater consideration. **Major or complex surgery** The results of the Delphi survey showed that clinicians would consider performing a resting ECG on ASA1 patients undergoing major or complex surgery. The GDG discussed this result and decided that a resting ECG in this population would be of limited use as a baseline measurement, and was unlikely to alter perioperative management. The GDG agreed with the Delphi survey that resting ECG should be obtained for ASA2 patients and ASA3 or ASA4 patients with obesity, diabetes, respiratory, cardiovascular or renal comorbidities undergoing major or complex surgery. The GDG therefore recommended the use of resting ECG preoperatively in all patients with ASA2 or above undergoing major or complex surgery. It was noted that these patients are at the most risk of surgical complications and an abnormal result is likely to change their perioperative management. It is therefore likely to be cost-effective to carry out a resting ECG in these patients. **Age** The GDG noted that the ESA 2014 guideline[25] recommended a resting ECG for patients over 65 undergoing major or complex surgery, however they noted that the guideline was not evidence based. The GDG discussed patients over 65 years of age and agreed they are at greater risk of asymptomatic changes that would be highlighted by a resting ECG and that may require preoperative treatment. The GDG therefore decided that resting ECG should be considered in ASA1 patients over 65 undergoing major or complex surgery if no previous ECG results within the last 12 months were available. For people under the age of 65 clinical judgement should be used to determine whether a test is necessary and the GDG acknowledged that variation in the prevalence of cardiac problems would be a factor taken into consideration.
Economic	In the absence of evidence, unit costs were provided alongside the results of the

considerations	Delphi survey. The unit cost of a resting ECG was found to be £31 from the NHS reference costs. However, it was noted that this cost was likely obtained from a specialist setting. In preoperative assessment the resting ECG would likely be done at a lower cost. Taking into account equipment time and staff time, it was estimated that the cost of a resting ECG could be as little as £8.66.If a resting ECG identifies an untreated cardiac condition then the costs of long-term management will be incurred. Downstream consequences of an abnormal resting ECG result could include an echocardiogram, ECG monitoring or an outpatient cardiology visit. These consequences have high costs, ranging from £65 to £222. Preoperative resting ECG testing would not be cost-effective if it does not change preoperative management that reduces the risk of related surgical complications. Therefore, resting ECG has been offered to individuals who are most at risk of undiagnosed heart conditions and are having surgery that will be most affected by these conditions.

6 Resting echocardiography

6.1 Introduction

Resting echocardiography is a non-invasive test that can assess ventricular function, heart valve anatomy and function and regional wall motion abnormalities (suggestive of previous myocardial infarction). In the preoperative setting, resting echocardiography is used to predict how the heart may respond to the physiological stress of surgery and to formulate a safe perioperative management plan for the patient. However resting echocardiography only assesses the heart at rest, not while under stress, and requires a skilled operator to perform and interpret the results. In addition, there is uncertainty about whether having a preoperative resting echocardiogram impacts on postoperative morbidity and mortality, especially in asymptomatic individuals.

See section 4.4 for a summary of the methodological approach taken for this preoperative test.

6.2 Review question (intervention): What is the usefulness of resting echocardiography as a preoperative test in altering perioperative management for adults and young people with mild to severe comorbidities undergoing major or complex elective surgery?

Table 29: PICO characteristics of review question

Population	All adults and young people ASA2 or above at risk of cardiovascular disease undergoing major or complex non-cardiac elective surgery Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory and renal diseases, obesity, diabetes
Intervention	Preoperative resting echocardiography
Comparison(s)	No preoperative resting echocardiography
Outcomes	**Critical:** • Change in healthcare management (for example cancellation of surgery or treating ischaemia, valvular disease or heart failure on the basis of the results of the tests). **Important:** • All-cause mortality • Complications related to surgery or anaesthesia • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing (time of testing) • Health-related quality of life • Intensive care unit (ICU) admission • Composite outcomes such as major adverse cardiovascular events (MACE) that incorporate cardiac deaths and non-fatal cardiac events • Optimisation of medical therapy

6.3 Clinical evidence

One retrospective non-randomised study (Wijeysundera 2011[94]) and one observational non-randomised study (Poso 2014[81]) were included in the review; these are summarised in Table 30 below. Evidence from these studies is summarised in the GRADE clinical evidence summary tables below (Table 31 and Table 32). See also the clinical study selection flow chart in Appendix E, clinical evidence tables in Appendix H, forest plots in Appendix J, GRADE tables in Appendix I and excluded clinical studies list in Appendix K.

Table 30: Summary of intervention studies included in the review

Study	Intervention and comparison	Population	Outcomes	Comments
Poso 2014 [81]	Intervention=preoperative resting echocardiography Control=no preoperative resting echocardiography	n=46. Echocardiography= 26 No echocardiography= 20 Morbidly obese subjects scheduled for bariatric surgery by laparoscopic Roux-en-Y gastric bypass surgery	Mean length of hospital stay (days) 30-day mortality	Non-randomised observational study. Not blinded. High risk of bias due to attrition and selection.
Wijeysundera 2011[94]	Intervention=preoperative resting echocardiography Control=no preoperative resting echocardiography	Preoperative adult population undergoing non-cardiac surgery n=70996 Echocardiography group=35498 No echocardiography group=35498 Propensity matching.	Length of stay 30-day mortality 30-day surgical site infection	Propensity matching. Retrospective. High risk of bias. Large sample size.

Table 31: Clinical evidence summary: Bariatric surgery

Outcomes	Number of participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with no echocardiography	Risk difference with preoperative resting echocardiography (95% CI)
Mean hospital stay (days)	46 (1 study)	VERY LOW due to risk of bias, imprecision	Not applicable	Only mean difference provided	The mean length of hospital stay in the intervention groups was 0.7 higher (0.13 to 1.53 higher)
30-day mortality	46 (1 study)	VERY LOW due to risk of bias, imprecision	Not applicable	Could not be meta-analysed as no event rate	Could not be meta-analysed as no event rate

[a] Downgraded by 2 increments if the majority of the evidence was at very high risk of bias

[b] Downgraded by 1 increment if the confidence interval crossed 1 MID or by 2 increments if the confidence interval crossed both MIDs

Table 32: Clinical evidence summary: Non-cardiac surgery

Outcomes	Number of participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with no echocardiography (non-cardiac surgery)	Risk difference with preoperative resting echocardiography (95% CI)
30-day mortality	70996 (1 study)	VERY LOW[a,b] due to risk of bias, imprecision	RR 1.14 (1.02 to 1.27)	17 per 1000	2 more per 1000 (from 0 more to 5 more)
Length of hospital stay	70996 (1 study)	VERY LOW[a] due to risk of bias	Not applicable	Only mean difference provided	The mean length of hospital stay in the intervention groups was 0.31 higher (0.17 to 0.45 higher)
Surgical site infection	70996 (1 study)	VERY LOW[(a)] due to risk of bias	RR 1.03 (0.98 to 1.08)	129 per 1000	4 more per 1000 (from 3 fewer to 10 more)

[(a)] Downgraded by 1 increment if the majority of the evidence was at high risk of bias

Outcomes	Number of participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with no echocardiography (non-cardiac surgery)	Risk difference with preoperative resting echocardiography (95% CI)
[(b)] Downgraded by 1 increment if the confidence interval crossed 1 MID or by 2 increments if the confidence interval crossed both MIDs					

6.4 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

Unit costs

In the absence of published economic evidence, unit costs were presented to the GDG to aid consideration of cost effectiveness. These are reported in the table below. Please see Appendix M for details.

Table 33: Cost of a resting echocardiogram

Test	Unit Cost (£)	Source
Resting echocardiogram	£75	NHS Reference Costs 2012-2013 (Code RA60A)[29]

Economic considerations

It is likely that patients who receive 'abnormal' results will be referred to a specialist. This is an additional cost that will need to be taken into account alongside the cost of the echocardiogram. The cost of one hour of a medical consultant's time is £99. The cost of a cardiology outpatient visit is £133 according to NHS reference costs 2012–2013.[29]

The clinical evidence suggests there is no clinical benefit derived from routine echocardiography testing. This is either because the prevalence of 'abnormal' results is minimal or that a significant portion of 'abnormal' results receive no differential management. However, it is likely that this evidence underestimates the true benefit of echocardiography testing due to confounders in the analysis.

6.5 Evidence statements

6.5.1 Clinical

For forest plots, see Section J.2 in Appendix J.

Bariatric surgery

One very low quality non-randomised observational study comprising 46 participants comparing preoperative resting echocardiography with no preoperative resting echocardiography demonstrated that the mean length of hospital stay was 0.7 higher in the group that had preoperative resting echocardiography. The evidence was at high risk of bias and showed serious imprecision. The same study reported zero events in both the intervention and control groups when looking at 30-day mortality. This evidence was also subject to a high risk of bias.

Non-cardiac surgery

One very low quality retrospective non-randomised observational study comprising 70996 participants compared preoperative resting echocardiography with no preoperative resting echocardiography. The study demonstrated no clinical benefit of preoperative resting echocardiography compared to control. The evidence was at high risk of bias and showed imprecision. The study also reported an increase in mean length of hospital stay in the group that

underwent preoperative resting echocardiography. This evidence was at high risk of bias. The study also showed that preoperative resting echocardiography did not demonstrate a clinical benefit for surgical site infection compared to control. This evidence was at high risk of bias.

6.5.2 Economic

No relevant economic evaluations were identified.

6.6 Recommendations and link to evidence

Recommendations	**5. Do not routinely offer resting echocardiography before surgery.** **6. Consider resting echocardiography if the person has:** • **a heart murmur <u>and</u> any cardiac symptom (including breathlessness, pre-syncope, syncope or chest pain) <u>or</u>** • **signs or symptoms of heart failure.** **Before ordering the resting echocardiogram, carry out a resting electrocardiogram (ECG) and discuss the findings with an anaesthetist.**
Relative values of different outcomes	The GDG considered change in healthcare management to be a critical outcome. All-cause mortality, complications related to surgery or anaesthesia, length of hospital stay after an operation, hospital readmission, adverse events caused by testing, health-related quality of life, ICU admission, optimisation of medical therapy and composite outcomes such as major adverse cardiovascular events were considered to be important outcomes.
Trade-off between clinical benefits and harms	The GDG discussed the evidence and noted the marginally longer length of stay in patients undergoing resting echocardiography. While the evidence was limited, they noted that there was no clinical benefit in conducting a preoperative echocardiogram on every patient. The GDG considered the consequences of delayed surgery, which may occur due to the additional time required for a patient to have a resting echocardiogram. They noted that unnecessary delays could potentially lead to poorer perioperative outcomes. The GDG also noted the importance of taking a history, performing a clinical examination and interpreting a resting 12-lead ECG in order to appropriately indicate whether to perform a resting echocardiogram or not. The GDG felt that there is sometimes a tendency to over-request echocardiograms based on a heart murmur alone. Therefore, the GDG determined that the presence of a cardiac murmur alone should not lead to a resting echocardiogram, as many cardiac murmurs are clinically insignificant. Instead, heart murmurs should be considered in conjunction with cardiac symptoms which indicate that the murmur reflects a haemodynamically significant cardiac problem affecting a valve or the outflow tract of the ventricle. These haemodynamically significant problems would necessitate, if present, alterations of management including perioperative management. Similarly, the presence of signs or symptoms of heart failure indicates the need for echocardiography to rule out cardiac dysfunction, or confirm it, at which point pharmacological and other interventions would be necessary prior to surgery. In all these subgroups, there is an increased likelihood of a cardiac problem that would be detectable by echocardiography and the perioperative management may well be altered accordingly. The GDG noted that the list of symptoms included in the recommendation is fairly comprehensive but not necessarily exhaustive. Echocardiograms are also frequently requested preoperatively in patients with breathlessness or poor exercise tolerance. There is a need for the referral for

	echocardiography and perioperative management of these patients to be standardised. The GDG also discussed the need for healthcare professionals referring patients for resting echocardiograms to be trained in correctly interpreting reports, and aware of the impact on perioperative management.
Economic considerations	No economic evaluations were identified that assessed the cost-effectiveness of resting echocardiograms. The unit cost of a resting echocardiogram was found to be £75. The GDG noted that there would also be significant follow-up costs with patients being referred for a cardiology opinion if the result was abnormal. There was no clinical evidence that suggested doing this test for all preoperative patients would improve patient outcomes. The GDG were concerned that in current practice resting echocardiograms are overused, resulting in high costs to the NHS with no improvement in patient outcomes, and in some cases causing unnecessary delays to surgery as well as increasing patient anxiety. The GDG therefore felt that their use should be limited to a targeted population who are most likely to have cardiac disease and therefore would most likely benefit from undergoing this test. Using resting echocardiograms on this targeted population would represent a cost-effective use of NHS resources as identifying cardiac disease and altering perioperative management accordingly is likely to improve patient outcomes.
Quality of evidence	All of the evidence was graded very low quality. Two retrospective cohort studies were included in addition to one prospective non-randomised study. One study propensity matched the cohort using a number of characteristics, including age, income, gender, comorbidities, preoperative consultation type, preoperative testing types, hospital type, surgical procedure and perioperative care. However, the retrospective design of the study is still limiting. The GDG discussed that the reason for overall mortality favouring the control over the echocardiography group may be explained by the fact that those in the echocardiography group were likely to have undergone echocardiography due to the identification of cardiac risk factors and/or poor functional capacity. Therefore the echocardiography group may have been a group with a higher baseline surgical risk, where a higher mortality rate would be expected. The second retrospective study had a very low sample size and reported incomplete data, leading to a very low quality rating. The prospective study was not randomised or blinded and was subject to a high degree of attrition. The GDG commented that due to this it would not be appropriate to base any recommendation on this evidence. The three studies could not be grouped for meta-analysis as they reported different surgery types.
Other considerations	The GDG commented that GPs should share recent echocardiograms when referring their patients for surgery, in order to prevent unnecessary duplication of investigations (see recommendation 1: communicating test results) The GDG also noted that a resting ECG should be performed and the results shared with an anaesthetist before considering further testing with echocardiography.

7 Cardiopulmonary exercise testing (CPET)

7.1 Introduction

Cardiopulmonary exercise testing (CPET) incorporates measurement of gas exchange variables, 12-lead electrocardiography, pulse oximetry, heart rate, and intermittent non-invasive blood pressure during exercise, typically using a cycle ergometer. In the preoperative setting, CPET is used to characterise an individual's functional capacity and to predict whether they will tolerate the physiological stress provoked by surgery. A major advantage of CPET is that it integrates assessment of cardiac, respiratory and metabolic variables in a situation mimicking that of surgery. It is also considered a safe test, with the risks the same as for mild-moderate exercise. However, it involves specialised facilities that are not universally available, is time-consuming, and requires a skilled practitioner to perform and analyse the test. In addition, there is uncertainty about the predictive value of CPET on perioperative morbidity and mortality, and also about how CPET results should be used in the clinical environment to inform preoperative optimisation and perioperative management.

See section 4.4 for a summary of the methodological approach taken for this preoperative test.

7.2 Review question (intervention): What is the clinical- and cost-effectiveness of using cardiopulmonary exercise test (CPET) as a preoperative test in improving patient outcomes in adults and young people with mild to severe comorbidities undergoing major or complex non-cardiac elective surgery?

Table 34: PICO characteristics of review question

Population	All adults and young people classified as ASA2 or above undergoing: • Major or complex non-cardiac elective surgery Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory and renal diseases, obesity, diabetes
Intervention	Cardiopulmonary exercise test (CPET)
Comparison(s)	No CPET test/clinical assessment only
Outcomes	**Critical:** • All-cause mortality • Health-related quality of life **Important:** • Complications related to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney failure, infection) • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing • Intensive care unit (ICU) admission

7.3 Clinical evidence

A single retrospective cohort study was included in the review (Goodyear 2013); this is summarised in Table 35 below. See also the clinical study selection flow chart in Appendix E, forest plots in Appendix J, GRADE tables in Appendix I, clinical evidence tables in Appendix H and excluded clinical studies list in Appendix K.

Table 35: Summary of studies included in the review

Study	Intervention and comparison	Population	Outcomes	Comments
Goodyear 2013	CPET versus no CPET	n=316 CPET=188 Historical control=128 Adult patients undergoing open abdominal aortic aneurysm (AAA) repair or endovascular aneurysm repair (EVAR)	• Length of inpatient stay • Duration of ITU stay • 30-day mortality	• Very high risk of bias due to retrospective nature of control group. • Non-randomised.

Table 36: Clinical evidence summary: Open AAA surgery

Outcomes	Number of participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with prior to recommendations	Risk difference with after implementation (95% CI)
Length of inpatient stay[c]	316 (1 study)	GRADE quality could not be assessed as no imprecision data			
30-day mortality	316 (1 study)	VERY LOW[a,b] due to risk of bias, imprecision	RR 0.32 (0.11 to 0.94)	126 per 1000	86 fewer per 1000 (from 8 fewer to 112 fewer)

[a] Downgraded by 2 increments if the majority of the evidence was at very high risk of bias

[b] Downgraded by 1 increment if the confidence interval crossed 1 MID or by 2 increments if the confidence interval crossed both MIDs

[c] The length of inpatient stay following open AAA surgery in the CPET era (median: 10 days, 95% CI: 10.3 to 13.5) was shorter than in the pre-CPET control era (median: 13 days, 95% CI: 13.9 to 19.0; p<0.001)

Table 37: Clinical evidence summary: EVAR AAA surgery

Outcomes	Number of participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with prior to recommendations	Risk difference with after implementation (95% CI)
Length of inpatient stay[c]	316 (1 study)	GRADE quality could not be assessed as no imprecision data			
30-day mortality	316 (1 study)	VERY LOW[a,b] due to risk of bias, imprecision	Peto OR 3.91 (0.05 to 329.71)	0 per 1000	145 more per 1000 (0 fewer to 77 more)

[a] Downgraded by 2 increments if the majority of the evidence was at very high risk of bias

[b] Downgraded by 1 increment if the confidence interval crossed 1 MID or by 2 increments if the confidence interval crossed both MIDs

[c] The length of inpatient stay following EVAR in the CPET era (median: 4.0 days, 95% CI: 4.6 to 6.7) was shorter than that in the pre-CPET era control (median: 6.0 days, 95% CI: 5.3 to 8.6; $p<0.05$)

7.4 Review question (prognostic): Does cardiopulmonary exercise testing (CPET) predict prognosis (patient outcomes after surgery) in adults and young people with mild to severe comorbidities undergoing major or complex non-cardiac elective surgery?

Table 38: Characteristics of review question

Population	All adults and young people with mild to severe comorbidities (classified as ASA2 or above) undergoing major or complex non-cardiac elective surgery. Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory and renal diseases, obesity, diabetes
Prognostic test	**Cardiopulmonary exercise test measures, including:** • VO_2 (oxygen uptake) • Peak VO_2 (highest value during test) • VO_2 max (maximal oxygen uptake) • VCO_2 (carbon dioxide exhaled) • AT – anaerobic threshold (exercise capacity) • VE/VO_2 and VE/VCO_2 – ventilatory equivalents
Key confounding factors	• Age • Comorbidities
Outcomes	**Critical:** • All-cause mortality **Important:** • Complications relating to surgery or anaesthesia • Length of hospital stay • Hospital readmission • Adverse events caused by testing • Health-related quality of life • Intensive care unit (ICU) admission

7.5 Clinical evidence

Sixteen observational studies reporting evidence of moderate to very low quality were included in the review; these are summarised in Table 39 below.

Evidence is presented by surgery type when two or more studies have reported results for the same surgery type; otherwise the results are summarised in the section 'other types of surgeries'.

The evidence is summarised in the following sections of surgeries:

- abdominal aortic aneurysm (AAA) repair
- lung resection
- colorectal surgery
  - pancreaticoduodenectomy

- major intra-abdominal surgery
- other types of surgeries

Four main measures of patients' ergometric capacity (oxygen consumption) during CPET were reported: anaerobic threshold (AT), peak oxygen consumption (VO_2), ventilator equivalents for oxygen (VE/VO_2) and carbon dioxide (VE/VCO_2).

See also the clinical study selection flow chart in Appendix E, clinical evidence tables in Appendix H, forest plots in Appendix J, GRADE tables in Appendix I and excluded clinical studies list in Appendix K.

Table 39: Summary of studies included in the review

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
Abdominal aortic aneurysm (AAA) repair surgery						
Barakat 2015 [11]	Prospective cohort n=130 Surgery: aortic aneurysm repair. CPET values didn't influence operative approach.	Multivariable logistic regression	Peak oxygen consumption (peak VO_2) Ventilator equivalent for carbon dioxide (V_E/VCO_2) Anaerobic threshold (AT)	Age, sex, method of repair, and CPET parameters (peak VO_2, V_E/VCO_2 and AT).	Cardiac complications Pulmonary complications	Not adjusted for comorbidities
Carlisle 2007[17]	Prospective cohort n=130 Surgery: aortic aneurysm repair. CPET values didn't influence operative approach.	Multivariable cox regression analysis	Anaerobic threshold (AT), peak oxygen consumption (peak VO_2), ventilator equivalents for oxygen (VE/VO_2) and carbon dioxide (VE/VCO_2)	Revised cardiac risk index score (RCRI)	Survival at 35 months	Unclear whether patients were consecutive. Multivariate analysis did not include all confounders identified in review protocol.
Grant 2015 [43]	Prospective cohort n=506 Overlaps with Hartley 2012 [49] cohort Surgery: aortic aneurysm repair. All had surgery – no information on fitness.	Multivariable Cox proportional hazards analysis	Peak oxygen consumption (peak VO_2 <15 ml/kg/minute), ventilator equivalent for carbon dioxide (VE/VCO_2 >42)	Stratified on operation type (open, EVAR), and adjusted for sex, age, diabetes, inducible cardiac ischaemia, statin, elevated urea, creatinine haemoglobin, VE/VCO_2 at AT <42, peak VO_2<15 ml/ kg/minute	Survival at 3 years	Only variables with a p value <0.20 at univariate analysis were entered into multivariable analyses.

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
Hartley 2012[49]	Prospective cohort n=415 Surgery: aortic aneurysm repair. All had surgery – no information on fitness.	Multivariable logistic regression	Anaerobic threshold (AT <10.2 ml/kg/minute), peak oxygen consumption (peak VO_2 <15 ml/kg/minute), ventilator equivalents for oxygen (VE/VO_2) and carbon dioxide (VE/VCO_2 >42)	Age, gender, diabetes, ischaemic heart disease, hypertension, preoperative medication, urea, creatinine, surgery type (open versus EVAR)	30-day and 90-day mortality	High risk of bias due to inaccurate outcome reporting.
Prentis 2012[82]	Prospective cohort n=212 Surgery: aortic aneurysm repair. CPET values didn't influence operative approach.	Backward stepwise multivariable regression	Anaerobic threshold (AT threshold of 10 ml/minute/kg), VE, VO_2, VCO_2, VE/VCO_2, peak VO_2	Age, BMI, comorbidities, aneurysm size, preoperative blood tests, medications	Complications (cardiovascular, respiratory, renal, gastrointestinal, wound complications, neurological) Length of hospital stay	Only 'important clinical variables' identified in univariate analysis were entered into multivariate analysis (peak VO_2, AT and watts), therefore did not include all confounders identified in review protocol
Lung resection surgery						
Brunelli 2009[16]	Prospective cohort n=204 Surgery: lung resection. 204/263 fit for surgery based on CPET.	Stepwise logistic regression	Peak oxygen consumption (VO_2) VO_2 <10 VO_2 10–15 VO_2 15–20 VO_2 >20	Age, gender, BMI, forced expiratory volume in 1 second (FEV1), diffusing capacity of the lung for carbon monoxide (DLCO), predicted postoperative FEV1 (ppoFEV1), predicted postoperative	30-day pulmonary complication	Only variables with a p value <0.10 at univariate analysis were entered into logistic regression analyses. Multivariate analysis did not include confounders identified in review protocol.

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
				diffusing capacity of the lung for carbon monoxide (ppoDLCO), coronary artery disease, type of operation (lobectomy versus pneumonectomy), neoadjuvant chemotherapy		
Brunelli 2012[15]	Prospective cohort n=225 Surgery: major lung resection. CPET values didn't influence operative approach.	Stepwise logistic regression analysis	Minute ventilation-to-carbon dioxide output slope (VE/VCO_2 threshold <35)	Age, sex, BMI, FEV1%, DLCO%, ppoDLCO%, chronic obstructive pulmonary disease (COPD) status, smoking, type of operation	30-day pulmonary complications	Only variables p<0.1 at univariate analysis were used in stepwise logistic regression analysis. MVA did not include all confounders identified in review protocol.
Licker 2011[63]	Prospective cohort n=210 Surgery: lung resection. CPET values didn't influence operative approach.	Logistic regression analysis	Peak oxygen uptake (peak VO_2 ≤10 ml/kg/minute, 11–17 ml/kg/minute, 17–20 ml/kg/minute, ≥20 ml/kg/minute).	Anaesthesia duration, age, extension of lung resection (major or minor), BMI, tidal volume (VT)	All complications Cardiovascular complications Pulmonary complications	Only variables with univariate probability <0.20 or those judged to be clinically important were selected for inclusion in the multivariable analysis. MVA did not include all confounders identified in review protocol.

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
Torchio 2010[88]	Retrospective cohort n=145 COPD patients. Surgery: lung resection. Connection between CPET and surgery decision unclear.	Multiple logistic regression analysis	Ventilatory inefficiency (VE/VCO_2), minute ventilation (VE), peak oxygen uptake (peak VO_2), carbon dioxide output (VCO_2)	Age, BMI, spirometric and CPET parameters	30-day mortality Cardiopulmonary complications	Multivariable analysis did not include all confounders identified in review protocol
Colorectal surgery						
West 2014[93]	Prospective cohort n=136 Surgery: major colonic surgery. CPET values didn't influence operative approach.	Multivariable logistic regression	Oxygen uptake (VO_2 at thresholds <10.6 or ≥10.6), peak VO_2, VE/VCO_2	Gender, operation type (laparoscopic or open)	Any complications	Only statistically significant variables from univariate logistic regression or Fisher's exact tests were retained for multivariable analysis. The rest presented as median (IQR) and p value only. Multivariable analysis did not include all confounders identified in review protocol.
Pancreaticoduodenectomy						
Ausania 2012[10]	Retrospective cohort n=124 Surgery: pancreatico-duodenectomy. CPET values didn't influence operative approach.	Multivariable logistic regression	Anaerobic threshold (AT 10.1 ml/kg/minute)	BMI, history of jaundice, preoperative biliary stent, pancreatic duct size	Pancreatic leak	Only variables significant in univariable analysis progressed to multivariable analysis. Multivariable analysis did not include all confounders identified in review protocol.
Junejo 2014[55]	Prospective cohort n=143	Simple logistic regression	Anaerobic threshold (AT), maximal oxygen	Obstructive jaundice	In-hospital mortality	Only if more than 1 variable predicted an

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
	Surgery: pancreatico-duodenectomy. 64/89 fit for surgery based on CPET.		consumption (VO_2 max), ventilator equivalent for carbon dioxide (VE/VCO_2), end-tidal values for oxygen ($PETO_2$)		30-day mortality Cardiopulmonary complications Any complications	outcome in simple logistic regression were they added to multivariable regression. No CPET variable was a significant predictor by simple logistic regression. Unclear which variables were entered into the multivariable analysis.
Other surgery						
Junejo 2012 [56]	Prospective cohort n=108 Surgery: hepatic resection. CPET values didn't influence operative approach.	Multivariable logistic regression	Anaerobic threshold, VE/VCO_2, VE/VO_2, VE, $PETO_2$, peak VO_2	Age, BMI	Cardiopulmonary complications Any complications	Variables with p>0.10 in univariate analysis were excluded from a multiple regression model. Multivariable analysis did not include all confounders identified in review protocol.
McCullough 2006 [67]	Prospective cohort n=109 Surgery: bariatric surgery (Roux-en-Y gastric bypass). Connection between CPET and surgery decision unclear.	Stepwise, multiple logistic regression	VO_2, VE, VCO_2,respiratory exchange ratio RER VCO_2/VO_2, VE/VCO_2 slope, peak oxygen consumption (peak VO_2) VO_2 <15.8 ml/kg/minute VO_2 >15.8 ml/kg/minute	Age, gender	Complications (death, unstable angina, MI, DVT, PE, renal failure, stroke)	Only peak VO_2 included in multivariable analyses. BMI and eGFR were p>0.10 at linear regression so were not entered into multivariate analysis. Multivariable analysis did not include all confounders identified in review protocol.
Prentis 2013	Prospective cohort	Logistic regression	Anaerobic threshold	Age, sex, BMI	Postoperative	Only CPET variables were

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
[83]	n=82 Surgery: radical cystectomy. 69/82 fit for surgery based on CPET.		(12 ml/minute/kg), VE, VO_2, VCO_2, peak VO_2, VE/VCO_2		complications	entered into multivariable regression. Unclear measurement, analysis and outcome reporting.
Snowden 2010 [86]	Prospective cohort n=123 Surgery: major elective surgery. CPET values didn't influence operative approach.	Multivariable logistic regression	VCO_2, VO_2, heart rate, minute ventilation, work rate, AT, peak VO_2, VE/VCO_2	Veterans activity score index (VASI), BMI, revised cardiac risk index, creatinine, POSSUM, age	Any complications	Only factors shown to be significant (p <0.1) at univariate analysis were entered in multivariable analysis. These did include confounders identified in review protocol.

Table 40: Clinical evidence summary: Abdominal aortic aneurysm (AAA) repair surgery

Risk factor	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Anaerobic threshold for predicting 30-day mortality (adjusted ORs)	1	Adjusted OR [95% CI]: 6.35 [1.84, 29.80]	No serious imprecision	LOW[a]
Anaerobic threshold for predicting survival at 35 months (adjusted HRs)	1	Adjusted HR [95% CI]: 0.84 [0.73, 0.98]	No serious imprecision	LOW[a]
Anaerobic threshold for predicting cardiac complications (adjusted ORs)	1	Adjusted OR [95% CI]: 0.55 [0.37, 0.82]	No serious imprecision	LOW[a]
Anaerobic threshold for predicting respiratory complications (adjusted ORs)	1	Adjusted OR [95% CI]: 0.85 [0.62, 1.17]	Serious	VERY LOW[ab]
Anaerobic threshold for predicting all	1	Adjusted OR [95% CI]: 0.71 [0.67, 0.88]	No serious	LOW[a]

Risk factor	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
major complications (adjusted ORs)			imprecision	
Peak VO_2 for predicting 90-day mortality (adjusted ORs)	1	Adjusted OR [95% CI]: 8.59 [2.33, 55.75]	No serious imprecision	VERY LOW[ab]
Peak VO_2 for predicting 3-year survival (adjusted HRs)	1	Adjusted HR [95% CI]: 1.68 [1.0, 2.8]	No serious imprecision	MODERATE[a]
Peak VO_2 for predicting cardiac complications (adjusted ORs)	1	Adjusted OR [95% CI]: 1.03 [0.81, 1.31]	Serious	VERY LOW[ab]
Peak VO_2 for predicting pulmonary complications (adjusted ORs)	1	Adjusted OR [95% CI]: 0.89 [0.69, 1.15]	Serious	VERY LOW[ab]
VE/VCO_2 for predicting survival at 35 months (adjusted HRs)	1	Adjusted HR [95% CI]: 1.13 [1.07, 1.20]	No serious imprecision	LOW[a]
VE/VCO_2 for predicting survival at 3 years (adjusted HRs)	1	Adjusted HR [95% CI]: 1.63 [1.01, 2.63]	No serious imprecision	MODERATE[a]
VE/VCO_2 for predicting cardiac complications (adjusted ORs)	1	Adjusted OR [95% CI]: 0.96 [0.86, 1.07]	Serious	VERY LOW[ab]
VE/VCO_2 for predicting respiratory complications (adjusted ORs)	1	Adjusted OR [95% CI]: 1.18 [1.05, 1.33]	No serious imprecision	LOW[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

Table 41: Clinical evidence summary: Lung resection surgery

Risk factor	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Peak VO_2 for predicting pulmonary	1	Adjusted OR [95% CI]: 0.87 [0.76, 0.99]	No serious	LOW[a]

Risk factor	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
complications (adjusted ORs)			imprecision	
Peak VO_2 for predicting pulmonary complications (adjusted ORs)	1	Adjusted OR [95% CI]: 0.84 [0.75, 0.94]	No serious imprecision	LOW[a]
Peak VO_2 for predicting cardiovascular complications (adjusted ORs)	1	Adjusted OR [95% CI]: 0.80 [0.68, 0.92]	No serious imprecision	LOW[a]
Peak VO_2 for predicting all complications (adjusted ORs)	1	Adjusted OR [95% CI]: 0.79 [0.71, 0.88]	No serious imprecision	LOW[a]
Peak VO_2 for predicting cardiopulmonary complications (adjusted ORs)	1	Adjusted OR [95% CI]: 0.05 [0.01, 0.25]	No serious imprecision	LOW[a]
VE/VCO_2 for predicting respiratory complications (adjusted ORs)	1	Adjusted OR [95% CI]: 1.09 [1.03, 1.16]	No serious imprecision	LOW[a]
VE/VCO_2 for predicting 30-day mortality (Adjusted ORs)	1	Adjusted HR [95% CI]: 1.24 [1.06, 1.45]	No serious imprecision	LOW[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*

Table 42: Clinical evidence summary: Major colonic surgery

Risk factor	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Peak VO_2 increase of 2.0 ml/kg/minute for predicting any complication (adjusted ORs)	1	Adjusted OR [95% CI]: 0.60 [0.45, 0.80]	No serious imprecision	LOW[a]
Peak VO_2 increase of 1.0 ml/kg/minute for predicting any complication (adjusted ORs)	1	Adjusted OR [95% CI]: 0.77 [0.66,0.90]	No serious imprecision	LOW[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*

Table 43: Clinical evidence summary: Pancreaticduodenectomy

Risk factor	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Anaerobic threshold for predicting in-hospital mortality (adjusted ORs)	1	Adjusted OR [95% CI]: 1.32 [0.14, 12.43]	Serious	VERY LOW[ab]
Anaerobic threshold for predicting in-hospital mortality (adjusted ORs)	1	Adjusted OR [95% CI]: 0.90 [0.52, 1.56]	Serious	VERY LOW[ab]
Anaerobic threshold for predicting cardiorespiratory complications (adjusted ORs)	1	Adjusted OR [95% CI]: 2.88 [0.66,12.64]	Serious	VERY LOW[ab]
Anaerobic threshold for predicting cardiopulmonary complications (adjusted ORs)	1	Adjusted OR [95% CI]: 1.05 [0.82, 1.34]	Serious	VERY LOW[ab]
Anaerobic threshold for predicting all complications (adjusted ORs)	1	Adjusted OR [95% CI]: 3.73 [1.33, 10.51]	Serious	LOW[a]
Anaerobic threshold for predicting all complications (adjusted ORs)	1	Adjusted OR [95% CI]: 1.07 [0.83, 1.38]	Serious	VERY LOW[ab]
Anaerobic threshold for predicting pancreatic leak (adjusted ORs)	1	Adjusted OR [95% CI]: 5.79 [1.62, 20.69]	Serious	LOW[a]
Peak VO_2 for predicting in hospital mortality (adjusted ORs)	1	Adjusted OR [95% CI]: 1.03 [0.77, 1.38]	Serious	VERY LOW[ab]
Peak VO_2 for predicting 30-day mortality (adjusted ORs)	1	Adjusted OR [95% CI]: 1.32 [0.91, 1.91]	Serious	VERY LOW[ab]
Peak VO_2 for predicting cardiopulmonary complications (adjusted ORs)	1	Adjusted OR [95% CI]: 1.0 [0.86, 1.16]	Serious	VERY LOW[ab]
VE/VCO_2 for predicting 30-day mortality (adjusted ORs)	1	Adjusted OR [95% CI]: 1.35 [1.03, 1.77]	No serious imprecision	LOW[a]
VE/VCO_2 for predicting in hospital	1	Adjusted OR [95% CI]: 1.26 [1.05, 1.51]	No serious	LOW[a]

Risk factor	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
mortality (adjusted ORs)			imprecision	
VE/VCO_2 for predicting all complications (adjusted ORs)	1	Adjusted OR [95% CI]: 0.97 [0.89, 1.06]	Serious	VERY LOW[ab]
VE/VCO_2 for predicting cardiopulmonary complications (adjusted ORs)	1	Adjusted OR [95% CI]: 0.98 [0.90, 1.07]	Serious	VERY LOW[ab]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

Table 44: Clinical evidence summary: Other surgery types

Risk factor	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Anaerobic threshold for predicting major postoperative morbidity after radical cystectomy (adjusted ORs)	1	Adjusted OR [95% CI]: 0.74 [0.57, 0.96]	No serious imprecision	LOW[a]
Anaerobic threshold for predicting any complication after major elective surgery (adjusted ORs)	1	Adjusted OR [95% CI]: 0.44 [0.30, 0.64]	No serious imprecision	MODERATE[a]
Anaerobic threshold for predicting length of hospital stay after radical cystectomy (adjusted ORs)	1	Adjusted OR [95% CI]: 0.47 [0.28, 0.79]	No serious imprecision	LOW[a]
Peak VO_2 <15.8 ml/kg/minute for predicting complications after bariatric surgery (adjusted ORs)	1	Adjusted OR [95% CI]: 12.89 [1.14, 145.76]	No serious imprecision	LOW[a]
Peak VO_2 (continuous) for predicting complications after bariatric surgery	1	Adjusted OR [95% CI]: 1.61 [1.19, 2.18]	No serious	LOW[a]

Risk factor	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
(adjusted ORs)			imprecision	
VE/VCO_2 for predicting surgical complications after hepatic surgery (adjusted ORs)	1	Adjusted OR [95% CI]: 3.97 [1.44, 10.95]	No serious imprecision	LOW[a]
VE/VCO_2 for predicting cardiopulmonary complications after hepatic surgery (adjusted ORs)	1	Adjusted OR [95% CI]: 3.45 [1.31, 9.09]	No serious imprecision	LOW[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*

7.6 Economic evidence

Published literature

One economic evaluation was identified with the relevant comparison and has been included in this review.[41] This study is summarised in the economic evidence profile below (Table 45).

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

Table 45: Economic evidence profile: CPET versus no CPET

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost-effectiveness	Uncertainty
Goodyear 2013[41] (UK)	Partially applicable[a]	Potentially serious limitations[b]	• Retrospective cohort study with historical control • Population: abdominal aortic aneurism (AAA) patients (>=5.5 cm). • Historical control group	• **Open surgery**: ○ saves £4,408[c]	• **Deaths averted per patient:** Open surgery: 0.086 (p<0.05) See clinical evidence review for details.	**Open surgery**: CPET is dominant compared to no CPET	• Differences in cost were statistically significant (p<0.01). • Separate analysis using influenceable (variable) ward costs instead of the fully

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost-effectiveness	Uncertainty
			(pre-CPET) and CPET era cohort (divided into four subgroups: CPET-pass, CPET-fail, CPET-submaximal and no CPET)				absorbed costs was reported. The results showed a similar trend to the base case analysis. • Adding CPET cost as reported in the paper (£150): ○ Open AAA: saves £4,258 (calculated by NCGC) • Endovascular aneurysm repair (EVAR) group: (calculated by NCGC) ○ Saves £500 ○ Increases mortality (0.014 more deaths per patient)

(a) Patients' ASA grades not reported. Twenty-two percent of patients in the CPET era group did not have CPET. Not all costs are reported (only ward and ICU stays). Quality of life and long-term outcomes are not considered.

(b) Comparator data are obtained from a historical cohort and it is not clear which risk-stratification strategy was used in this group. Univariate analysis was used with no control for confounders. Costs could vary between interventions as a result of a change in clinical practice, not because of the intervention itself. Data on other important health outcomes, such as complications, adverse events due to testing and quality of life are not reported. Sensitivity analysis was not undertaken.

(c) Cost components reported in the paper included total non-operative inpatient costs, including ward and ITU stay.

Unit costs

In addition to the published economic evidence, unit costs were presented to the GDG to aid consideration of cost effectiveness. These are reported in the tables below.

Table 46: Unit cost of CPET

Currency description	National average unit cost (£)[a]
Complex lung function exercise testing	183

(a) Source: NHS National Schedule of Reference Costs 2012-2013.[29]

Table 47: NHS reference cost of adult critical care

Currency description	National Average Unit cost (£)[a]
Adult critical care, 6 or more organs supported	1,867
Adult critical care, 5 organs supported	1,697
Adult critical care, 4 organs supported	1,573
Adult critical care, 3 organs supported	1,422
Adult critical care, 2 organs supported	1,236
Adult critical care, 1 organ supported	852
Adult critical care, 0 organs supported	619

(a) Source: NHS National Schedule of Reference Costs 2012-2013.[29]

7.7 Evidence statements

7.7.1 Clinical

For forest plots, see Section J.3 in Appendix J.

7.7.1.1 Intervention review

Open AAA surgery

Evidence from one very low quality retrospective cohort study comprising 316 participants comparing preoperative CPET with a historic control receiving no preoperative CPET demonstrated that CPET decreased the length of inpatient stay from a median of 13 days (13.9 to 19.0) to a median of 10 days (10.3 to 13.5). The same study demonstrated that preoperative CPET reduced 30-day mortality. Evidence was at very high risk of bias and showed imprecision.

EVAR AAA surgery

Evidence from one very low quality retrospective cohort study comprising 316 participants comparing preoperative CPET with a historic control receiving no preoperative CPET demonstrated that CPET decreased the length of inpatient stay from a median of 6 days (5.3 to 8.6) to a median of 4 days (4.6 to 6.7). The same study demonstrated that preoperative CPET increased 30-day mortality. Evidence was at very high risk of bias and showed imprecision.

7.7.1.2 Prognostic review

7.7.1.2.1 Abdominal aortic aneurysm (AAA) repair surgery

Anaerobic threshold (AT)

Three prospective cohort studies with a total of 647 patients reported that a lower AT was predictive of increased mortality (measured by 30-day mortality or survival at 35 months) [Low quality].

Four prospective cohort studies reported on the predictive value of AT for complications. Two of these reported on cardiac complications and showed inconsistent results. One study of 130 patients found that a lower AT was predictive of increased risk for cardiac complications [Low quality]; whereas a study of 102 patients found no predictive value of AT [Very low quality]. AT was found not to be predictive of cerebrovascular complications or length of ICU stay [1 study, 102 patients, Very low quality] or respiratory complications [2 studies, 232 patients, Very low quality]; whereas a lower AT did predict an increased risk of all complications [2 studies, 314 patients, Low quality].

VO_2

Two overlapping prospective cohort studies of 506 and 415 patients reported that a higher peak VO_2 improved 90-day mortality and 3-year survival [Moderate to Low quality], but was not predictive of 30-day mortality [Low quality].

One further study of 130 patients found no predictive value of peak VO_2 for cardiac or pulmonary complications [Very low quality].

VE/VO_2

One prospective cohort study of 102 patients reported that VE/VO_2 had no predictive value for 30-day mortality, cardiac complications, cerebrovascular complications, respiratory complications or all complications [Very low quality].

VE/VCO_2

Two prospective cohort studies with a total of 636 patients reported that a lower VE/VCO_2 was predictive of increased 3-year survival [Low to Moderate quality]. However, one prospective study of 102 patients found no predictive value of VE/VCO_2 for 30-day mortality [Very low quality].

Two prospective cohort studies with a total of 232 patients reported on the predictive value of VE/VCO_2 for complications. No predictive value was found for cardiac, cerebrovascular or all complications, or for length of ICU stay [Very low quality]. Inconsistent results were found for respiratory complications, as for AT. One study of 130 patients found that a lower VE/VCO_2 was predictive of increased risk for cardiac complications [Very low quality]; whereas a study of 102 patients found no predictive value of VE/VCO_2 [Very low quality].

7.7.1.2.2 Lung resection surgery

VO_2

Four studies reported on the predictive value of VO_2 for postoperative complications. Three studies demonstrated that a lower peak VO_2 preoperatively was predictive of increased risk of postoperative complications (pulmonary [2 studies, 414 patients, Low quality], cardiovascular [1 study, 210 patients, Low quality], cardiopulmonary [2 studies, 356 patients, Low quality] and all complications [2 studies, 336 patients, Low quality]). One poorly reported retrospective cohort study of 110 patients reported no association of peak VO_2 with cardiopulmonary complications [Low quality].

VE/VCO_2

One prospective cohort study of 225 patients reported that a higher VE/VCO_2 was predictive of increased risk of respiratory complications [Low quality], while one poorly reported retrospective cohort study of 110 patients reported no association of VE/VCO_2 with cardiopulmonary complications [Low quality].

One retrospective cohort study of 145 COPD patients reported that a higher VE/VCO_2 was predictive of increased risk of 30-day mortality [Low quality].

7.7.1.2.3 Pancreaticoduodenectomy

AT

Two studies in a total of 267 patients reported that AT was not predictive of in-hospital mortality, cardiorespiratory complications or cardiorespiratory complications [Very low quality]. These 2 studies showed inconsistent findings for the effect of AT on all complications, with one study showing that lower AT increased the risk [Low quality] and the other showing no predictive value [Very low quality].

VO_2

One study in 143 patients reported that VO_2 max was not predictive of in-hospital mortality, 30-day mortality or cardiopulmonary complications [Very low quality].

VE/VCO_2

One study in 143 patients reported that higher VE/VCO_2 was predictive of increased risk of 30-day mortality and in-hospital mortality [Low quality], but not of all complications or cardiopulmonary complications [Very low quality].

7.7.1.2.4 Other surgery types

AT

One prospective cohort study in 82 patients showed that a lower AT was predictive of increased risk for major postoperative morbidity and increased length of stay after radical cystectomy [Low quality].

Two further prospective cohort studies from the same research group reported that a lower AT was predictive of increased risk for complications [123 patients, Moderate quality] and for postoperative mortality [389 patients; Low quality].

VO_2

One prospective cohort study of 109 patients demonstrated that a lower peak VO_2 is predictive of increased risk for postoperative complications after bariatric surgery [Low quality].

VE/VCO_2

One prospective cohort study of 108 patients demonstrated that a higher peak VE/VCO_2 is predictive of increased risk for cardiopulmonary or any complications [Low quality].

Peak VO_2

One prospective cohort study of 105 patients reported that a lower peak VO_2 was predictive of increased risk of in-hospital morbidity 5 days after surgery [Low quality].

7.7.2 Economic

One cost-consequences analysis found that using CPET as a risk stratification strategy prior to AAA surgery was dominant (more effective and less costly) in the open surgery group compared to no testing. This study was assessed as partially applicable with potentially serious limitations.

7.8 Recommendations and link to evidence

Recommendations	No recommendations made
Relative values of different outcomes	The GDG considered all-cause mortality and health-related quality of life to be critical outcomes for the intervention review. Complications related to surgery or anaesthesia, length of hospital stay after an operation, hospital readmission, adverse events caused by testing and ICU admission were considered to be important outcomes. For the prognostic review all-cause mortality was considered critical while all other listed outcomes from the intervention review were considered important.
Trade-off between clinical benefits and harms	The GDG considered whether the results of CPET enable the clinician to predict individual patient outcomes: that is, does poor CPET performance indicate likelihood of death or perioperative complications? A number of studies within the evidence indicated that this might be the case. The GDG were also interested in studies that demonstrated if perioperative management was altered in some way, based on the test, impacting on patient outcomes. In practice, various parameters are used to inform patient management decisions. All CPET measures need to be considered in the context of the surgical procedure that is being performed. For example, in lung cancer surgery, peak exercise capacity is deemed a more appropriate predictor of complications, whereas anaerobic threshold is considered more relevant for other surgery types. CPET results require careful interpretation by an experienced physiologist or clinician, otherwise a patient who is suitable for surgery may not be offered surgery (and vice versa). The GDG were concerned that incorrect assessments may be made regarding the risk of surgery based on the outcome of CPET, in particular when there is no strong evidence to support such assessments. The GDG agreed that CPET is one of several assessments that may be taken into consideration when having an informed discussion with patients about their overall surgical risk and treatment options. CPET is considered a safe test, with the risks the same as for mild-moderate exercise. Major adverse events including death, myocardial infarction, arrhythmia, haemodynamic instability and orthopaedic injury are reported in study populations at a rate of <1 to 5 per 10,000 tests.
Economic considerations	Only one economic evaluation was included. [41] This study was rated as partially applicable with potentially serious limitations. The study was a cost-consequences analysis that was also included in the clinical review. The study results showed that using CPET in the preoperative assessment of adults undergoing abdominal aortic aneurysm (AAA) repair surgery was more effective and less costly (that is, dominant) compared to no test, as it resulted in reduction of perioperative mortality, improved mid-term survival, reduced length of stay and lower costs. However, the GDG noted that the analysis considered non-operative costs only and did not consider the cost of the operation itself. The GDG agreed that the results of CPET would normally be used to identify patients for whom endovascular aneurysm repair (EVAR) would be more appropriate. This means that a CPET-fail result would normally lead to offering the patient a more costly intervention (EVAR). Additionally, it was not clear from the study which risk stratification methods were used in the pre-CPET era. Overall the GDG felt that the study results could not be generalised to other types of surgery or other populations. Additionally, the GDG highlighted the availability of cheaper alternatives to CPET, such as the 6-minute walk test, which may represent a more cost-effective use of resources and can be more suitable for those who may be less able to exercise, such as the elderly.

	The GDG also noted the variability in access to CPET among hospitals. Undertaking CPET testing requires the presence of trained staff who can confidently run the test and interpret and communicate the results to the preoperative assessment clinicians. Hence, making the test available in all hospitals will have cost implications. In the absence of strong evidence, the cost-effectiveness of this test is uncertain.
Quality of evidence	One retrospective study (with a historical control) comparing CPET with no CPET for patients undergoing AAA repair surgery demonstrated a benefit in terms of mortality and length of hospital stay for those receiving the test. However, the study had serious limitations: it used historic controls and conducted only univariate analysis. The GDG noted that there was a lack of comparative studies comparing CPET versus no CPET, and therefore requested prognostic data. The majority of the prognostic studies included were non-randomised prospective cohort single-centre studies, which did not adjust for confounding factors. The GDG recognised that the outcomes from these studies might be partially or wholly explained by the confounders, rather than the CPET measures. However, the GDG also acknowledged that CPET could in effect summarise all these confounders into a single measurement, such as oxygen consumption at anaerobic threshold, diminishing its independence yet making it a useful overall indicator of cardiopulmonary fitness for surgery. The GDG agreed that, based on the evidence, there is not enough robust evidence to recommend or not recommend CPET testing before surgery.
Other considerations	Cardiac surgery was not included in the scope of this guideline and therefore studies on cardiac or other types of surgery intimately connected to the heart were excluded. However certain types of vascular surgery (including peripheral vascular and AAA surgery) do fall within the remit of the scope and were included when evidence was found. The GDG considered the question of how to evaluate patients who are unable to fully or partially complete the CPET test to be a very important limitation. Patients having CPET need to be aware of the risks associated with the test (major adverse events including death, myocardial infarction, arrhythmia, haemodynamic instability and orthopaedic injury are reported in study populations at a rate of <1 to 5 per 10000 tests). Validated information for patients, which is available nationally, would be useful. The GDG noted that there seems to be a drive towards the use of CPET nationally but that there are much simpler tests, for example the 6-minute walk test or shuttle walk test, that do not require specialist equipment and are cheaper. However, there is a lack of consensus regarding how clinicians should best determine how to assess a patient's fitness for surgery and individual risk. Further research is required to evaluate the best method of assessing a patient's perioperative risk by comparing a number of different tools, including the 6-minute walk test and the incremental shuttle walk test. The GDG noted that the METS study (Pearse et al) is ongoing; this compares CPET with a physician assessment of functional activity and the Duke Activity Status Index (DASI) questionnaire of physical activity. In addition, another UK-based multicentre study examining whether different levels of postoperative care according to CPET results influences outcomes has finished recruiting with results pending. Based on these ongoing studies, the GDG decided not to make recommendations or a research recommendation.

8 Chest X-ray

8.1 Introduction

Chest X-rays (formally known as chest radiographs) can detect diseases of the lungs, pleura, heart, major vasculature, mediastinum, chest wall and diaphragm. In the preoperative setting, chest X-rays are used to assess known chronic medical conditions or to detect previously undiagnosed diseases. Conditions that are frequently detected in this setting include chronic obstructive pulmonary disease (COPD), heart failure, tuberculosis and lung cancers. However, chest X-rays involve exposure to a dose of radiation and are of questionable benefit in asymptomatic individuals, in whom the rate of abnormality detection is low. In addition, there is uncertainty concerning whether chest X-ray findings impact on perioperative management and whether rates of perioperative pulmonary complications are affected by the performance of a preoperative chest X-ray.

See section 4.4 for a summary of the methodological approach taken for this preoperative test.

8.2 Delphi survey results

As no new evidence on the use of chest X-ray as a routine preoperative test was identified during the scoping phase of this guideline, it was decided not to carry out an evidence review, but to include chest X-ray in the modified Delphi survey in order to re-evaluate the consensus held amongst health professionals on the value of routinely conducting the test before elective surgery.

The survey participants were asked if they would use chest X-ray as a routine preoperative test for patients undergoing elective surgery. As the consensus results clearly showed at the first round that chest X-ray was not used by over 70% of respondents, no further questions on the use of chest X-ray were asked in subsequent rounds of the survey. Please see Appendix L for full details on the survey method and results.

8.2.1 Delphi statements where consensus was reached

Table 48: All surgery grades and comorbidities (cardiovascular, renal, respiratory, diabetes, obesity)

	Results % (round in which consensus was achieved)
I do use chest X-ray tests for routine elective surgery	23.20%
I do not use chest X-ray tests for routine elective surgery	74.74%
I do not have the expertise to answer	2.06%

8.3 Economic evidence

Unit costs were provided for consideration alongside the Delphi survey results. Please see Appendix M for details. These are reported in Table 49 below.

Table 49: Unit cost of chest X-ray

Equipment/staff	Quantity (units/minutes)	Cost	Source
Direct access plain film[(a)]	1	£29.60	NHS reference costs 2013-14[30]

Equipment/staff	Quantity (units/minutes)	Cost	Source
	Total per patient	£29.60	

(a) Includes medical and staffing cost involved in the procedure

8.4 Recommendations and link to evidence

Recommendations	7. Do not routinely offer chest X-rays before surgery.
Delphi	Delphi survey respondents strongly agreed that chest X-ray should not be used as a routine preoperative test in any population. The GDG agreed with the findings of the Delphi survey and considered that patient management would not be altered by chest X-ray results for any patient population considered within the guideline. The GDG noted the risk of ionising radiation associated with chest X-ray. This provided the GDG with further rationale to restrict the use of chest X-ray in routine preoperative testing.
Economic considerations	The cost of performing a chest X-ray was found to be £29.60, including staff time and equipment. The GDG did not feel that the results of a chest X-ray would often improve health outcomes as results were unlikely to influence any management strategies. The Delphi survey also showed a strong consensus against the use of chest X-rays, therefore supporting the notion that the test provides minimal clinical benefit. It is therefore unlikely that a chest X-ray is a cost-effective preoperative test.

9 Polysomnography

9.1 Introduction

Polysomnography is used to diagnose and monitor treatment responsiveness in obstructive sleep apnoea (OSA) and other sleep disorders. Formal polysomnography is conducted in a hospital setting and involves monitoring parameters including pulse oximetry, electroencephalography (EEG), surface electromyography (EMG), respiratory effort, electrooculography (EOG) and electrocardiography (ECG) during a night's sleep. Simpler sleep studies may also be performed by issuing an individual with a sleep study device to use in their own home. In the preoperative setting, polysomnography is used to diagnose OSA and institute appropriate management with the aim of reducing postoperative morbidity and mortality. Polysomnography is non-invasive and safe, with the only recognised complication being self-limiting skin irritation from electrodes. However, it is time-consuming and requires a skilled practitioner to run and interpret the tests. In addition, there is uncertainty regarding the impact of preoperative polysomnography on perioperative outcomes and how to identify which patients may benefit from polysomnography in the preoperative setting.

See section 4.4 for a summary of the methodological approach taken for this preoperative test.

9.2 Review question (intervention): What is the clinical- and cost-effectiveness of using polysomnography as a preoperative test (to detect obstructive sleep apnoea) in improving patient outcomes in adults and young people with obesity undergoing major or complex elective non-cardiac surgery?

Table 50: PICO characteristics of review question

Population	All adults and young people with obesity (ASA2 or above) undergoing major or complex elective non-cardiac surgery. Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade
Intervention	Polysomnography
Comparison(s)	No polysomnography
Outcomes	**Critical outcomes:** • All-cause mortality **Important outcomes:** • Complications related to surgery or anaesthesia • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing • Health-related quality of life • Intensive care unit (ICU) admission • Optimisation of therapy • Change in management

9.3 Clinical evidence

One observational study was included in the review. [22] The study examined obese patients scheduled for elective surgery that had preoperative polysomnography and compared postoperative outcomes with those who had no preoperative polysomnography. See also the clinical study selection flow chart in Appendix E, forest plots in Appendix J, GRADE tables in Appendix I, clinical evidence tables in Appendix H and excluded clinical studies list in Appendix K.

Table 51: Summary of observational studies included in the review

Study	Population	Intervention	Comparison	Outcome	Limitations
Chung 2008[22]	n=416 Obese patients aged 18 years or older who had an ASA physical status of I-IV and were scheduled to undergo elective procedures in general surgery, gynaecology, orthopaedics, urology, plastic surgery, ophthalmology, or neurosurgery.	Polysomnography	No polysomnography	Respiratory complications Cardiac complications Neurological complications Unplanned ICU admission Readmission within 30 days	Non-randomised observational study. Patients were not matched at baseline for significant confounders such as smoking. The paper did not state what effect the polysomnography results had on perioperative care. However, if the apnoea-hypopnoea index (AHI) of a patient was >30/hour the anaesthesiologist and surgeon were informed. Whilst good definitions of postoperative complications were given, the outcomes were extracted from medical notes by a blinded research anaesthesiologist and assumed initial accuracy. Contained a small percentage of patients who had undergone neurosurgery which is considered an indirect population.

Table 52: Clinical evidence summary: Elective surgery

Outcomes	Number of participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with prior to recommendations	Risk difference with after implementation (95% CI)

Outcomes	Number of participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with prior to recommendations	Risk difference with after implementation (95% CI)
Respiratory complications	416 (1 study)	VERY LOW[a,b,c] due to risk of bias, imprecision, indirectness	RR 1.43 (0.96 to 2.06)	12.2%	52 more per 1000 (from 5 fewer to 129 more)
Cardiac complications	416 (1 study)	VERY LOW[a,b,c] due to risk of bias, imprecision, indirectness	RR 1.94 (0.74 TO 5.08)	2.9%	27 more per 1000 (from 8 fewer to 118 more
Neurological complications	416 (1 study)	VERY LOW[a,b,c] due to risk of bias, imprecision, indirectness	RR 0.65 (0.11 TO 3.84)	1.5%	5 fewer per 1000 (from 13 fewer to 43 more)
Unplanned ICU admission	416 (1 study)	VERY LOW[a,b,c] due to risk of bias, imprecision, indirectness	RR 3.26 (0.56 to 19)	0.5%	11 more per 1000 (from 2 fewer to 82 more)
Readmission within 30 days	416 (1 study)	VERY LOW[a,b,c] due to risk of bias, imprecision, indirectness	RR 0.78 (0.21 to 2.85)	2.4%	5 fewer per 1000 (from 19 fewer to 44 more)

[a] Downgraded by 2 increments if the majority of the evidence was at very high risk of bias

[b] Downgraded by 1 increment if the confidence interval crossed 1 MID or by 2 increments if the confidence interval crossed both MIDs

[c] The evidence included a small percentage of patients undergoing neurosurgery

9.4 Review question (prognostic): Does polysomnography predict prognosis (patient outcomes after surgery) in adults and young people with obesity undergoing major or complex elective non-cardiac surgery?

Table 53: PICO characteristics of review question

Population	All adults and young people with obesity (ASA2 or above) undergoing (major or complex elective non-cardiac surgery. Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory, renal, obesity
Prognostic test	Polysomnography
Key confounding factors	*Minimum set of confounders that should be adjusted for (will vary per outcome)* • Comorbidities • BMI • Older age • Male • Hypertension
Outcomes	**Critical:** • All-cause mortality **Important:** • Complications relating to surgery or anaesthesia • Length of hospital stay (post-operation) • Hospital readmission • Adverse events after surgery (wound infection) • Health-related quality of life • Intensive care unit (ICU) admission

9.5 Clinical evidence

A single retrospective cohort study was included in the review. [92] It examined cohorts of people who had polysomnography prior to surgery and compared postoperative outcomes between those who had an apnoea–hypopnoea index (AHI) above 5 (considered a positive test) with those who had an AHI below 5 (considered a negative test). Evidence from the study is summarised in Table 55 below. See also the clinical study selection flow chart in Appendix E, forest plots in Appendix J, clinical evidence tables in Appendix H and excluded clinical studies list in Appendix K.

Table 54: Summary of prognostic studies included in the review

Study	Population	Analysis	Prognostic variable	Confounder list	Outcomes	Limitations
Weingarten 2011[92]	Retrospective cohort Single centre n=797 patients over 18 years old who had first-time bariatric surgery and were referred from preoperative testing clinic to have poly-somnography at 1 attached centre only.	Logistic regression adjusting for covariates. Patients divided into two groups: AHI≥5 and AHI<5.	AHI≥5 on poly-somnography	Age Sex Operative approach (laparoscopic or open) BMI	Pulmonary complications Surgical complications Other complications Any complication	Outcomes extracted from medical notes (assuming initial accuracy). All anaesthetists aware of poly-somnography results and likely treated patients differently in peri- and postoperative settings. Assumed compliance with the use of positive airway ventilation but not measured. Only included patients who chose to have poly-somnography performed at a specific site.

Table 55: Clinical evidence summary: Bariatric surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
AHI >5 for predicting postoperative pulmonary complications (adjusted ORs)	1	Adjusted OR [95% CI]: 1.0 [0.44, 2.27]	Serious imprecision	LOW[ab]
AHI >5 for predicting postoperative surgical complications (adjusted ORs)	1	Adjusted OR: 1.33 [0.79, 2.24]	Serious imprecision	LOW[ab]
AHI >5 for predicting other postoperative complications (adjusted ORs)	1	Adjusted OR: 0.79 [0.49, 1.27]	Serious imprecision	LOW[ab]
AHI >5 for predicting all postoperative	1	Adjusted OR: 0.86 [0.59, 1.25]	Serious imprecision	LOW[ab]

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
complications (adjusted ORs)				

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

9.6 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

Unit costs

In the absence of published economic evidence, unit costs were presented to the GDG to aid consideration of cost effectiveness. These are reported in Table 56 below.

Table 56 - Unit cost of polysomnography

Test	Unit cost (£)	Setting	Source
Polysomnography	£227	Outpatient	NHS references costs 2013-14 HRG code: DZ50Z[29]
	£195	Directly accessed diagnostic service	

9.7 Evidence statements

9.7.1 Clinical

For forest plots, see Section J.4 in Appendix J.

9.7.1.1 Intervention review

One very low quality non-randomised observational study comprising 416 participants demonstrated no clinical benefit of preoperative polysomnography on postoperative respiratory complications, cardiac complications, neurological complications, unplanned ICU admission, and readmission within 30 days, compared to no preoperative polysomnography. All outcomes were at high risk of bias and showed imprecision and indirectness.

9.7.1.2 Prognostic review

One retrospective study reported no increased risk in pulmonary complications (defined as aspiration, pneumonia, new requirement for CPAP or biPAP, use of naloxone, postoperative tracheal intubation, mechanical ventilator support or respiratory arrest) in patients with an AHI of 5 or greater on polysomnography prior to bariatric surgery (OR=1.00, 95% CI 0.44–2.30, p=0.992) [Low quality].

One retrospective study also reported no increased risk in surgical complications (bleeding, wound dehiscence, anastomotic leak, wound infection or the need for reoperation) in those with an AHI of 5 or greater on polysomnography prior to surgery (OR=1.33, 95% CI 0.79–2.25, p=0.284) [Low quality].

One retrospective study also reported no increased risk in other complications (myocardial infarction, dysrhythmia, stroke, thromboembolic events, sepsis, liver failure, acute kidney injury, hospital readmission, or death within 30 postoperative days) in those with an AHI of 5 or greater on polysomnography prior to surgery (OR=0.79, 95% CI 0.49–1.25, p=0.310) [Low quality].

The same retrospective study also reported no increased risk in postoperative complications in those who tested with an AHI of 5 or greater on polysomnography prior to bariatric surgery (OR= 0.86, 95% CI 0.59–1.29, p=0.47) [Low quality].

9.7.2 Economic

- No relevant economic evaluations were identified.

9.8 Recommendations and link to evidence

Recommendations	**Research recommendations:** **1. Does preoperative screening of people who are at risk of obstructive sleep apnoea (OSA) with polysomnography identify those at higher risk of postoperative complications?** **2. Does treating OSA perioperatively improve outcomes?**
Relative values of different outcomes	The GDG considered all-cause mortality to be a critical outcome for the intervention and prognostic reviews. Complications related to surgery or anaesthesia, length of hospital stay after an operation, hospital readmission, adverse events caused by testing, health-related quality of life and ICU admission were considered to be important outcomes for both the intervention and prognostic reviews; with the addition of optimisation of therapy and change in management as important outcomes for the intervention review.
Trade-off between clinical benefits and harms	The comparative evidence demonstrated no clinical benefit of preoperative polysomnography for any clinical outcome and was of very low quality. The GDG therefore decided to consider the prognostic evidence in order to make a recommendation. The single prognostic studies found no association between patients with obstructive sleep apnoea (OSA) diagnosed on preoperative polysomnography and postsurgical clinical outcomes. The GDG discussed the clinical evidence and noted from their experience that preoperatively optimising patients with OSA through the use of continuous positive airway pressure (CPAP) devices had a positive impact in both improving postoperative outcomes and general quality of life. This improvement in postoperative outcome was reflected in one paper included in this review, but not measured in the other studies. The GDG also commented on the potential benefit they have noticed in their own practice of diagnosing OSA using polysomnography prior to surgery, informing perioperative care changes. The GDG noted poor adherence to CPAP treatment in a large amount of patients, which was reflected in the evidence. The GDG were unsure, given the large non-adherence to treatment, whether having a diagnosis provided by polysomnography actually improves perioperative management. The GDG commented on the delay that might be incurred prior to surgery whilst waiting for polysomnography to be undertaken. The GDG discussed other means of screening (STOP Bang, Epworth sleepiness scale) for OSA that were quicker and cheaper that may be of benefit, but they felt that polysomnography is still the definitive preoperative test to diagnose OSA. The GDG felt that test tools could suggest a patient has OSA, but could not diagnose it. Despite this, there was no evidence to indicate that polysomnography provides a clinical benefit to allow the GDG to make a recommendation. They therefore chose

	to make a research recommendation.
Economic considerations	The GDG were presented with the unit cost of a polysomnography test taken from the 2014 NHS reference costs. In a preoperative setting, the polysomnography would most likely be performed in an outpatient setting and therefore the unit cost associated with this test was found to be £227. The GDG noted that this was high compared to other preoperative tests. Comparative evidence showed no benefit of polysomnography, however this was of very low quality, meaning the clinical effectiveness could not inform the cost-effectiveness of the test. Although evidence from the prognostic review showed that OSA was associated with worse postsurgical outcomes, the extent to which treated OSA lead to better postsurgical outcomes could not be inferred. There was no clinical evidence on polysomnography to demonstrate the health benefits needed to justify its high cost. A research question was therefore set out to obtain the clinical evidence needed to indicate whether or not polysomnography represents an efficient use of NHS resources.
Quality of evidence	The evidence for the reported outcomes was very low quality. A single interventional study was found, comparing the incidence of postoperative complications in obese patients undergoing polysomnography and no polysomnography before elective surgery. The method of patient allocation to either group was not randomised or blinded. Furthermore, the study did not adjust for significant confounders, specifically smoking. The GDG commented that due to the very low quality of this paper, it would not be appropriate to base any recommendation on the results. Only a single prognostic study presented multivariate assessment analysis. It was unclear from the study whether the surgeon or the anaesthetist were made aware of the findings of the polysomnography, and if so whether they had treated the patient differently during surgery compared to the control group. The GDG commented that the type of surgery may not be relatable to all major or complex (grade 3 or 4) surgeries as, given the known high prevalence of OSA in this group, this is a highly specialised type of surgery with strong existing guidance and protocols on the peri- and postoperative management of these patients. The GDG also commented that the cut-off for diagnosis of OSA in both studies was an apnoea–hypopnoea index (AHI) >5 and they felt that this was much lower than the threshold they apply in clinical practice.
Other considerations	The GDG discussed other disadvantages of polysomnography including the large waiting lists that some sleep study centres have that can cause long delays for surgery, adding to the patient's morbidity in this time. The GDG were aware of one study, (Mutter et al., 2014[70]) that suggests diagnosing OSA using polysomnography and treatment prior to surgery does improve cardiac postoperative outcomes, compared to patients who were diagnosed with OSA using polysomnography after surgery. However, the study included a mixed population of patients undergoing cardiac and non-cardiac surgery and did not meet the protocol. The GDG commented on the use of questionnaires as a screening tool for OSA in the preoperative setting and using the results to base referrals for polysomnography. The GDG also discussed the different kinds of polysomnography that are available and the different settings they can be performed in, ranging from basic night time pulse oximetry measured at home, to a complex series of investigations including electroencephalograms undertaken in sleep clinics. These have a variety of different costs and benefits that were discussed by the GDG.

	The GDG noted that there was not enough good quality evidence to recommend polysomnography as a preoperative test for obese patients. They discussed the clinical importance of this topic and noted the need for a better evidence base and awareness amongst healthcare professionals on which to make further decisions. The GDG therefore agreed to make a research recommendation.

10 Lung function tests

10.1 Introduction

Lung function tests can assess lung volumes, capacities, rates of flow and gas exchange, enabling the diagnosis and monitoring of respiratory diseases. In the preoperative setting, lung function tests are used to assess individuals with known or suspected respiratory disease. Tests used include spirometry, which measures inhaled and exhaled lung volumes and flow over time, as well as more sophisticated tests to measure static lung volumes and the diffusing capacity of the lungs. Spirometry is simple and easy to perform, but the other tests are more time-consuming and require specialised equipment. The tests are considered safe, although some individuals may experience light-headedness during the test, and testing can precipitate asthmatic episodes. The results must be interpreted carefully as all lung function tests are effort dependent. There is uncertainty about which patients with known or suspected respiratory disease require preoperative lung function tests, what time period prior to surgery these are required, and whether spirometry or more sophisticated tests are indicated. The impact of testing on perioperative morbidity and mortality is also not established.

See section 4.4 for a summary of the methodological approach taken for this preoperative test.

10.2 Health technology assessment 2012

The starting point of this review was the Health Technology Assessment (HTA)[24] published in 2012, that investigated routine testing before elective surgery in patients with no apparent clinical indication (and in subgroups with common comorbidities). The HTA 2012 was based on an earlier HTA published in 1997[69], but the protocols and outcomes differ.

Although the protocol for the HTA 2012 only considered comparative studies for inclusion, the included studies have a prognostic design. It concluded that there was insufficient evidence to indicate that the included tests are clinically- and cost-effective in patients without apparent clinical indications.

This update did not restrict studies to minor surgery and patients with no apparent clinical indications, but included a wider patient population (including people who are obese who were excluded from the HTA) and all surgery types that are specified in the scope of the guideline. This meant that some of the excluded studies from the HTA were included in this review. Otherwise, we used the same approach as the HTA in that we aimed to find comparative studies in the first instance. In case of an absence of such evidence, prognostic studies were also considered.

10.3 Review question (intervention): What is the usefulness of lung function tests in predicting outcome or altering perioperative management for adults and young people undergoing any type of elective surgery?

Table 57: Characteristics of review question – comparative protocol

Population	All adults and young people classified as patients ASA grade 1 to 4 undergoing: • Minor, intermediate, or major or complex surgery Stratified analysis if data available for: • Surgery type or surgery grade (if specified)

	• ASA grade • Selected comorbidities: cardiovascular, respiratory, renal, obesity, diabetes.
Intervention	Lung function tests (also including blood gas analysis)
Comparison(s)	No lung function tests
Outcomes	**Critical:** • All-cause mortality • Complications relating to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney injury, infection) **Important:** • Change in healthcare management (for example cancellation of surgery) • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing • Health-related quality of life • Intensive care unit (ICU) admission

10.4 Clinical evidence

No relevant clinical studies comparing preoperative resting lung function testing with no preoperative lung function testing were identified.

10.5 Review question (prognostic): Do lung function tests (also including blood gas analysis) predict prognosis (patient outcomes after surgery) in adults and young people ASA 1–4 undergoing any type of elective non-cardiac surgery?

Table 58: Characteristics of review question – prognostic protocol

Population	Adults and young people classified as patients ASA grade 1 to 4 undergoing: • Minor, intermediate, or major or complex surgery Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory, renal, obesity, diabetes
Prognostic test	Lung function tests (also including blood gas analysis)
Key confounding factors	• Age • Comorbidities
Outcomes	**Critical:** • All-cause mortality • Complications relating to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney injury, infection) • Change in healthcare management (for example cancellation of surgery) • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing (time of testing) • Health-related quality of life

- Intensive care unit (ICU) admission

10.6 Clinical evidence

Two studies[46,,53] were included in the review; these are summarised in Table 59.

See also the clinical study selection flow chart in Appendix E, clinical evidence tables in Appendix H, forest plots in Appendix J, GRADE tables in Appendix I and excluded clinical studies list in Appendix K.

Table 59: Summary of studies included in the review

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
Lung function tests						
Hamoui et al., 2006[46]	Prospective cohort. n=146 consecutive, morbidly obese patients who had duodenal switch operation surgery (bariatric surgery) during a 12-month period.	Multivariable logistic regression using variables identified as significant in the univariate analysis	Vital capacity (VC), functional residual capacity (FRC) and total lung capacity (TLC), forced vital capacity (FVC) and forced expiratory volume in 1 second (FEV1), maximal voluntary ventilation (MVV), and pO_2	• Age • Sex • BMI • Variables identified as statistically significant on univariate analysis	Postoperative complications	ASA and surgery grade not stated Uncommon type of surgery
Jeong et al., 2013[53]	Retrospective cohort. n=538 patients who underwent elective gastric cancer surgery. Mixed ASA grades with some patients in grade 3.	Multivariate analysis was carried out using variables that were significant at the p≤0.05 in the univariate analysis as covariates	Normal/abnormal lung function test: defined based on FEV1/FVC ratios and FEV1 values with FEV1/FVC ≥0.7 classified as normal; FEV1/FVC <0.7, FEV1 ≥80% predicted as mild; FEV1/FVC <0.7, FEV1 50–80% predicted as moderate; (FEV1/FVC <0.7, FEV1 30-50% predicted as severe; FEV1/FVC <0.7, FEV1 <30% predicted as very severe.	Postoperative surgical complications: • Age • Resection type • Operative approach • Tumour node metastasis stage Postoperative systemic complications: • Age • History of pulmonary disease	Postsurgical complication Systemic complication	Outcomes not clearly defined

Table 60: Clinical evidence summary: Bariatric surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
10% decrease in vital capacity for predicting all postoperative complications	1	Adjusted OR [95% CI]: 2.29 (2.20 to 2.38)	No serious imprecision	Low[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*

Table 61: Clinical evidence summary: Cancer surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Abnormal lung function test (based on FEV1/FVC ratios and FEV1 values) for predicting surgical postoperative complications	1	Adjusted OR [95% CI]: 1.75 (1.03 to 2.97)	No serious imprecision	Low[a]
Abnormal lung function test (based on FEV1/FVC ratios and FEV1 values) for predicting surgical postoperative complications	1	Adjusted OR [95% CI]: 1.11 (0.32 to 3.85)	Serious imprecision	Very low[ab]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

10.7 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

Unit costs

Unit costs were provided for consideration alongside the Delphi survey results. Please see Appendix M for details. These are reported in Table 62 below.

Table 62: Unit cost of lung function test

Equipment/staff	Quantity (units/minutes)	Cost	Source
Lung function test	1	£66	HTA 2012[24]
	Total per patient	£66	

10.8 Evidence statements

10.8.1 Clinical

For forest plots, see Section J.5.1 in Appendix J.

10.8.1.1 Intervention review

No relevant studies were identified.

10.8.1.2 Prognostic review

Bariatric surgery

One retrospective cohort study of 146 patients reported that postoperative respiratory complications were not predicted by the percentage values of forced vital capacity (FVC), forced expiratory volume in 1 second (FEV1) and TV, but no statistics were provided [Low quality].

One prospective cohort study of 146 patients investigated the following pulmonary test measures: vital capacity (VC), functional residual capacity (FRC) and total lung capacity (TLC), forced vital capacity (FVC) and forced expiratory volume in 1 second (FEV1), maximal voluntary ventilation (MVV), and pO_2. They reported that for each 10% decrease in vital capacity, the risk of postoperative complications increased more than two-fold [Low quality]. The remaining tests were not found to predict risk of complications [Low quality].

Cancer surgery

One retrospective cohort study of 538 patients compared abnormal with normal lung function tests. They reported an increased risk of postoperative surgical complications in those with abnormal findings [Very low quality]. However, lung function tests were not predictive of postoperative systemic complications [Low quality].

One retrospective cohort study of 213 patients reported that an abnormal spirometry result was not an independent predictor of postoperative complications after laparoscopy-assisted gastrectomy [Very low quality].

10.8.2 Economic

No relevant economic evaluations were identified.

10.9 Delphi survey results

Preoperative lung function tests were included in the modified Delphi consensus survey.

The survey participants were asked if lung function tests and blood gases should be used as a routine preoperative test for patients undergoing elective surgery. Participants rated their response from strongly disagree to strongly agree using a nine-point Likert scale. Each question was considered to have reached consensus if greater than 70% of responses were in a single category (0–3 strongly disagree, 4–6 unclear, 7–9 strongly agree). Please see Appendix L for full details on the survey method and results.

10.9.1 Delphi statements where consensus was reached

10.9.1.1 Lung function tests

Table 63: Patients: ASA2 with respiratory comorbidity

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	89.80 strongly disagree (round 1)
Intermediate surgery	71.43 strongly disagree (round 1)

Table 64: Patients: ASA3 or ASA4 with respiratory comorbidity

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	72.60 strongly disagree (round 1)
Major or complex surgery	73.24 strongly agree (round 3)

Table 65: Patients: ASA2 with cardiovascular, diabetes, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	89.58 strongly disagree (round 1)
Intermediate surgery	79.45 strongly disagree (round 1)

Table 66: Patients: ASA3 or ASA4 with cardiovascular, diabetes, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	80.55 strongly disagree (round 1)
Intermediate surgery	70.83 strongly disagree (round 1)

10.9.1.2 Blood gases

Table 67: Patients: ASA2 with obesity

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	98.69 strongly disagree (round 1)
Intermediate surgery	94.77 strongly disagree (round 1)
Major or complex surgery	85.62 strongly disagree (round 1)

Table 68: Patients: ASA3 or ASA4 with obesity

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	90.85 strongly disagree (round 1)
Intermediate surgery	84.31 strongly disagree (round 1)
Major or complex surgery	74.51 strongly disagree (round 1)

Table 69: Patients: ASA2 with respiratory comorbidity

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	95.39 strongly disagree (round 1)
Intermediate surgery	89.33 strongly disagree (round 1)
Major or complex surgery	78.28 strongly disagree (round 1)

Table 70: Patients: ASA3 or ASA4 with respiratory comorbidity

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	89.95 strongly disagree (round 1)
Intermediate surgery	80.79 strongly disagree (round 1)

Table 71: Patients: ASA2 with cardiovascular or renal comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	97.36 strongly disagree (round 1)
Intermediate surgery	96.00 strongly disagree (round 1)
Major or complex surgery	86.09 strongly disagree (round 1)

Table 72: Patients: ASA3 or ASA4 with cardiovascular or renal comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	96.02 strongly disagree (round 1)
Intermediate surgery	92.00 strongly disagree (round 1)
Major or complex surgery	88.74 strongly disagree (round 1)

10.9.2 Delphi statements where consensus was not reached

10.9.2.1 Lung function

Table 73: Patients: ASA2 with respiratory comorbidity

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Major or complex surgery	1	45.89	34.24
	2	24.05	51.9
	3	21.13	43.66

Table 74: Patients: ASA3 or ASA4 with respiratory comorbidity

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Intermediate surgery	1	49.32	23.28
	2	21.05	39.47
	3	26.76	30.97

Table 75: Patients: ASA2 with cardiovascular, diabetes, renal or obesity comorbidities

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Major or complex surgery	1	62.76	20.00
	2	54.54	19.48
	3	46.48	30.99

Table 76: Patients: ASA3 or ASA4 with cardiovascular, diabetes, renal or obesity comorbidities

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Major or complex surgery	1	54.11	32.19
	2	35.9	43.59
	3	37.14	40.0

10.9.2.2 Blood gases

Table 77: Patients: ASA3 or ASA4 with respiratory comorbidity

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Major or complex surgery	1	69.07	21.70
	2	19.74	51.32
	3	36.37	33.34

10.10 Recommendations and link to evidence

Recommendations	**8. Do not routinely offer lung function tests or arterial blood gas analysis before surgery.**

	9. Consider seeking advice from a senior anaesthetist as soon as possible after assessment for people who: • **are ASA grade 3 or 4 due to known or suspected respiratory disease and** • **are having intermediate or major or complex surgery.**
Relative values of different outcomes	The GDG considered all-cause mortality to be a critical outcome for the intervention and prognostic reviews. Change in healthcare management, complications related to surgery or anaesthesia, length of hospital stay after an operation, hospital readmission, adverse events caused by testing, health-related quality of life and ICU admission were considered to be important outcomes for the intervention and prognostic reviews.
Trade-off between clinical benefits and harms	Two types of surgery were covered by the evidence: bariatric surgery and gastric cancer surgery. For bariatric surgery in people with severe obesity, respiratory complications were not independently predicated by lung function tests, whereas rates of all postoperative complications were more common in people with normal preoperative lung function test results. The GDG noted that an anaesthetist sees people before surgery if they have an existing condition, and if this was severe, they would refer the patient to a specialist prior to surgery.
Economic evidence	No economic evaluations were identified for lung function testing.
Quality of evidence	The evidence was mainly of low or very low quality. The ideal evidence would have been testing as an intervention rather than a prognostic factor. However, no such evidence was identified. The prognostic evidence is problematic since it is often unclear whether test results led to management changes prior to surgery or whether the physician was aware of the test results prior to surgery. Even though we restricted evidence to studies using multivariable analyses to identify test results as independent factors leading to postsurgical outcomes, some studies only adjusted for a minimum number of characteristics, and in other studies it was unclear which factors were accounted for. The studies using bariatric surgery were not conducted in the UK, and the specific surgery types used are uncommon in the UK. It is therefore difficult to know whether the results would be similar in an NHS setting. Furthermore, study results were not consistent; one study reported lung function tests to be an independent predictor, and the other not an independent predictor, of complications. Inconsistent patterns of results were also seen in people having tests prior to gastric cancer surgery. Confidence that the results are representative is therefore low. For this reason, the GDG agreed to add this review question to the Delphi survey.
Other considerations	The GDG noted that recommendations for lung function tests in this guideline are not applicable to thoracic surgery, where spirometry is frequently performed as part of the diagnostic and surgical planning process prior to surgery. Thoracic surgical units have their own specialist guidelines for spirometry indications. The GDG discussed whether having access to spirometry results would be of benefit in pre-optimising some patients prior to surgery. For example, performing surgery on patients with poorly controlled asthma poses considerable perioperative risks, and performing spirometry may be useful to identify these patients and institute appropriate therapies to improve asthma control. However in the case of asthma, clinical history is usually more informative, since it is an intermittent and fluctuating

	disease. Spirometry might also highlight patients that would benefit from critical care management postoperatively, even if there are no modifiable spirometry abnormalities (for example in morbid obesity with functional lung volume loss or irreversible chronic respiratory disease). The GDG emphasised the importance of good clinical assessment in the preoperative setting prior to requesting tests and referring patients for preoperative optimisation. The GDG discussed the importance of referring GPs sharing information about a patient's respiratory disease and previous spirometry results. NICE asthma guidelines recommend annual spirometry testing for patients with asthma, and GPs should share these results with the preoperative assessment clinic to avoid unnecessary duplication of tests. The GDG discussed the need for full lung function tests versus spirometry alone in the preoperative setting. The general consensus was that spirometry alone was sufficient for guiding perioperative management in most patients. The GDG noted that their draft consensus recommendations were similar to those from the initial guideline. This is because there had been no significant changes in the indications for spirometry and strategies for respiratory pre-optimisation since the initial guideline. The GDG also noted a systematic review (Johansson 2013)[54] with specific tests for the respiratory system (including spirometry, chest X-ray and blood gases) that concluded there was no high quality evidence for routine preoperative testing in otherwise healthy adults undergoing elective non-cardiac surgery. This did not conflict with the included evidence. Due to the limited quality and applicability of the identified evidence, the GDG agreed that lung function tests should be entered into the Delphi consensus survey.
Delphi	**Lung function tests** The GDG considered the Delphi survey results and agreed with the consensus survey view that lung function tests should not be routinely considered. They commented that the test was commonly used to provide reassurance both to the patients and clinician before surgery, but had no real impact on perioperative or general clinical management. This may explain why some areas did not achieve a 'do not use' consensus, particularly in patients undergoing major surgery. The GDG noted where consensus was achieved in favour of carrying out the test, namely in those with respiratory problems undergoing major or complex surgery. The GDG then discussed the use of the test in patients with respiratory comorbidities, and felt that lung function tests do not offer much value above risk stratification even in ASA3 or 4 for major or complex surgery. The GDG also pointed out that for many patients with severe chronic respiratory disease, optimisation would not be possible and thus the benefit of the test would be limited. Moreover, the test would potentially require multiple outpatient visits. However, the GDG accepted that lung function tests could have a role in optimising the patient before surgery, leading to better postoperative management under advice from a senior anaesthetist. The GDG felt this would have the added value of encouraging earlier anaesthetist input in preoperative testing.

	Arterial blood gas analysis The GDG agreed with the results of the Delphi survey, which did not support the use of blood gas analysis for any populations. They indicated that patients with an obesity or respiratory comorbidity undergoing major surgery may be candidates for blood gas analysis, but that this was unlikely to lead to a change in management. The GDG therefore made a similar recommendation as for lung function tests, encouraging direction from a senior anaesthetist before offering the test.
Economic considerations	In the absence of evidence, unit costs were provided alongside the results of the Delphi survey. The cost of performing simple lung function tests (spirometry) was found to be £66. It was noted that the NHS reference cost of full lung function testing was £174, however the GDG agreed that full lung function testing was unnecessary in this circumstance. No clinical evidence was identified that helped determine whether undertaking lung function tests leads to better postoperative outcomes. Without knowing the clinical effectiveness of the test, the cost-effectiveness remains uncertain. The only clinical evidence found was prognostic and although this suggested that impaired lung function identified patients with poorer surgical outcomes, it could not be shown whether knowing the results preoperatively could improve post-surgical outcomes. The GDG also considered the cost of further investigations that may be prompted by abnormal lung function and blood gas analyses such as a CT scan (£71–£146) or a respiratory outpatient visit (£150). The GDG noted that for individuals ASA3 or above and a known respiratory condition (such as asthma or COPD), knowing the lung function results could have an impact on postoperative outcomes. However, the GDG felt that a senior anaesthetist would be in the best position to decide whether testing would inform management and therefore influence post-surgical outcomes. The cost of arterial blood gas analysis was found to be £6.42–£9.84. The GDG noted that abnormalities are only likely to occur in individuals who are severely symptomatic, limiting the need for the test as they would receive a change in management regardless. Routinely the GDG did not feel this test would have any impact on management and therefore health outcomes.

11 Full blood count test

11.1 Introduction

The full blood count test is used in the preoperative setting to detect anaemia, bleeding disorders, inherited and acquired haematological disorders, and the effects of other systemic diseases. The results may be used to plan the use of blood products and blood salvage techniques in the perioperative period. The test is considered safe and is straightforward to perform and analyse, but may be painful for the patient. There is a low risk of complications including haematoma formation, vasovagal reactions and infection. However, detection of abnormalities is uncommon in asymptomatic, fit and healthy individuals and there is uncertainty about the clinical effectiveness of performing routine preoperative full blood count testing in all individuals.

See section 4.4 for a summary of the methodological approach taken for this preoperative test.

11.2 Health technology assessment 2012

Full blood count test is part of the HTA 2012 update – see section 10.2 for details.

11.3 Review question (intervention): What is the usefulness of full blood count (haemoglobin, white blood cell count and platelet count) in predicting outcome or altering perioperative management for adults and young people undergoing any type of elective surgery?

Table 78: Characteristics of review question – comparative protocol

Population	Adults and young people classified as patients ASA grade 1 to 4 undergoing: • Minor, intermediate, or major surgery Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory, renal, obesity, diabetes
Intervention	Full blood count (haemoglobin, white blood cell count and platelet count)
Comparison(s)	No full blood count
Outcomes	**Critical:** • All-cause mortality • Complications relating to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney injury, infection) **Important:** • Change in healthcare management (for example cancellation of surgery) • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing • Health-related quality of life • Intensive care unit (ICU) admission

See section 10.2 for further details.

11.4 Clinical evidence

No relevant clinical studies comparing preoperative resting full blood count testing with no preoperative full blood count testing were identified.

11.5 Review question (prognostic): Do full blood count tests (haemoglobin, white blood cell count and platelet count) predict prognosis (patient outcomes after surgery) in adults and young people ASA 1–4 undergoing any type of elective non-cardiac surgery?

Table 79: Characteristics of review question – prognostic protocol

Population	Adults and young people classified as patients ASA grade 1 to 4 undergoing: • Minor, intermediate, or major or complex surgery Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory, renal, obesity, diabetes
Prognostic test	Full blood count tests (haemoglobin, white blood cell count and platelet count)
Key confounding factors	• Age • Comorbidities
Outcomes	**Critical:** • All-cause mortality • Complications relating to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney injury, infection) • Change in healthcare management (for example cancellation of surgery) • Length of hospital stay • Hospital readmission • Adverse events caused by testing (time of testing) • Health-related quality of life • Intensive care unit (ICU) admission

11.6 Clinical evidence

Ten studies[7,12,13,36,39,44,51,52,90,95] were included in the prognostic review; these are summarised in Table 80 below. Studies are further subdivided by surgery type and by relevant patient comorbidities:

- All elective surgery
- Orthopaedic surgery
  - Knee and hip arthroplasty
- Elective vascular surgery
- Cancer surgery
- All non-cardiac surgery

See also the clinical study selection flow chart in Appendix E, clinical evidence tables in Appendix H, forest plots in Appendix J, GRADE tables in Appendix I and excluded clinical studies list in Appendix K.

Table 80: Summary of studies included in the review

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
Full blood count						
Amaranto 2011[7]	Retrospective study n=1773 Adult vascular patients with normal preoperative WBC (3.5–10.5 K/microlitre) Surgery: carotid endarterectomy (CEA), carotid artery stenting (CAS), open repair of abdominal aortic aneurysm (AAA), endovascular repair of abdominal aortic aneurysm (EVAR), open repair of thoracoabdominal aortic aneurysm (TAAA), endovascular repair of thoracoabdominal aneurysm (TEVAR), lower extremity bypass grafting (LEB), or lower extremity stenting (LES)	Multivariable analysis – logistic regression	White blood cell count	• Age • Gender • Diabetes • Congestive heart failure • Myocardial infarction, • Renal insufficiency, • Hypertension, • Hyperlipidaemia • Emergent presentation	Complications:, stroke, MI, transient ischaemic attack (TIA), infection, bleeding, reoperation, and amputation Major adverse events: death, stroke, and MI	ASA not stated but 25% have diabetes, 84% have hypertension, 60% have hyperlipidaemia Surgery grade not stated
Beattie 2009[12]	Retrospective cohort study n=7679	Multivariable analysis – logistic regression using	Preoperative anaemia defined as: <13 g/litre for males,	• Age >70 years • In-hospital status • History of CHF	Mortality within 90 days	ASA and surgery grade not stated

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
	Surgery: included vascular and oncology surgery in head and neck, urology, thoracic, hepatobiliary, general, and gynaecological procedures	variables identified as significant on the univariate analysis	<12 g/litre for women	• Preoperative renal dysfunction • Perioperative medications: ○ No beta-blockers ○ Metoprolol ○ Atenolol or bisoprolol ○ ACE inhibitors ○ Calcium channel blockers • Postoperative NSAID • Transfusion: ○ No blood products ○ 1–2 units ○ 3–4 units ○ 5–10 units ○ >10 units		Mixed (all elective) surgeries: 35% major surgery
Bedke 2012[13]	Retrospective observational study: n=327 Surgery: partial or radical nephrectomy for clear cell RCC	Multivariable Cox model	White blood cell count measured 1–2 days before surgery	• TNM stage • Tumour size • Fuhrman grade • Karnofsky index • CRP • Leucocytes	Mortality	ASA and surgery grade not stated
Dunkelgrun et al. (2008)[36]	Retrospective cohort study n=1,211 Surgery: elective non-cardiac open vascular	Cox proportional hazard regression	Preoperative anaemia defined as: <13 g/litre for males <12 g/litre for women	• Anaemia • Renal dysfunction • Heart failure • Age • Gender • Type of vascular surgery (central or peripheral	30-day major adverse cardiac event (MACE) including nonfatal myocardial infarction and cardiac death	ASA and surgery grade not stated

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
	surgery			open procedure) • Diabetes mellitus • Chronic obstructive pulmonary disease • Hypertension • Ischaemic heart disease Stroke		
Glance et al. (2014)[39]	Retrospective cohort study n=316,644 Non-cardiac surgery	Multivariable analysis with multiple imputation for missing values of preoperative serum creatinine	Platelet count: moderate-to-severe thrombocytopenia (<100,000 μl^{-1}); mild thrombocytopenia (101,000–150,000 μl^{-1}); low–normal (151,000–200,000 μl^{-1}); normal (201,000–450,000 μl^{-1}); thrombocytosis (≥450,000 μl^{-1}).	• Haematocrit • Age • Sex • BMI (underweight, overweight, obesity, morbid obesity, and super obesity) • Admission source (home, transfer from other hospital, chronic care facility), race, inpatient status (versus outpatient), emergency status, surgical complexity (work-relative value units) • Previous operation within 30 days • Comorbidities: diabetes, pulmonary, cardiac, hypertension, peripheral vascular disease, renal disease, central nervous system	Receipt of any erythrocyte transfusion 30-day mortality 30-day complications: • cardiac • pulmonary • renal • central nervous system • sepsis • wound infection • graft failure	ASA and surgery grade not stated 18% of original sample excluded because no coagulation testing was performed (may over-estimate prognostic relevance)
Greenky et	Retrospective cohort	A multivariable	Preoperative anaemia	Demographic details:	Periprosthetic joint	ASA and surgery grade

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
al. (2012)[44]	study n=15,222 Surgery: total hip arthroplasty or total knee arthroplasty	logistic regression analysis A propensity score analysis generated through a regression model	defined as: <13 g/litre for males <12 g/litre for women	• Gender • Race • Age • BMI • Time in operating room • Surgery type (primary or revision) Comorbidities: • Atrial fibrillation • Congestive heart failure • Coronary artery disease • Hypercholesterolemia • Hypertension • Pneumonia • Renal failure • Renal transplant • Cerebrovascular disease • Diabetes • Gastroesophageal reflux disease • Peptic ulcer disease • Cancer • Coagulopathies • Systemic lupus • Peripheral vascular disease • Rheumatoid arthritis	infections 30-day mortality 90-day mortality 1-year mortality	not stated
Jamsen 2015[51]	Prospective cohort study n=191	A multivariable binary logistic	Preoperative anaemia defined as:	• Age • Sex	Hyperglycaemia during the	Small sample size

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
	Surgery: primary hip and knee replacements	regression analysis	<136 g/litre for males <117 g/litre for women	• Operated joint (hip, knee) • ASA risk score	hospitalization, defined as glucose >7.8 mmol/litre in two consecutive measurements Severe hyperglycaemia, defined as glucose >10 mmol/litre at any time point.	Unclear how/when haemoglobin measurements were made
Jans 2014[52]	Prospective observational cohort study n=5165 episodes, or 4940 unique patients. Surgery: unilateral primary total hip arthroplasty or total knee arthroplasty	Multivariate logistic regression used for the confounders listed	Preoperative anaemia defined as: <13 g/litre for males <12 g/litre for women	• Age • Procedure (THA versus TKA) • Female sex • Hypertension • Cardiac disease • Pulmonary disease • Cerebrovascular disease Preoperative walking aid	Risk of RBC transfusion during primary admission Length of stay >5 days All-cause readmission within 90 days after surgery	No information on whether preoperative Hb was acted on prior to operation or if the anaesthetist or surgeon were made aware of its value. The assessors extracting the information from computer databases and patient's notes were blinded to the patient's preoperative anaemic status. The study also reported that perioperative transfusion of red

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
						blood cells had a clinically significant increased risk of readmission and length of stay over 5 days. When perioperative transfusion was added to the multivariate analysis the preoperative Hb was no longer predictive of postoperative complications and length of stay. Mixed ASA grades with some grade 3 participants
Wang 2015	Retrospective cohort study	Multivariable Cox regression	Platelet count (≤178 x 10^9 l)	• Lymph node metastasis • TNM stage • Tumour location	Overall survival	ASA grade not stated
Yoshihara 2014[95]	Retrospective cohort study n=605,665 Surgery: total hip and knee arthroplasty	Multivariable logistic regression	Anaemia	• Age • Sex • Race • Comorbidity • Elixhauser Comorbidity Score • Autologous-related blood transfusion • Hospital size	Allogenic blood transfusion	ASA grade not stated Unclear how anaemia was recorded

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
				• Hospital caseload • Hospital region • Payer information		

Table 81: Clinical evidence summary: All surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Anaemia for predicting mortality in all anaemic patients	1	Adjusted OR [95% CI]: 2.36 (1.57 to 3.55)	No serious imprecision	Moderate[a]
Anaemia for predicting mortality in patients with severe anaemia	1	Adjusted OR [95% CI]: 1.79 (1.17 to 2.74)	No serious imprecision	Moderate[a]
Anaemia for predicting mortality in all patients (excluding those who received RBC transfusions)	1	Adjusted OR [95% CI]: 3.04 (1.80 to 5.13)	No serious imprecision	Moderate[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*

Table 82: Clinical evidence summary: Orthopaedic surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Anaemia for predicting RBC transfusion	1	Adjusted OR [95% CI]: 4.70 (3.80 to 5.81)	No serious imprecision	Moderate[a]
Anaemia for predicting allogeneic blood transfusion (hip arthroplasty)	1	Adjusted OR [95% CI]: 2.03 (1.86 to 2.22)	No serious imprecision	Moderate[a]

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Anaemia for predicting allogeneic blood transfusion (knee arthroplasty)	1	Adjusted OR [95% CI]: 2.70 (2.52 to 2.89)	No serious imprecision	Moderate[a]
Anaemia for predicting increased length of stay (>5 days)	1	Adjusted OR [95% CI]: 2.50 (1.90 to 3.29)	No serious imprecision	High
Anaemia for predicting readmission within 90 days	1	Adjusted OR [95% CI]: 1.40 (1.10 to 1.78)	No serious imprecision	Moderate[a]
Anaemia for predicting peri-prosthetic joint infections	1	Adjusted OR [95% CI]: 1.95 (1.41 to 2.70)	No serious imprecision	Low[a]
Anaemia for predicting 30-day mortality	1	Adjusted OR [95% CI]: 0.59 (0.10 to 3.59)	Serious imprecision	Very Low[ab]
Anaemia for predicting 90-day mortality	1	Adjusted OR [95% CI]: 1.54 (0.50 to 4.73)	Serious imprecision	Very Low[ab]
Anaemia for predicting 1-year mortality	1	Adjusted OR [95% CI]: 1.81 (1.00 to 3.29)	No serious imprecision	Low[a]
Anaemia for predicting hyperglycaemia	1	Adjusted OR [95% CI]: 3.90 (0.91 to 17.00)	Serious imprecision	Low[ab]
Anaemia for predicting severe hyperglycaemia	1	Adjusted OR [95% CI]: 2.00 (0.50 to 8.10)	Serious imprecision	Low[ab]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

Table 83: Clinical evidence summary: Vascular surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Anaemia for predicting major adverse cardiac event (mild anaemia)	1	Adjusted OR [95% CI]: 1.80 (0.80 to 4.05)	Serious imprecision	Low[ab]
Anaemia for predicting major adverse	1	Adjusted OR [95% CI]: 2.30 (1.10 to 4.81)	No serious	Moderate[a]

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
cardiac event (moderate anaemia)			imprecision	
Anaemia for predicting major adverse cardiac event (severe anaemia)	1	Adjusted OR [95% CI]: 4.70 (2.60 to 8.50)	No serious imprecision	Moderate[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

White blood cell

Table 84: Clinical evidence summary: Vascular surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
White blood cell count for predicting surgical complications (endovascular cohort)	1	Adjusted OR [95% CI]: 1.32 (1.11 to 1.58)	No serious imprecision	Moderate[a]
White blood cell count for predicting surgical complications (open cohort)	1	Adjusted OR [95% CI]: 0.97 (0.86 to 1.08)	Serious imprecision	Low[ab]
White blood cell count for predicting major adverse effects (endovascular cohort)	1	Adjusted OR [95% CI]: 1.67 (1.23 to 2.27)	No serious imprecision	Moderate[a]
White blood cell count for predicting major adverse effects (open cohort)	1	Adjusted OR [95% CI]: 1.07 (0.98 to 1.17)	Serious imprecision	Low[ab]
White blood cell count for predicting death (endovascular cohort)	1	Adjusted OR [95% CI]: 1.82 (1.12 to 2.96)	No serious imprecision	Moderate[a]
White blood cell count for predicting death (open cohort)	1	Adjusted OR [95% CI]: 1.17 (1.05 to 1.30)	Serious imprecision	Moderate[ab]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

Table 85: Clinical evidence summary: Cancer surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
White blood cell count for predicting cancer-specific survival (WBC <9.5 versus >9.5)	1	Adjusted OR [95% CI]: 1.91 (1.10 to 3.32)	No serious imprecision	Very Low[ab]
White blood cell count for cancer-specific survival (WBC <10.0 versus >10.0)	1	Adjusted OR [95% CI]: 1.56 (0.86 to 2.83)	Serious imprecision	Very Low[abc]
White blood cell count for predicting cancer-specific survival (WBC <11.0 versus >11.0)	1	Adjusted OR [95% CI]: 1.97 (1.00 to 3.88)	No serious imprecision	Very Low[ab]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Downgraded by 1 increment as patients were at high risk of mortality*
[c] *Imprecision was considered serious if the confidence intervals crossed the null line*

Platelet count

Table 86: Clinical evidence summary: Cancer surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Platelet count for predicting overall survival (platelet count <178 versus >178)	1	Adjusted OR [95% CI]: 1.54 (1.04 to 2.29)	No serious imprecision	Low[ab]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Downgraded by 1 increment as patients were at high risk of mortality*

Table 87: Clinical evidence summary: Non-cardiac surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Platelet count for predicating blood transfusion (moderate to severe thrombocytopenia)	1	Adjusted OR [95% CI]: 1.76 (1.49 to 2.08)	No serious imprecision	Low[a]
Platelet count for predicting blood transfusion (mild thrombocytopenia)	1	Adjusted OR [95% CI]: 1.28 (1.18 to 1.39)	No serious imprecision	Moderate[a]
Platelet count for predicting blood transfusion (low to normal thrombocytopenia)	1	Adjusted OR [95% CI]: 1.01 (0.96 to 1.06)	Serious imprecision	Very Low[ab]
Platelet count for predicting blood transfusion (thrombocytosis)	1	Adjusted OR [95% CI]: 1.44 (1.30 to 1.60)	No serious imprecision	Low[a]
Platelet count for predicting mortality (moderate to severe thrombocytopenia)	1	Adjusted OR [95% CI]: 1.93 (1.43 to 2.60)	No serious imprecision	Very serious[ab]
Platelet count for predicting mortality (mild thrombocytopenia)	1	Adjusted OR [95% CI]: 1.31 (1.11 to 1.55)	No serious imprecision	Moderate[a]
Platelet count for predicting mortality (low normal thrombocytopenia)	1	Adjusted OR [95% CI]: 0.91 (0.80 to 1.04)	Serious imprecision	Very Low[ab]
Platelet count for predicting mortality (thrombocytosis)	1	Adjusted OR [95% CI]: 0.91 (0.72 to 1.23)	Serious imprecision	Very Low[ab]
Platelet count for predicting mortality or major complication (moderate to severe thrombocytopenia)	1	Adjusted OR [95% CI]: 1.52 (1.32 to 1.75)	No serious imprecision	Moderate[a]
Platelet count for predicting mortality or major complication (mild thrombocytopenia)	1	Adjusted OR [95% CI]: 1.12 (1.04 to 1.21)	No serious imprecision	Moderate[a]

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Platelet count for predicting mortality or major complication (low to normal thrombocytopenia)	1	Adjusted OR [95% CI]: 1.00 (0.96 to 1.04)	Serious imprecision	Very Low[ab]
Platelet count for predicting mortality or major complication (thrombocytosis)	1	Adjusted OR [95% CI]: 1.36 (1.25 to 1.48)	No serious imprecision	Moderate[a]
Platelet count for predicting cardiac complication (moderate to severe thrombocytopenia)	1	Adjusted OR [95% CI]: 1.02 (0.67 to 1.55)	Serious imprecision	Very Low[ab]
Platelet count for predicting cardiac complication (mild thrombocytopenia)	1	Adjusted OR [95% CI]: 0.99 (0.81 to 1.21)	Serious imprecision	Very Low[ab]
Platelet count for predicting cardiac complication (low to normal thrombocytopenia)	1	Adjusted OR [95% CI]: 1.08 (0.95 to 1.23)	Serious imprecision	Very Low[ab]
Platelet count for predicting cardiac complication (thrombocytosis)	1	Adjusted OR [95% CI]: 1.13 (0.84 to 1.52)	Serious	Very Low[ab]
Platelet count for predicting pulmonary complication (moderate to severe thrombocytopenia)	1	Adjusted OR [95% CI]: 1.87 (1.50 to 2.33)	No serious imprecision	Moderate[a]
Platelet count for predicting pulmonary complication (mild thrombocytopenia)	1	Adjusted OR [95% CI]: 1.08 (0.95 to 1.23)	Serious imprecision	Very Low[ab]
Platelet count for predicting pulmonary complication (low to normal thrombocytopenia)	1	Adjusted OR [95% CI]: 1.06 (0.99 to 1.14)	Serious imprecision	Very Low[ab]
Platelet count for predicting pulmonary complication (thrombocytosis)	1	Adjusted OR [95% CI]: 1.30 (1.12 to 1.51)	No serious imprecision	Low[a]
Platelet count for predicting renal complication (moderate to severe thrombocytopenia)	1	Adjusted OR [95% CI]: 2.05 (1.48 to 2.84)	No serious imprecision	Moderate[a]

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
Platelet count for predicting renal complication (mild thrombocytopenia)	1	Adjusted OR [95% CI]: 1.45 (1.20 to 1.75)	No serious imprecision	Moderate[a]
Platelet count for predicting renal complication (low to normal thrombocytopenia)	1	Adjusted OR [95% CI]: 1.06 (0.92 to 1.22)	Serious imprecision	Very Low[ab]
Platelet count for predicting renal complication (thrombocytosis)	1	Adjusted OR [95% CI]: 1.48 (1.14 to 1.92)	No serious imprecision	Low[a]
Platelet count for predicting CNS complication (moderate to severe thrombocytopenia)	1	Adjusted OR [95% CI]: 0.73 (0.34 to 1.57)	Serious imprecision	Very Low[ab]
Platelet count for predicting CNS complication (mild thrombocytopenia)	1	Adjusted OR [95% CI]: 1.13 (0.85 to 1.50)	Serious imprecisions	Very Low[ab]
Platelet count for predicting CNS complication (low to normal thrombocytopenia)	1	Adjusted OR [95% CI]: 1.01 (0.83 to 1.23)	Serious imprecision	Very Low[ab]
Platelet count for predicting CNS complication (thrombocytosis)	1	Adjusted OR [95% CI]: 1.09 (0.69 to 1.72)	Serious imprecision	Very Low[ab]
Platelet count for predicting sepsis complication (moderate to severe thrombocytopenia)	1	Adjusted OR [95% CI]: 1.17 (0.92 to 1.49)	Serious imprecision	Very Low[ab]
Platelet count for predicting sepsis complication (mild thrombocytopenia)	1	Adjusted OR [95% CI]: 1.00 (0.89 to 1.12)	Serious imprecisions	Very Low[ab]
Platelet count for predicting sepsis complication (low to normal thrombocytopenia)	1	Adjusted OR [95% CI]: 0.95 (0.88 to 1.03)	Serious imprecision	Very Low[ab]
Platelet count for predicting sepsis complication (thrombocytosis)	1	Adjusted OR [95% CI]: 1.27 (1.12 to 1.44)	No serious imprecision	Low[a]
Platelet count for predicting wound complication (moderate to severe	1	Adjusted OR [95% CI]: 1.24 (0.97 to 1.59)	Serious imprecision	Very Low[ab]

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
thrombocytopenia)				
Platelet count for predicting wound complication (mild thrombocytopenia)	1	Adjusted OR [95% CI]: 1.11 (0.98 to 1.26)	Serious imprecisions	Very Low[ab]
Platelet count for predicting wound complication (low to normal thrombocytopenia)	1	Adjusted OR [95% CI]: 0.94 (0.88 to 1.00)	No serious imprecision	Low[a]
Platelet count for predicting wound complication (thrombocytosis)	1	Adjusted OR [95% CI]: 1.49 (1.31 to 1.69)	No serious imprecision	Low[a]
Platelet count for predicting thromboembolic complication (moderate to severe thrombocytopenia)	1	Adjusted OR [95% CI]: 1.08 (0.74 to 1.58)	Serious imprecision	Very Low[ab]
Platelet count for predicting thromboembolic complication (mild thrombocytopenia)	1	Adjusted OR [95% CI]: 1.09 (0.90 to 1.32)	Serious imprecisions	Very Low[ab]
Platelet count for predicting thromboembolic complication (low to normal thrombocytopenia)	1	Adjusted OR [95% CI]: 1.04 (0.93 to 1.16)	Serious imprecision	Very Low[ab]
Platelet count for predicting thromboembolic complication (thrombocytosis)	1	Adjusted OR [95% CI]: 1.74 (1.43 to 2.12)	No serious imprecision	Low[a]
Platelet count for predicting graft failure (moderate to severe thrombocytopenia)	1	Adjusted OR [95% CI]: 1.09 (0.55 to 2.16)	Serious imprecision	Very Low[ab]
Platelet count for predicting graft failure (mild thrombocytopenia)	1	Adjusted OR [95% CI]: 0.81 (0.56 to 1.17)	Serious imprecisions	Very Low[ab]
Platelet count for predicting graft failure (low to normal thrombocytopenia)	1	Adjusted OR [95% CI]: 0.87 (0.70 to 1.08)	Serious imprecision	Very Low[ab]
Platelet count for predicting graft	1	Adjusted OR [95% CI]: 1.31 (0.91 to 1.89)	Serious imprecision	Very Low[ab]

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
failure (thrombocytosis)				

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

11.7 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

Unit costs

Unit costs were provided for consideration alongside the Delphi survey results. Please see Appendix M for details. These are reported in Table 88 below.

Table 88: Unit cost of full blood count test

Equipment/staff	Quantity (units/minutes)	Cost	Source
Full blood count[(a)]	1	£6.00	HTA 2012[24]
	Total per patient	£6.00	

11.8 Evidence statements

11.8.1 Clinical

For forest plots, see Section J.5.2 in Appendix J.

11.8.1.1 Intervention review

No relevant studies were identified.

11.8.1.2 Prognostic review

All elective surgeries

One retrospective cohort study of 7,679 patients compared those with and without preoperative anaemia. They reported an increased risk of mortality at 90 days for those with anaemia, and this association held when those with severe anaemia or those who received red blood cell transfusions were excluded from the analysis [Moderate quality].

Orthopaedic surgery

Each of the studies in this category compared those with and without preoperative anaemia.

One prospective cohort study of 4,940 patients reported an increased risk of peri- or postoperative red blood cell transfusion, length of stay over 5 days and readmission at 90 days in those with preoperative anaemia [Moderate quality].

One retrospective cohort study of 15,222 patients reported an increased risk of periprosthetic joint infections in those with anaemia [Low quality]. This study also reported no clear difference in 30- or 90-day mortality between those with and without preoperative anaemia [Very low quality], although an increased risk of mortality at 1 year was seen for those with anaemia [Low quality].

One prospective study of 191 patients reported no clear increased risk of postoperative hyperglycaemia among those with anaemia [Low quality].

One retrospective cohort study of 605,655 patients reported an increased risk of peri- or postoperative allogenic blood transfusion among those with preoperative anaemia [Low quality].

Vascular surgery

One retrospective cohort study of 1,211 patients compared those with and without preoperative anaemia. They reported an increased risk of major adverse cardiac events in those with anaemia, and this risk increased with severity of anaemia [Moderate to Low quality].

One retrospective cohort study of 1,773 patients compared levels of preoperative white blood cell (WBC) count within the normal range. They reported an increased risk of postoperative complications, major adverse events and death for those with higher WBC count undergoing endovascular surgery [Moderate quality]. However, no clear difference in risk of these outcomes with variation in WBC count within the normal range was observed for those undergoing open surgery [Moderate to Low quality]. Note that the overall odds ratio for death in the open cohort masked an effect that both low and high values of preoperative WBC count in the open cohort were predictive of an increased risk of death.

Cancer surgery

One retrospective cohort study of 327 patients investigated the preoperative WBC count at three different thresholds. They reported that WBC count was predictive of survival when using the threshold of 9.5 per microlitre [Very low quality].

One retrospective study of 223 patients investigated the preoperative platelet count at a threshold of 178 x 10^9 l, which was reported to be predictive of overall survival [Low quality].

Non-cardiac surgery

One retrospective cohort study of 316,644 patients investigated the effect of increased preoperative platelet count on the incidence of blood transfusion, death and major complications. They stratified results according to preoperative platelet count and compared each of the following with normal platelet counts: moderate-to-severe thrombocytopenia, mild thrombocytopenia, low-to-normal platelet count and thrombocytosis.

They reported the following findings:

- Mild thrombocytopenia, moderate-to-severe thrombocytopenia, and thrombocytosis were each associated with increased risk of blood transfusion [Low quality].
- Mild and moderate-to-severe thrombocytopenia were also associated with increased risk of 30-day mortality [Low quality].
- Moderate-to-severe thrombocytopenia was associated with increased risk of postoperative pulmonary and renal complications [Low quality].
- Mild thrombocytopenia was associated with increased risk of renal complications [Low quality].
- Thrombocytosis was associated with increased risk of with pulmonary, renal, sepsis, wound and thromboembolic complications [Low quality].
- There was no clear association between platelet count and cardiac complications, central nervous system complications or graft failure [Very low quality].

11.8.2 Economic

- No relevant economic evaluations were identified.

11.9 Delphi survey results

Preoperative full blood count tests were included in the modified Delphi consensus survey.

The survey participants were asked if full blood count tests should be used as a routine preoperative test for patients undergoing elective surgery. Participants rated their response from strongly disagree to strongly agree using a nine-point Likert scale. Each question was considered to have reached consensus if greater than 70% of responses were in a single category (0–3 strongly disagree, 4–6 unclear, 7–9 strongly agree). Please see Appendix L for full details of the survey method and results.

11.9.1 Delphi statements where consensus was reached

Table 89: Patients: ASA1

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	88.55 strongly disagree (round 1)
Major or complex surgery	79.16 strongly agree (round 1)

Table 90: Patients: ASA2 with cardiovascular, diabetes, respiratory, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Major or complex surgery	83.33 strongly agree (round 1)

Table 91: Patients: ASA3 or ASA4 with cardiovascular, diabetes, respiratory, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	74.28 strongly agree (round 3)
Intermediate surgery	75.59 strongly agree (round 1)
Major or complex surgery	87.50 strongly agree (round 1)

11.9.2 Delphi statements where consensus was not reached

Table 92: Patients: ASA1

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Intermediate surgery	1	46.75	17.16
	2	59.74	13.25
	3	53.43	17.81

Table 93: Patients: ASA2 with cardiovascular, diabetes, respiratory, renal or obesity comorbidities

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Minor surgery	1	44.38	17.16
Intermediate surgery		21.76	48.82
Minor surgery	2	35.81	30.86
Intermediate surgery		10.84	63.86

Surgery grade		Results %	
Minor surgery	3	36.11	31.94
Intermediate surgery		8.22	57.54

11.10 Recommendations and link to evidence

Recommendations

10.Full blood count test

	Surgery grade		
ASA grade	**Minor**	**Intermediate**	**Major or complex**
ASA 1	**Do not routinely offer**	**Do not routinely offer**	**Offer**
ASA2	**Do not routinely offer**	**Do not routinely offer**	**Offer**
ASA3 or ASA4	**Do not routinely offer**	**Consider for people with cardiovascular or renal disease if any symptoms not recently investigated**	**Offer**

Relative values of different outcomes	The GDG considered all-cause mortality to be a critical outcome for the intervention and prognostic reviews. Change in health care management, complications relating to surgery or anaesthesia, length of stay after an operation, hospital readmission, adverse events caused by testing, health-related quality of life and ICU admission were considered to be important outcomes for both the intervention and prognostic reviews.
Trade-off between clinical benefits and harms	The majority of the evidence used full blood count testing to identify people with anaemia. There was only one study relating to platelet count. Furthermore, the evidence was restricted to major surgery and higher ASA grades. In the majority of studies, people without anaemia had better outcomes with regards to mortality, infections, length of stay and readmission rates. One study reported that people with anaemia had received more blood transfusions, which seems to indicate a change in management in relation to testing.
Economic evidence	No economic evaluations were identified for this review question.
Quality of evidence	The evidence was mainly of low or very low quality. The ideal evidence would have been testing as an intervention rather than a prognostic factor. However, no such evidence was identified. The prognostic evidence is problematic since it is often unclear whether test results led to management changes prior to surgery or whether the physician was aware of the full blood count results prior to surgery. Even though we restricted evidence to studies using multivariable analyses to

	identify test results as independent factors leading to postsurgical outcomes, some studies only adjusted for a minimum number of characteristics, and in other studies it was unclear which factors were accounted for. For instance, red blood cell transfusions were either not reported as an outcome or not adjusted for in the analyses, which makes results difficult to interpret. No evidence was identified for people with lower ASA grades for minor elective surgery. Therefore the evidence is not generalisable to all people covered in the remit of the guideline. Therefore, the GDG decided to put this into the Delphi Survey.
Other considerations	The GDG considered the impact of performing preoperative full blood count testing on rates of blood transfusion in the perioperative period. The normal haemoglobin (Hb) level is 115–180 g/litre, but blood transfusions are not normally given until the Hb is less than 70–100 g/litre (dependent on age and other comorbidities). As such, patients with higher baseline Hb levels will tolerate greater surgical blood loss before requiring a blood transfusion. Some GDG members felt that knowing the baseline Hb was important for informing transfusion decisions and enabled reduction in transfusion rates, thus avoiding unnecessarily exposing patients to the serious hazards of blood transfusion. However, other GDG members disagreed and felt that it was the haemodynamic status of the patient that was of greatest importance in blood transfusion decisions and not preoperative Hb levels. The GDG also noted that the need for blood transfusion in patients having minor and intermediate surgery was extremely low. The GDG considered the impact of preoperative anaemia on postoperative outcomes. There was agreement that preoperative anaemia is associated with worse postoperative outcomes. The GDG discussed whether investigating for preoperative anaemia using the full blood count and then instituting anaemia management prior to surgery could improve postoperative outcomes. The evidence considered in this review did not answer this question. However, national guidelines exist that recommend this practice, including the NATA (National Association of Testing Authorities, Australia) guidelines[40], which recommend testing for and treating anaemia in patients undergoing elective orthopaedic surgery. The use of intravenous iron to treat iron-deficiency anaemia was also discussed as it is increasingly being used in the preoperative setting. The GDG were aware of a systematic review and meta-analysis published by Litton et al in 2013[64] that found that this is generally safe, effective and reduces the need for blood transfusion. The GDG discussed that full blood count testing would be required to direct and monitor this therapy. There were no known equality issues that related to this testing. The GDG noted that there are genetic causes of anaemia, but these were considered most likely to have already been detected during childhood. The higher prevalence of anaemia in the older population was also discussed, but it was felt that this would be adequately considered as this population would be more likely to have a higher ASA grade and therefore be offered full blood count testing. The GDG noted an RCT[21] (Chung 2009) which was a battery of tests (full blood count, electrolytes, blood glucose, creatinine, ECG, and chest X-ray). As this study presented their outcomes based on a battery of tests versus no testing, it did not match our protocol. However, the study did not conflict with the included evidence. The study concluded that there was no significant difference in perioperative adverse events between the testing and no testing groups, and that there was no association between abnormal test results and perioperative adverse events. The GDG also discussed the benefit of doing a battery of tests. Venepuncture is used

	to obtain samples for blood tests, including the full blood count. Venepuncture is safe and generally well-tolerated by patients with only minor side effects, such as bruising, fainting and infection. The majority of patients having major or complex surgery will undergo venepuncture to have a blood group test and antibody screen prior to surgery. The GDG discussed whether patients could also have a full blood count performed at this time, as it would pose no additional risk to the patient and is an inexpensive test. The consensus view was that patients undergoing minor surgery did not routinely need a full blood count. In contrast, the GDG felt that all patients ASA3 or above having major or complex surgery should have a baseline full blood count, irrespective of comorbidities. Opinion regarding full blood count testing for patients undergoing intermediate surgery and patients ASA1 or ASA2 undergoing major or complex surgery was split, so the GDG felt that the full blood count test should be entered into the Delphi survey.
Delphi	Both the GDG and Delphi survey results were in agreement that full blood count should be offered to all patients undergoing major surgery. These surgeries were considered to be a higher risk and more likely to offer benefits to patients. Moreover, this was supported by the clinical evidence previously considered. The GDG did not agree with the Delphi survey results with regards to minor surgery in those with comorbidities, and recommended that full blood count should not be offered as part of routine preoperative testing, primarily as it was unlikely to change patient management in this population. It was noted that consensus was not reached for ASA2 patients. The GDG felt that the Delphi results may reflect current practice, but that additional testing in these patients could cause stress, inconvenience and increase risk of spurious findings. The GDG decided not to recommend this test for any minor procedure. The GDG debated the use of full blood count before intermediate surgery and suggested that the haemoglobin and possibly platelets were the only parameters which would inform perioperative management. However it was felt that this only applied to the ASA3 grade, and in particular to cardiovascular and renal comorbidities. The GDG pointed out that the majority of intermediate surgeries were day cases, and at lower risk of postoperative complications, and noted the lack of evidence to suggest improved outcomes following full blood count measurement. The GDG decided not to recommend full blood count testing for ASA1 and ASA2 patients, and to recommend consideration of the tests in ASA3 and ASA4 patients, taking into account underlying conditions.
Economic considerations	In the absence of evidence, unit costs were provided alongside the results of the Delphi survey. The cost of performing a full blood count was found to be £6, including staff time and equipment. Although the cost of performing the test is low, the GDG felt that the results of the test would not be likely to change perioperative management in individuals undergoing minor surgery and most individuals undergoing intermediate surgery, regardless of whether there was an abnormal result. However the GDG felt that individuals with cardiovascular or renal disease are more likely to be anaemic and would benefit from preoperative full blood count testing. It is therefore unlikely to be a cost-effective use of NHS resources to carry out preoperative full blood count tests for minor and intermediate surgery except in the outlined subset of patients. In major surgery, due to the higher risk of blood loss, the GDG felt that the results from a full blood count test could have an impact of patient outcomes, therefore justifying its use from a cost-effectiveness perspective.

12 Kidney function tests

12.1 Introduction

Kidney function tests involve sampling venous blood to test for creatinine, electrolytes and sometimes urea to examine the functional status of the kidneys. Estimated glomerular filtration rate is also frequently reported. In the preoperative setting the test is used to establish a baseline for the patient, to inform prediction of postoperative risks and to plan medical management in the perioperative period. The test is considered safe and is straightforward to perform and analyse, but may be painful for the patient. There is a low risk of complications including haematoma formation, vasovagal reactions and infection. However, detection of abnormalities is uncommon in asymptomatic, fit and healthy individuals and there is uncertainty about the clinical effectiveness of performing routine preoperative kidney function tests in all individuals.

See section 4.4 for a summary of the methodological approach taken for this preoperative test.

12.2 Health technology assessment 2012

Kidney function tests are part of the HTA 2012 update – see section 10.2 for details.

12.3 Review question (intervention): What is the usefulness of kidney function tests (urea, estimated glomerular filtration rate and electrolyte tests) in predicting outcome or altering perioperative management for adults and young people undergoing any type of elective non-cardiac surgery?

Table 94: Characteristics of review question – comparative protocol

Population	Adults and young people classified as patients ASA grade 1 to 4 undergoing: • Minor, intermediate, or major or complex surgery Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory, renal, obesity, diabetes
Intervention	Kidney function tests (urea, estimated glomerular filtration rate and electrolyte tests)
Comparison(s)	No kidney function tests
Outcomes	**Critical:** • All-cause mortality • Complications relating to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney injury, infection) **Important:** • Change in healthcare management (for example cancellation of surgery) • Length of hospital stay after an operation • Hospital readmission • Adverse events caused by testing • Health-related quality of life

	• Intensive care unit (ICU) admission

See section 10.2 for further details.

12.4 Clinical evidence

No relevant clinical studies comparing preoperative resting kidney function testing with no preoperative kidney function testing were identified.

12.5 Review question (prognostic): Do kidney function tests (urea, estimated glomerular filtration rate and electrolyte tests) predict prognosis (patient outcomes after surgery) in adults and young people ASA 1–4 undergoing any type of elective non-cardiac surgery?

Table 95: Characteristics of review question – prognostic question

Population	Adults and young people classified as patients ASA grade 1 to 4 undergoing: • Minor, intermediate, or major or complex surgery Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory, renal, obesity, diabetes
Prognostic test	Kidney function tests (urea, estimated glomerular filtration rate and electrolyte tests)
Key confounding factors	• Age • Comorbidities
Outcomes	**Critical:** • All-cause mortality • Complications relating to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney injury, infection) • Change in healthcare management (for example cancellation of surgery) • Length of hospital stay • Hospital readmission • Adverse events caused by testing (time of testing) • Health-related quality of life • Intensive care unit (ICU) admission

12.6 Clinical evidence

Three studies[5,57,66,87] were included in the prognostic review; these are summarised in Table 96 below. Studies are further subdivided by surgery type and by relevant patient comorbidity:

- Vascular surgery
  - Carotid endarterectomy
- Endovascular abdominal artery repair
- All non-cardiac surgery

See also the clinical study selection flow chart in Appendix E, clinical evidence tables in Appendix H, forest plots in Appendix J, GRADE tables in Appendix I and excluded clinical studies list in Appendix K.

Table 96: Summary of studies included in the review

Study	Population	Analysis	Prognostic test variable(s)	Confounders (list)	Outcomes	Limitations
Kidney function tests						
AbuRahma 2013 [5]	Retrospective cohort n=940 procedures (881 patients) Carotid endarterectomy	Multivariable logistic regression but unclear variables	eGFR	• Race • Age • Sex • Serum albumin Serum urea nitrogen	30-day stroke and/or death	ASA and surgery grade not stated No GFR data available in 15/940 operations
Mases 2014[66]	Post-hoc analysis of prospectively collected data n=2323 Non-cardiac surgery	Logistic regression	eGFR	• Race • Age • Sex • Serum albumin • Serum urea nitrogen	All-cause mortality MAACE	Data missing from 34% of original sample Restricted to those aged ≥40 years Unclear analysis
Soong 2008 [87]	Retrospective cohort n=155 Endovascular abdominal aortic aneurysm repair	Multiple regression analysis stated but unclear reporting eGFR equation adjusts for some factors in the formula	eGFR	• Race • Age • Sex • Serum albumin • Serum urea nitrogen	Perioperative and long-term mortality Postoperative deterioration in renal function Development of renal failure	Poor reporting Small sample size ASA and surgery grade not stated

Table 97: Clinical evidence summary: Vascular surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
eGFR for predicting postoperative mortality or stroke (platelet count <178 versus >178)	1	Adjusted OR [95% CI]: 3.70 (1.30 to 10.53)	No serious imprecision	Low[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*

Table 98: Clinical evidence summary: Endovascular repair of abdominal aortic aneurysm

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
eGFR for predicting postoperative mortality	1	Adjusted OR [95% CI]: 0.25 (0.03 to 2.32)	Serious imprecision	Very Low[ab]
eGFR for predicting postoperative renal failure	1	Adjusted OR [95% CI]: 0.07 (0.03 to 0.21)	No serious imprecision	Low[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

Table 99: Clinical evidence summary: Non-cardiac surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
eGFR for predicting peri- or postoperative mortality	1	Adjusted OR [95% CI]: Stage 2: 0.8 (0.3–1.8) Stage 3a: 2.2 (0.9–5.4) Stage 3b: 2.8 (0.9–8.5) Stage 4: 11.3 (4.3–29.9)	Serious imprecision	Very Low[ab]

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
		Stage 5: 5.8 (1.5–21.9)		
eGFR for predicting peri- or postoperative operative MACE	1	Adjusted OR [95% CI]: Stage 2: 1.5 (0.9–2.5) Stage 3a: 1.8 (0.9–3.5) Stage 3b: 3.9 (0.9–8.0) Stage 4: 4.8 (1.9–11.8) Stage 5: 3.9 (1.3–012.0)	Serious imprecision	Very Low[ab]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

Table 100: Clinical evidence summary: Radical nephrouretectomy

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
eGFR for predicting early postoperative eGFR	1	Adjusted beta coefficient [95% CI]: 5.04 (2.76–7.31)	No serious imprecision	Low[ab]
eGFR for predicting late postoperative eGFR	1	Adjusted beta coefficient [95% CI]: 3.33 (1.35–5.30)	No serious imprecision	Low[ab]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

12.7 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

Unit costs

Unit costs were provided for consideration alongside the Delphi survey results. Please see Appendix M for details. These are reported in Table 101 below.

Table 101: Kidney function test

Equipment/staff	Quantity (units/minutes)	Cost	Source
Testing for creatinine, urea and electrolytes[(a)]	1	£6.00	CG 182 (Chronic Kidney disease 2014)[78]
	Total per patient	£6.00	

12.8 Evidence statements

12.8.1 Clinical

For forest plots, see Section J.5.3 in Appendix J.

12.8.1.1 Intervention review

No relevant studies were identified.

12.8.1.2 Prognostic review

Vascular surgery

One retrospective cohort study of 881 patients compared those with an eGFR value of <60 ml/minute/1.73m^2 with those with higher values. They reported an increased risk for the composite outcome of postoperative mortality or stroke among those with eGFR <60 ml/minute/1.73m^2 [Low quality].

Endovascular repair of abdominal aortic aneurysm

One retrospective cohort study of 155 patients also compared those with an eGFR value of <60 ml/minute/1.73m^2 with those with higher values. They reported an increased risk of perioperative mortality and postoperative renal failure among those with eGFR <60 ml/minute/1.73m^2, but there is considerable uncertainty for the mortality outcome [Very low quality].

All non-cardiac surgery

One post-hoc analysis of a prospective study of 2323 patients compared those with an eGFR value of >90 ml/minute/1.73m^2 with those with lower values. They reported a general increase in risk of all-

cause mortality and major adverse cardiovascular and cerebrovascular events (MACE) with declining eGFR [Very low quality].

12.8.2 Economic

- No relevant economic evaluations were identified.

12.9 Delphi survey results

Preoperative kidney function tests were included in the modified Delphi consensus survey.

The survey participants were asked if kidney function tests should be used as a routine preoperative test for patients undergoing elective surgery. Participants rated their response from strongly disagree to strongly agree using a nine-point Likert scale. Each question was considered to have reached consensus if greater than 70% of responses were in a single category (0–3 strongly disagree, 4–6 unclear, 7–9 strongly agree). Please see Appendix L for full details on the survey method and results.

12.9.1 Delphi statements where consensus was achieved

Table 102: Patients: ASA1

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	90.26 strongly disagree (round 1)

Table 103: Patients: ASA2 with cardiovascular, respiratory, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Major or complex surgery	85.91 strongly agree (round 3)

Table 104: Patients: ASA3 or ASA4 with cardiovascular, respiratory, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Intermediate surgery	77.46 strongly agree (round 3)
Major or complex surgery	95.77 strongly agree (round 3)

12.9.2 Delphi statements where consensus was not reached

Table 105: Patients: ASA1

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Intermediate surgery	1	62.99	14.28
Major surgery		26.31	60.53
Intermediate surgery	2	61.25	15.0
Major surgery		21.25	66.25
Intermediate surgery	3	59.15	14.09
Major surgery		12.68	59.16

Table 106: Patients: ASA2 with cardiovascular, respiratory, renal or obesity comorbidities

Surgery grade		Results %

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Minor surgery	1	24.84	35.29
Intermediate surgery		20.53	54.30
Minor surgery	2	34.61	39.75
Intermediate surgery		8.75	63.75
Minor surgery	3	28.17	38.03
Intermediate surgery		12.67	59.16

Table 107: Patients: ASA3 or ASA4 with cardiovascular, respiratory, renal or obesity comorbidities

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Minor	1	20.00	58.00
	2	12.66	68.35
	3	4.29	67.14

12.10 Recommendations and link to evidence

Recommendations

11.Kidney function tests

	Surgery grade		
ASA grade	**Minor**	**Intermediate**	**Major or complex**
ASA 1	**Do not routinely offer**	**Do not routinely offer**	**Consider in people at risk of AKI[e]**
ASA2	**Do not routinely offer**	**Consider in people at risk of AKI[f]**	**Offer**
ASA3 or ASA4	**Consider in people at risk of AKI[g]**	**Offer**	**Offer**

Relative values of different outcomes	The GDG considered all-cause mortality to be a critical outcome for the intervention and prognostic reviews. Change in health care management, complications relating to surgery or anaesthesia, length of stay after an operation, hospital readmission, adverse events caused by testing, health-related quality of life and ICU admission were considered to be important outcomes for both the intervention and prognostic reviews.
Trade-off between clinical benefits and harms	Four studies using blood kidney function testing prior to major surgery were identified. All of the studies associated an increased glomerular filtration rate (eGFR) >60 ml/minute/1.73 m^2 with lower rates of post- or perioperative mortality or post-surgical renal failure. The GDG noted that if acute kidney injury was suspected, performing a kidney function test would assist in planning the patient's management post-surgery and should be considered in this population
Economic evidence	No economic evaluations were identified for this review question.
Quality of evidence	The evidence was mainly of low or very low quality. The ideal evidence would have been testing as an intervention rather than a prognostic factor. However, no such evidence was identified. The prognostic evidence is problematic since it is often unclear whether test results led to management changes prior to surgery or whether the physician was aware of the test results prior to surgery. We searched for studies using multivariable analyses to identify test results as

e See recommendation 1.1.8 in NICE CG169: Acute kidney injury: prevention, detection and management (2013)
f See recommendation 1.1.8 in NICE CG169: Acute kidney injury: prevention, detection and management (2013)
g See recommendation 1.1.8 in NICE CG169: Acute kidney injury: prevention, detection and management (2013)

	independent factors leading to postsurgical outcomes. In one study it was unclear which variables were used in the multivariable analysis. A further study did not adjust for variables in the analysis, but as the eGFR measure accounts for other factors (such as age and race) the study was included. No evidence was identified for people with lower ASA grades for minor elective surgery. The evidence is therefore not generalisable to all people covered in the remit of the guideline. Due to the limited quality and applicability of the identified evidence, the GDG agreed that kidney function tests should be entered into the Delphi consensus survey.
Other considerations	The GDG discussed the NICE guideline on acute kidney injury (AKI),[73] which states that adults undergoing surgery are considered an at-risk group for AKI. The GDG discussed whether this therefore mandated surgical patients having baseline urea, electrolyte and creatinine tests taken preoperatively. However, most of the AKI risk factors listed in the AKI guideline would be reflected in an ASA grade of 3 or above, and these patients would be offered the test under this guideline The recommendations from the 2003 preoperative tests guideline[71] were reviewed, including the recommendation to offer urea, electrolyte and creatinine testing to everyone with known chronic kidney disease (CKD), regardless of the grade of surgery. The GDG felt that minor surgery is unlikely to lead to significant physiological changes that might result in kidney injury in ASA grade 1 and 2 patients, so having a baseline preoperative test would be unnecessary. However, the GDG felt that kidney function tests should be considered in ASA grade 3 and 4 patients at risk of AKI due to the potential impact these results may have on postoperative drug therapy. The GDG discussed the ideal time period for kidney function tests to be undertaken prior to surgery. They felt that this would depend on the nature of the surgery and pre-existing conditions, so considered it difficult to formulate overall guidance on this. However, it was considered good practice for a GP to include recent blood results in the initial referral to the surgeon, and the GDG commented that patients should be encouraged to bring all healthcare documents to the preoperative clinic. The GDG agreed that a general recommendation should be included at the beginning of the guideline (see section 3.4.1.2) stating that any results of tests undertaken in primary care should be provided when referring the patient on for preoperative assessment. This would prevent unnecessary duplication of tests reducing both costs and delays for the patient. There were no known equality issues that related to this test. The higher prevalence of impaired renal function in the older population was discussed, but it was felt that this would be adequately considered as this population would be more likely to have a higher ASA grade and therefore be offered urea, electrolyte and creatinine testing.
Delphi	The GDG considered the results of the Delphi to be more conservative than their own individual practice. In particular, the GDG considered the minor surgery population and suggested a 'do not routinely offer' recommendation for ASA grade 1 and 2 patients. The GDG noted that minor procedures present a limited risk to these patients with the likelihood of postoperative complications being very low. For this reason minor surgeries are commonly undertaken as day cases so a baseline result is unnecessary as patients are rarely monitored post-surgery. Moreover, the GDG did not feel that the test would alter perioperative management in these patients. The GDG considered that baseline kidney function tests should be considered in ASA grade 3 and 4 patients at risk of AKI due to the potential for these results to alter the postoperative management of these patients, with specific regard to drug therapy. For intermediate surgery the GDG agreed with the Delphi survey to offer the test to

	all populations in the ASA3 and ASA4 categories. The GDG agreed that ASA1 patients should not be offered the test as it is unlikely to change perioperative management or outcome. The GDG felt that the ASA2 population was ill-defined and decided to err on the side of caution with a 'consider' recommendation. The GDG felt that the test should only be offered as a baseline reference to patients suspected to be at risk of AKI,[73] in order to inform management prior to surgery. The GDG considered the results of the Delphi survey and agreed to offer the test for major or complex surgery in patients ASA2 and above. They agreed that kidney function tests were useful in major surgery as a preoperative baseline value and can be used to measure postoperative recovery. The GDG considered a 'do not offer' recommendation for ASA1 patients, but felt that there is a high incidence of asymptomatic advanced chronic kidney disease within the general population. The GDG agreed with the Delphi consensus and made a 'consider' recommendation to be informed by current guidance for AKI populations.[73] The GDG debated the current AKI guideline[73] which recommends assessment of risk in all patients at high risk of AKI, including patients above 65 years of age (although 60 was suggested by some Delphi respondents). They agreed that risk is assessed in the preoperative tests guideline by surgery type and ASA grade, so felt an age distinction was not necessary. However, the GDG wanted to cross-refer to the AKI guideline[73] in order to highlight populations at increased risk of AKI.
Economic considerations	In the absence of evidence, unit costs were provided alongside the results of the Delphi survey. The cost of performing kidney functions tests was found to be £6.00. Although the cost of performing the test is low, no clinical evidence was identified that suggested whether doing the test lead to a change in postoperative outcomes. The only clinical evidence found was prognostic and although this suggested that the outcomes the test identified lead to poorer surgical outcomes, it could not be shown whether knowing these outcomes preoperatively would improve postsurgical outcomes. The GDG considered the cost of further investigations that may arise through an abnormal test result including the cost of a nephrology outpatient visit (£145) and an ultrasound scan (£48 –£59). The GDG recognised that the test had value as a baseline indicator for those suspected of acute kidney injury (AKI) which could inform postoperative management, thus improving health outcomes for individuals undergoing intermediate or complex surgery. The GDG did not feel it was necessary to test those who were ASA1 undergoing minor or intermediate surgery as the prevalence of AKI is low and the health outcomes post-surgery are unlikely to be significantly altered by knowing the test results beforehand. For minor surgery in ASA grade 2 patients, the GDG felt that the results of the test would not alter management and would therefore not improve health outcomes. Due to this the test is unlikely to be a cost-effective use of NHS resources in the outlined set of patients.

13 Haemostasis tests

13.1 Introduction

Haemostasis tests involve sampling venous blood to detect congenital and acquired coagulation disorders and to examine the effects of anticoagulant drugs. In the preoperative setting, the test is used to establish a baseline for the patient and may be used to plan the use of blood products and blood salvage techniques in the perioperative period. The test is considered safe and is straightforward to perform and analyse, but may be painful for the patient. There is a low risk of complications including haematoma formation, vasovagal reactions and infection. However, detection of abnormalities is uncommon in asymptomatic, fit and healthy individuals and there is uncertainty about the clinical effectiveness of performing routine preoperative haemostasis tests in all individuals.

See section 4.4 for a summary of the methodological approach taken for this preoperative test.

13.2 Delphi survey results

As no new evidence on the use of haemostasis tests as a routine preoperative test was identified during the scoping phase of this guideline it was decided not to carry out an evidence review, but to include haemostasis tests in the modified Delphi survey to re-evaluate the consensus held amongst health professionals on the value of routinely conducting the test prior to elective surgery.

The survey participants were asked if haemostasis tests should be used as a routine preoperative test for patients undergoing elective surgery. Participants rated their response from strongly disagree to strongly agree using a nine-point Likert scale. Each question was considered to have reached consensus if greater than 70% of responses were in a single category (0–3 strongly disagree, 4–6 unclear, 7–9 strongly agree). Please see Appendix L for full details on the survey method and results.

13.2.1 Delphi statements where consensus was reached

Table 108: Patients: ASA1

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	96.31% strongly disagree (round 1)
Intermediate surgery	88.95% strongly disagree (round 1)
Major or complex surgery	73.45% strongly disagree (round 1)

Table 109: Patients: ASA2 with cardiovascular, diabetes, respiratory, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	85.80 strongly disagree (round 1)
Intermediate surgery	72.04 strongly disagree (round 1)

Table 110: Patients: ASA3 or ASA4 with cardiovascular, diabetes, respiratory, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	78.88 strongly disagree (round 1)

13.2.2 Delphi statements where consensus was not reached

Table 111: Patients: ASA2 with cardiovascular, diabetes, respiratory, renal or obesity comorbidities

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Major or complex	1	55.63	18.75
	2	40.74	34.56
	3	46.27	32.84

Table 112: Patients: ASA3 or ASA4 with cardiovascular, diabetes, respiratory, renal or obesity comorbidities

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Intermediate surgery	1	59.88	19.14
Major or complex surgery		50.93	40.00
Intermediate surgery	2	51.22	25.62
Major or complex surgery		30.49	51.23
Intermediate surgery	3	59.15	16.9
Major or complex surgery		33.8	47.89

13.3 Economic evidence

Unit costs were provided for consideration alongside the Delphi survey results. Please see Appendix M for details. These are reported in Table 113 below.

Table 113: Unit cost of haemostasis test

Equipment/staff	Quantity (units/minutes)	Cost	Source
Testing prothrombin time (PT)	1	£26.00[(a)]	NICE DG13[74]
Testing activated partial thromboplastin time (APTT)			
Testing platelet count (PC)			
Testing plasma fibrinogen concentration (PFC)			
Testing activated clogging/coagulation time (ACT)			
Phlebotomy[(b)]	1	£3.42	NHS reference costs 2013-14[30]

Equipment/staff	Quantity (units/minutes)	Cost	Source
	Total per patient	£29.42	

(a) This cost does not include staff cost of taking blood
(b) Includes staff time and equipment required to take the blood

13.4 Recommendations and link to evidence

Recommendations	**12. Do not routinely offer haemostasis tests before surgery.** **13. Consider haemostasis tests in people with chronic liver disease having intermediate or major or complex surgery.** • **If people taking anticoagulants need modification of their treatment regimen, make an individualised plan in line with local guidance.** • **If clotting status needs to be tested before surgery (depending on local guidance) use point–of-care testing.**[h]
Delphi	The GDG considered the Delphi results and noted the consensus for not using the test preoperatively for healthy individuals or those undergoing minor surgery. The group agreed that the test should not be offered as a routine test for any type of surgery. The GDG suggested that assessment of clinical bleeding history and physical examination as part of standard preoperative assessment should identify patients at risk of bleeding during surgery. While comments from the Delphi survey did not explicitly highlight patients with chronic liver disease, the GDG suggested that patients suffering from chronic liver failure have an increased risk of bleeding and this may require monitoring prior to intermediate and major or complex surgery. The GDG also identified patients on anticoagulant therapy as a high risk population for bleeding. These patients may require anticoagulant modification prior to surgery. The GDG indicated that this is regularly conducted using point-of-care testing both pre-surgery and perioperatively, but noted that the effects of direct oral anticoagulants cannot currently be measured by routine testing.
Economic considerations	The overall cost of performing standard laboratory haemostasis tests was found to be £29.42, including staff time and equipment. The cost of a clinical haematology outpatient visit was noted to be £160, which may occur if further investigations were needed due to an abnormal test result. The GDG discussed that patients with known blood clotting disorders or on anticoagulants are more likely to have abnormalities and therefore potential complications such as platelet transfusions or increased blood transfusions. In this scenario it is usually the status of the individual's clotting at the time of surgery that is important, rather than a value from a few days, or even weeks, before. However, in individuals who are not on anticoagulants or who do not suffer from chronic liver disease, the prevalence of abnormalities identified by haemostasis testing that would alter management is low. It is therefore likely to be cost-effective to perform haemostasis tests on patients with increased risk of related complications, however it would not be cost-effective to perform this test routinely.

h Note that currently the effects of direct oral anticoagulants (DOACs) cannot be measured by routine testing.

14 Glycated haemoglobin (HbA1c) test

14.1 HbA1c testing in people with diabetes

14.1.1 Introduction

The glycated haemoglobin (HbA1c) test is a venous blood test used to diagnose diabetes mellitus and monitor glucose control in patients known to have diabetes. In the preoperative setting, the test is used in those with known diabetes and may also be used to screen for previously undiagnosed diabetes. The information from the test may be used to alter diabetes management both pre and perioperatively, with the aim of reducing postoperative morbidity and mortality. The test is considered safe and is straightforward to perform and analyse, but may be painful for the patient. There is a low risk of complications including haematoma formation, vasovagal reactions and infection. However there is uncertainty regarding the optimal timing of the test in individuals known to have diabetes and regarding the clinical effectiveness of preoperative screening in asymptomatic patients without diabetes.

See section 4.4 for a summary of the methodological approach taken for this preoperative test.

14.1.2 Review question (intervention): What is the clinical- and cost-effectiveness of using HbA1c (glycated haemoglobin) as a preoperative test in improving patient outcomes in adults and young people with diabetes and mild to severe comorbidities undergoing non-cardiac elective surgery?

Table 114: PICO characteristics of review question

Population	Adults and young people with diabetes (all types) undergoing non-cardiac related surgery Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory, renal, obesity
Intervention	HbA1c (glycated haemoglobin)
Comparison(s)	No HbA1c (glycated haemoglobin)/clinical assessment only
Outcomes	**Critical:** • All-cause mortality • Health-related quality of life **Important:** • Complications related to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney failure, infection) • Length of hospital stay • Hospital readmission • Adverse events caused by testing • Intensive care unit (ICU) admission

14.1.3 Clinical evidence

No relevant clinical studies comparing preoperative HbA1c testing in patients diagnosed with diabetes patients with no preoperative HbA1c testing in patients diagnosed with diabetes were identified.

14.1.4 Review question (prognostic): Does HbA1c (glycated haemoglobin) predict prognosis (patient outcomes after surgery) of adults and young people with diabetes (all types) and mild to severe comorbidities undergoing major or complex non-cardiac elective surgery?

Table 115: Characteristics of review question

Population	Adults and young people with diabetes (all types) undergoing non-cardiac related surgery Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory and renal diseases, obesity
Prognostic test	Level of glycated haemoglobin (HbA1c)
Key confounding factors	*Minimum set of confounders that should be adjusted for (will vary per outcome)* • Age • BMI • Comorbidities (cardiovascular, respiratory and renal diseases, obesity) • Patients taking drugs that cause a rapid rise in glucose (such as corticosteroids or antipsychotic drugs (≤2 months). HbA1c can be used in patients taking these drugs longer term (>2 months) who are not clinically unwell. • Ethnic groups • Patients with acute pancreatic damage or who have undergone pancreatic surgery • Patients with renal failure • Patients with HIV infection
Outcomes	**Critical:** • All-cause mortality **Important:** • Complications relating to surgery or anaesthesia • Length of hospital stay (post-operation) • Hospital readmission • Adverse events after surgery (wound infection) • Health-related quality of life • Intensive care unit (ICU) admission

14.1.5 Clinical evidence

Four retrospective cohort studies[6,20,35,47] were included in the prognostic review. All of these studies looked at the level of HbA1c as a predictor of outcome after surgery.

Studies addressing the following surgery types were reported:

- Arteriovenous fistula surgery[6]
- Non-cardiac surgery[35]
- Joint arthroplasty[2047]

Evidence from these studies is summarised below. See also the clinical study selection flow chart in Appendix E, forest plots in Appendix J, clinical evidence tables in Appendix H and excluded clinical studies list in Appendix K.

Table 116: Summary of studies included in the prognostic review

Study	Population	Analysis	Prognostic variable(s)	Confounders (list	Outcomes	Limitations
Afsar 2012[6]	Retrospective cohort Single centre n=73/233 patients with diabetes and non-dialysis stage 5 chronic kidney disease undergoing arteriovenous fistula surgery	Multivariate logistic regression analysis	HbA1c	• Age • Gender • Smoking status • Fistula location • BMI • Presence of coronary artery disease • Peripheral artery disease • Fasting glucose • HbA1c • Use of antiplatelet drugs	Primary arteriovenous fistula failure	Less than 10 events per variable included in the multivariate analysis High risk of attrition bias 15.2% Not stratified by ASA grade
Chrastil 2015[20]	Retrospective cohort Multi centre n=328/13272 patients with diabetes undergoing either primary total knee arthroplasty or primary total hip arthroplasty	Multivariable Cox proportional hazard model	HbA1c	• HbA1c • Preoperative glucose • Age • Gender • BMI • Charlson comorbidity index • Smoking status • Diabetic complications	Periprosthetic joint infection	Retrospective Unclear which variables adjusted for Not stratified by ASA grade
Dronge 2006 [35]	Retrospective cohort Single centre n=490/647 patients with diabetes undergoing major non-cardiac surgery	Logistic regression	HbA1c	Age, ASA grade, ADL assessment, case status, operation length, wound class, HbA1c	Postoperative infectious complications	Retrospective Multivariate analysis performed only on factors significant in

Study	Population	Analysis	Prognostic variable(s)	Confounders (list	Outcomes	Limitations
						univariate analysis
Harris 2013 [47]	Retrospective cohort Single centre n=6088 patients with diabetes undergoing joint arthroplasty	Boosted regression (non-parametric)	HbA1c	Thirty-eight variables including: age at time of surgery, gender, race, BMI, ASA physical score status, alcohol consumption, smoking status, comorbidities, VASQIP functional health status score, anaesthesia technique, total operation time, postgraduate year of surgeon and other preoperative lab values.	Any complications 90-day mortality	Retrospective Unclear which variables adjusted for

Table 117: Clinical evidence summary: Arteriovenous fistula surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
HbA1c <7% versus HbA1c >7% for predicting primary arteriovenous fistula failure (adjusted ORs) [adults with diabetes]	1	Adjusted OR [95% CI]: 2.78 [1.3, 5.32]	No serious imprecision	LOW[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*

Table 118: Clinical evidence summary: Joint arthroplasty

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
HbA1c <7% versus HbA1c >7% for predicting peri-prosthetic joint infection (adjusted HRs) [adults with diabetes]	1	Adjusted HR [95% CI]: 0.86 [0.68, 1.09]	Serious	LOW[ab]
HbA1c <7% versus HbA1c >7% for	1	Adjusted HR [95% CI]: 1.30 [1.08, 1.56]	Serious	MODERATE[ab]

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
predicting death (adjusted HRs) [adults with diabetes]				
HbA1c <7% versus HbA1c >7% for predicting 90-day mortality (Adjusted ORs) [Adults with diabetes]	1	Adjusted OR [95% CI]: 1.37 [0.82, 2.29]	Serious	VERY LOW[ab]
HbA1c <7% versus HbA1c >7% for predicting number of complications (adjusted ORs) [adults with diabetes]	1	Adjusted OR [95% CI]: 1.18 [0.97, 1.44]	Serious	VERY LOW[ab]
HbA1c <7% versus HbA1c >7% for predicting all complications (adjusted ORs) [adults with diabetes]	1	Adjusted OR [95% CI]: 1.22 [1.01, 1.47]	No serious imprecision	LOW[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

Table 119: Clinical evidence summary: Non-cardiac surgery

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
HbA1c <7% versus HbA1c >7% for predicting postoperative infectious complications (adjusted ORs) [adults with diabetes]	1	Adjusted OR [95% CI]: 2.13 [1.23, 3.69]	No serious imprecision	LOW[a]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*

14.1.6 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

Unit costs

Unit costs were provided for consideration alongside the Delphi survey results. Please see Appendix M for details. These are reported in Table 120 below.

Table 120: Unit cost of blood glucose (HbA1c) test

Equipment/staff	Quantity (units/minutes)	Cost	Source
Haematology[(a)]	1	£3.00	NHS reference costs 2013-14[30]
Phlebotomy[(b)]	1	£3.42	NHS reference costs 2013-14[30]
	Total per patient	£6.42	

(a) Includes medical and staffing cost involved in analysing the result
(b) Includes staff time and equipment required to take the blood

14.1.7 Evidence statements

14.1.7.1 Clinical

For forest plots, see Section J.6.1 in Appendix J.

14.1.7.1.1 Intervention review

No relevant studies were identified.

14.1.7.1.2 Prognostic review

Arteriovenous fistula surgery

One retrospective cohort found preoperative HbA1c >7% to be a predictor of primary arteriovenous fistula failure (OR 2.79 [1.31-5.32]) in multivariate analysis [Low quality]

Joint arthroplasty

One retrospective cohort found preoperative HbA1c ≥7% to be a predictor of 90-day mortality (OR=1.37 [0.82-2.29]) in multivariate analysis (adjusted for 38 variables) [Very low quality]. The same study found preoperative HbA1c ≥7% to be an independent predictor of complications (OR=1.22 [1.01-1.47]) and a weak predictor of total number of complications (OR=1.18 [0.97-1.43]) in multivariate analysis (adjusted for 38 variables) [Low to Very low quality].

A further retrospective cohort found that preoperative HbA1c >7% was not predictive of periprosthetic joint infection (HR 0.86 [0.68-1.09]). However, the same study found preoperative HbA1c >7% to be a weak predictor of death (HR 1.3 [1.083-1.564]) [Very low quality]

Non-cardiac surgery

One retrospective cohort found preoperative HbA1c <7% to be a predictor of reduced postoperative infections complications (OR=2.13 [1.23 to 3.69]) in multivariate analysis [Low quality]

14.1.7.2 Economic

No relevant economic evaluations were identified.

14.1.8 Delphi survey results

Preoperative HbA1c tests were included in the modified Delphi consensus survey.

The survey participants were asked if HbA1c tests should be used as a routine preoperative test for patients with diabetes undergoing elective surgery. Participants rated their response from strongly disagree to strongly agree using a nine-point Likert scale. Each question was considered to have reached consensus if greater than 70% of responses were in a single category (0–3 strongly disagree, 4–6 unclear, 7–9 strongly agree). Please see Appendix L for full details on the survey method and results.

14.1.8.1 Delphi statements where consensus was reached

Table 121: Patients: ASA2 with diabetes comorbidity

Surgery grade	Results % (round in which consensus was achieved)
Major or complex surgery	75.72 strongly agree (round 2)

Table 122: Patients: ASA3 or ASA4 with diabetes comorbidity

Surgery grade	Results % (round in which consensus was achieved)
Intermediate surgery	71.83 strongly agree (round 3)
Major or complex surgery	77.47 strongly agree (round 2)

14.1.8.2 Delphi statements where consensus was not reached

Table 123: Patients: ASA2 with diabetes comorbidity

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Minor surgery	1	59.45	28.37
Intermediate surgery		44.59	39.12
Minor surgery	2	45.68	33.33
Intermediate surgery		25.0	55.0
Minor surgery	3	39.43	38.02
Intermediate surgery		11.27	52.12

Table 124: Patients: ASA3 or ASA4 with diabetes comorbidity

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Minor	1	50.34	36.05

Surgery grade		Results %	
	2	30.38	49.36
	3	23.94	57.75

14.1.9 Recommendations and link to evidence

Recommendations	**14. People with diabetes who are being referred for surgical consultation from primary care should have their most recent HbA1c test results included in their referral information.** **15. Offer HbA1c testing to people with diabetes having surgery if they have not been tested in the last 3 months.** **Research recommendation:** **3. Does optimisation of HbA1c in people with poorly controlled diabetes improve surgical outcomes?**
Relative values of different outcomes	The GDG considered all-cause mortality to be a critical outcome for the intervention and prognostic reviews. For the intervention review health-related quality of life was also considered critical. Complications related to surgery or anaesthesia, length of hospital stay after an operation, hospital readmission, adverse events caused by testing and ICU admission were considered to be important outcomes for both the intervention and prognostic reviews with the addition of health-related quality of life as an important outcome for the prognostic review.
Trade-off between clinical benefits and harms	The GDG noted that the studies were in patients undergoing different surgery types, but that many surgery types were not included. It was noted that one of the included surgery types, surgical treatment of diabetic foot osteomyelitis, was a complication of poor diabetes control. This already indicates that the study population had a history of poor diabetes control. The GDG discussed the thresholds used by the included studies. The studies used an upper threshold of 7% (53 mmol/mol) with a HbA1c level above this considered to indicate poor diabetes control. Some GDG members felt that using a threshold of 8% (64 mmol/mol), or examining HbA1c as a continuous measure, might yield different results. The GDG also considered if there was any impact on the decision to continue with surgery as planned, based on the test results. If surgery is delayed in order to optimise control of the patient's diabetes, there is a need to consider any potential consequences of delaying surgery. It is also important to note that there is no guarantee that every patient will achieve improved diabetic control during this delay. The GDG noted that it is not currently known whether optimisation, and to what level, improves postoperative outcome.
Economic evidence	There were no economic evaluations that addressed this review question.
Quality of evidence	The reported outcomes of the included studies were all very low quality, including four retrospective and one prospective cohort. The studies all conducted multivariate analysis but adjusted for different confounders. There was inconsistency in outcomes. For example, the evidence concerning postoperative infection in non-cardiac surgery included one study that showed

	HbA1c was an independent predictor of postoperative infection, and one study that did not show a predictive ability of HbA1c on postoperative infection. The GDG identified a need for higher quality evidence from RCTs that covers different populations in different surgical contexts. The GDG were unable to make a recommendation based on the clinical evidence and decided to include this test in the Delphi survey.
Other considerations	Patients should already be having their diabetes optimally managed by their GP regardless of whether they are undergoing surgery or not. The GDG commented that GPs should share information about the patient's diabetes control when referring them for surgery. The GDG noted that some hospitals currently optimise patients based on their HbA1c levels and have programmes that focus on diet, lifestyle and pharmacological therapies. The GDG recognised that patients do not like variable rate intravenous insulin infusions, but noted that these are often used perioperatively in patients with poorly controlled diabetes. Patient experience could potentially be improved if diabetes control were improved preoperatively, reducing the requirement for variable rate intravenous insulin infusions. Patient groups have also noted that perioperative management of patients with diabetes is often suboptimal (see Management of adults with diabetes undergoing surgery and elective procedures: Improving standards[31])
Delphi	The GDG considered the Delphi survey results for patients with diagnosed diabetes and agreed that knowing the HbA1c level of a patient with diabetes was important for perioperative management. The GDG acknowledged the areas where consensus was reached and noted the divided opinion on whether to test people having minor or intermediate surgery. The GDG agreed that all patients should have their HbA1c measured regularly by their GP and if a recent test result was not available the test should be undertaken for all surgery types.
Economic considerations	In the absence of evidence, unit costs were provided alongside the results of the Delphi survey. The unit cost of measuring HbA1c was found to be £6.42. The GDG noted that patients should have their most recent HbA1c test results included in their referral information. Where this is provided and from the last 3 months, there is no clinical benefit for testing HbA1c again. Where this information is not available, HbA1c testing in patients diagnosed with diabetes would be a cost-effective use of NHS resources as the information could be used to improve health outcomes. This was supported by strong consensus from the Delphi survey.

14.2 HbA1c testing in people without diagnosed diabetes

14.2.1 Review question (intervention): What is the clinical- and cost-effectiveness of using HbA1c (glycated haemoglobin) as a preoperative test in improving patient outcomes in adults and young people with mild to severe comorbidities undergoing non-cardiac elective surgery?

Table 125: PICO characteristics of review question

Population	Adult patients without diagnosed diabetes (all types) undergoing non-cardiac related

	surgery Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory, renal, obesity
Intervention	HbA1c (glycated haemoglobin)
Comparison(s)	No HbA1c (glycated haemoglobin)/clinical assessment only
Outcomes	**Critical:** • All-cause mortality • Health-related quality of life **Important:** • Complications related to surgery or anaesthesia (for example arrhythmias, myocardial infarction, heart failure, respiratory failure, acute kidney failure, infection) • Length of hospital stay after an operation • Hospital readmission • Intensive care unit (ICU) admission

14.2.2 Clinical evidence

No relevant clinical studies comparing preoperative HbA1c testing in patients without diagnosed diabetes with no preoperative HbA1c testing in patients without diagnosed diabetes were identified.

14.2.3 Review question (prognostic): Does HbA1c (glycated haemoglobin) predict prognosis (patient outcomes after surgery) of people with mild to severe comorbidities undergoing major or complex non-cardiac elective surgery?

Table 126: PICO characteristics of review question

Population	Adult patients without diagnosed diabetes (all types) undergoing non-cardiac related surgery Stratified analysis if data available for: • Surgery type or surgery grade (if specified) • ASA grade • Selected comorbidities: cardiovascular, respiratory, renal, obesity
Prognostic test	Level of HbA1c (glycated haemoglobin)
Key confounding factors	*Minimum set of confounders that should be adjusted for (will vary per outcome)* • Age • BMI • Comorbidities (cardiovascular, respiratory, renal, obesity) • Patients taking drugs that cause a rapid rise in glucose (such as corticosteroids or antipsychotic drugs (≤2 months). HbA1c can be used in patients taking these drugs longer term (>2 months) who are not clinically unwell. • Ethnic groups • Patients with acute pancreatic damage or who have undergone pancreatic surgery • Patients with renal failure • Patients with HIV infection
Outcomes	**Critical:** • All-cause mortality

Important:

- Complications relating to surgery or anaesthesia
- Length of hospital stay (post-operation)
- Hospital readmission
- Adverse events after surgery (wound infection)
- Health-related quality of life
- Intensive care unit (ICU) admission

14.2.4 Clinical evidence

A single prospective observational study for colorectal surgery was included in the prognostic review and is summarised below.

The main surgical procedures within this surgery were arterior resection, abdominoperineal resection, total colectomy, right hemicolectomy, left hemicolectomy, and other resection. Patients in this study were grouped based on the preoperative measurements of HbA1c: 31 patients were found to have HbA1c above the normal range (over 6%) and 89 patients had HbA1c within the normal range (4.5–6%).

See also the clinical study selection flow chart in Appendix E.

Table 127: Summary of studies included in the review

Study	Study design/ sample size (response rate)/ Population characteristics	Analysis	Prognostic variable	Confounders list	Outcomes	Limitations
Major colorectal surgery						
Gustafsson 2009[45]	Prospective cohort n=120 (85.1%)/range of age: 31–90	Univariate and multi-variate (multiple logistic regression)	HbA1c (two categories: >6% and within normal range (4.5–6%) measured 3 months before surgery	Age, sex, BMI, ASA grade, preoperative bleeding and duration of surgery	Surgical complications Infection	The outcome of complications was heterogeneous and the authors stated that although they found differences across the different types of complications based on the HbA1c grouping, this finding was not further explored.

Table 128: Clinical evidence summary

Risk factors/outcomes/population	Number of studies	Pooled effect with 95% CIs [if meta-analysed] OR Effect and CI in single study	Imprecision	GRADE
HbA1c <6% versus HbA1c >6% for predicting surgical complications	1	Adjusted OR [95% CI]: 2.51 (1.07 to 5.90)	No serious imprecision	Low[a]
HbA1c <6% versus HbA1c >6% for predicting infection	1	Adjusted OR [95% CI]: 2.02 (0.78 to 5.24)	Serious imprecision	Very Low[a,b]

[a] *Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias*
[b] *Imprecision was considered serious if the confidence intervals crossed the null line*

14.2.5 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

Unit costs

Unit costs were provided for consideration alongside the Delphi survey results. Please see Appendix M for details. These are reported in Table 129 below.

Table 129: Blood glucose (HbA1c)

Equipment/staff	Quantity (units/minutes)	Cost	Source
Haematology[(a)]	1	£3.00	NHS reference costs 2013-14[30]
Phlebotomy[(b)]	1	£3.42	NHS reference costs 2013-14[30]
	Total per patient	£6.42	

(a) Includes medical and staffing cost involved in analysing the result
(b) Includes staff time and equipment required to take the blood

14.2.6 Evidence statements

14.2.6.1 Clinical

For forest plots, see Section J.6.2 in Appendix J.

14.2.6.1.1 Intervention review

No relevant clinical studies were identified.

14.2.6.1.2 Prognostic review

A single prospective study demonstrated that for patients undergoing major colorectal surgery, preoperative HbA1c at a 6% threshold could predict postsurgical complications but not length of hospital stay or infection after surgery [Moderate quality].

14.2.6.2 Economic

- No relevant economic evaluations were identified.

14.2.7 Delphi survey results

Preoperative HbA1c tests were included in the modified Delphi consensus survey.

The survey participants were asked if HbA1c tests should be used as a routine preoperative test for patients undergoing elective surgery. Participants rated their response from strongly disagree to strongly agree using a nine-point Likert scale. Each question was considered to have reached consensus if greater than 70% of responses were in a single category (0–3 strongly disagree, 4–6 unclear, 7–9 strongly agree). Please see Appendix L for full details on the survey method and results.

14.2.7.1 Delphi statements where consensus was reached

Table 130: Patients: ASA1

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	97.29 strongly disagree (round 1)
Intermediate surgery	95.27 strongly disagree (round 1)
Major or complex surgery	92.52 strongly disagree (round 1)

Table 131: Patients: ASA2 with cardiovascular, respiratory, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	96.62 strongly disagree (round 1)
Intermediate surgery	83.67 strongly disagree (round 1)
Major or complex surgery	76.35 strongly disagree (round 1)

Table 132: Patients: ASA3 or ASA4 with cardiovascular, respiratory, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	85.14 strongly disagree (round 1)
Intermediate surgery	78.08 strongly disagree (round 1)
Major or complex surgery	73.10 strongly disagree (round 1)

14.2.8 Recommendations and link to evidence

Recommendations	16. Do not routinely offer HbA1c testing before surgery to people without diagnosed diabetes.
Relative values of different outcomes	The GDG considered all-cause mortality and health-related quality of life to be critical outcomes for the intervention review. Only all-cause mortality was considered critical for the prognostic question. Complications related to surgery or anaesthesia, length of hospital stay after an operation, hospital readmission and ICU admission were considered to be important outcomes for both the intervention and prognostic reviews with the addition of health-related quality of life and adverse events after surgery as important outcomes for the prognostic review.
Trade-off between clinical benefits and harms	The GDG discussed the prognostic evidence which indicated that HbA1c testing could predict postsurgical complications, but agreed that the evidence did not allow them to justify blanket testing. The GDG commented that people with undiagnosed diabetes have poorer surgical outcomes than those with diagnosed diabetes (in the general population rather than the preoperative population specifically). They considered whether the results of HbA1c testing enable clinicians to predict individual patient outcomes: that is, does the HbA1c level indicate the likelihood of death and/or perioperative complications, but ultimately had limited evidence to suggest this was the case. The GDG were also interested in studies that demonstrated if perioperative management was altered in some way based on the test, which impacted on patient outcomes.
Economic evidence	No economic evaluations were identified for this question. The GDG considered the clinical evidence presented to be inconclusive, with no comparative studies available.
Quality of evidence	There were no comparative studies found which examined HbA1c testing versus no HbA1c testing in the intervention review.

	A single study with low quality evidence was found for the prognostic review. The GDG agreed that there was not enough robust evidence to inform the debate and therefore decided to add this to the Delphi survey in order to gain a wider consensus view on the use of the test for people without diagnosed diabetes.
Other considerations	The GDG noted that HbA1c is now routinely measured in the standardised international unit of mmol/mol. However, many of the studies reported HbA1c in units of percentage of total haemoglobin. The GDG noted that the prevalence of diagnosed diabetes is approximately 6% in the general UK population (not the preoperative population).[32] In addition, it is estimated that a further 1–2% of the general population (not the preoperative population) have undiagnosed Type 2 diabetes.[33] However, the prevalence of diagnosed and undiagnosed diabetes among the preoperative population is currently unknown. The GDG noted that preoperative HbA1c measurement should not be used as a screening tool for diabetes within this context and agreed that it would be most appropriate to focus on groups that are considered to be at high risk of diabetes, rather than the whole preoperative population. The GDG noted that the following factors may indicate an increased risk of type 2 diabetes: age, obesity, personal history of hypertension, ischaemic heart disease, cerebrovascular disease, hypothyroidism or hyperlipidaemia, personal history of gestational diabetes or polycystic ovary syndrome, family history of diabetes, ethnicity (for example there is a higher prevalence of diabetes in South Asian, Chinese, African-Caribbean and Black African populations). The GDG discussed that the UK National Screening Committee does not recommend universal screening but does recommend selective screening based on a diabetes risk assessment.[91]
Delphi	The results of the Delphi survey were in accordance with the consensus of the GDG and supported the recommendation that HbA1c testing should not be considered preoperatively in patients not previously diagnosed with diabetes.
Economic considerations	In the absence of evidence, unit costs were provided alongside the results of the Delphi survey. The GDG considered the unit cost of HbA1c testing (£6.42) and the prevalence of undiagnosed diabetes in people admitted for elective surgery. The GDG felt that screening all adults undergoing all grades and types of surgery would not represent a cost-effective use of NHS resources. The GDG noted that there were pathways in place for routine diabetic screening and these should be capturing the majority of undiagnosed cases. The Delphi survey supported the view of not conducting this test routinely, adding strength to the notion that the clinical benefits of doing so are small.

15 Sickle cell disease or sickle cell trait tests

15.1 Introduction

Sickle cell disease/trait testing involves venous blood sampling to detect haemoglobinopathies including sickle cell anaemia, which may have clinical implications in the perioperative setting. The test is considered safe and is straightforward to perform and analyse, but may be painful for the patient. There is a low risk of complications including haematoma formation, vasovagal reactions and infection. Universal newborn screening is performed in the UK to detect major haemoglobinopathies, so many individuals will already be aware of their condition(s). However, it is uncertain whether routine preoperative sickle disease/trait screening should be offered to all individuals in particularly affected ethnic groups to detect sickle cell disease/trait in those who are unaware of their newborn screening results or who were born outside the UK.

See section 4.4 for a summary of the methodological approach taken for this preoperative test.

15.2 Delphi survey results

As no new evidence on the use of sickle cell disease/trait tests as a routine preoperative test was identified during the scoping phase of this guideline it was decided not to carry out an evidence review, but to include sickle cell disease/trait tests in the modified Delphi survey to re-evaluate the consensus held amongst health professionals on the value of routinely conducting the test prior to elective surgery.

The survey participants were asked if screening for sickle cell disease/trait should be undertaken as a routine preoperative test for certain groups of patients undergoing elective surgery. Participants rated their response from strongly disagree to strongly agree using a nine-point Likert scale. Each question was considered to have reached consensus if greater than 70% of responses were in a single category (0–3 strongly disagree, 4–6 unclear, 7–9 strongly agree). Please see Appendix L for full details on the survey method and results.

15.2.1 Delphi statements where consensus was achieved

Table 133: Patients of West African origin

Surgery grade	Results % (round in which consensus was achieved)
All surgery grades	78.95 strongly agree (round 1)

Table 134: Patients of South/sub-Saharan African origin

Surgery grade	Results % (round in which consensus was achieved)
All surgery grades	79.56 strongly agree (round 1)

Table 135: Patients of African/Caribbean origin

Surgery grade	Results % (round in which consensus was achieved)
All surgery grades	84.21 strongly agree (round 1)

15.2.2 Delphi statements where consensus was not achieved

Table 136: Patients of North African origin

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Minor	1	41.25	42.65
	2	22.07	48.06
	3	16.17	69.12

15.3 Economic evidence

Unit costs were provided for consideration alongside the Delphi survey results. Please see Appendix M for details. These are reported in Table 137 below.

Table 137: Sickle cell

Equipment/staff	Quantity (units/minutes)	Cost	Source
Sickle cell lab testing[(a)]	1	£4	HTA 2010[34]
Phlebotomy[(b)]	1	£3.42	NHS reference costs 2013-14[30]
	Total per patient	£7.42	

(a) This cost is from a published source and does not include staff and equipment costs of taking blood
(b) Includes medical and staffing cost involved in the procedure

15.4 Recommendations and link to evidence

Recommendations	**17. Do not routinely offer testing for sickle cell disease or sickle cell trait before surgery.** **18. Ask the person having surgery if they or any member of their family have sickle cell disease.** **19. If the person is known to have sickle cell disease and has their disease managed by a specialist sickle cell service, liaise with this team before surgery.**
Delphi	The GDG discussed the results from the Delphi survey, querying in particular the strength of agreement with sickle cell testing. The group felt this was likely to reflect traditional medical teaching, which tends to be based on historical studies from the 1950s and does not necessarily reflect the current state of society in which there is increased movement of people and increased mixed race populations. The GDG debated the implications of sickle cell trait and suggested that screening sickle cell trait would not influence perioperative management, and was of no added value as a preoperative test. As screening for sickle cell disease is offered to all neonates in the NHS, adults are very unlikely to have undiagnosed sickle cell disease. Moreover, the GDG felt that even when missed, it is likely that most people with sickle cell disease would have experienced symptoms before the age of 16. People with known sickle cell disease will be linked to specialist teams for management of the condition. Intraoperative hypoxia is unlikely with modern anaesthesia and therefore there is a low risk of

	sickle cell crisis in patients with mild sickle cell disease. Anyone with more severe forms of the disease will be already known to specialist teams. On this basis, the GDG felt that clinicians should first enquire as to whether the person has ever experienced symptoms of sickle cell disease. This is of particular relevance in cases where the patient's clinical history is uncertain and a full blood count test is not being done. The GDG therefore agreed that routine testing for sickle cell disease should not be done, but that clinicians should use their judgement to determine whether a test is necessary.
Economic considerations	The cost of performing a sickle cell test was found to be £7.42, including staff time and equipment. The GDG felt that most individuals with sickle cell disease would most likely be diagnosed before the age of 16. In the rare circumstance whereby the individual had not been diagnosed through childhood screening, the GDG felt that sickle cell disease would be very symptomatic and easily identifiable, reducing the need for routine testing. The GDG also felt that the occurrence of sickle cell trait would not change management if detected, therefore leading to no change in health outcomes post-surgery. It is therefore unlikely that routinely performing sickle cell testing is a cost-effective use of NHS resources.

16 Urinalysis

16.1 Introduction

Urinalysis is the physical, chemical and microscopic analysis of urine. In the preoperative setting, it may be used to detect urinary tract infections, renal diseases and poorly controlled diabetes. The test is safe with no known risks. However, it is uncertain whether the test provides valuable information in asymptomatic individuals, and other more specific tests may be used to diagnose and monitor diabetes and renal disease. In addition, the clinical effectiveness of routine preoperative urinalysis is uncertain.

See section 4.4 for a summary of the methodological approach taken for this preoperative test.

16.2 Delphi survey results

As no new evidence on the use of urine dipstick tests as a routine preoperative test was identified during the scoping phase of this guideline it was decided not to carry out an evidence review, but to include urinalysis in the modified Delphi survey to re-evaluate the consensus held amongst health professionals on the value of routinely conducting the test prior to elective surgery.

The survey participants were asked if urine tests (urine dipstick) should be used as a routine preoperative test for patients undergoing elective surgery. Participants rated their response from strongly disagree to strongly agree using a nine-point Likert scale. Each question was considered to have reached consensus if greater than 70% of responses were in a single category (0–3 strongly disagree, 4–6 unclear, 7–9 strongly agree). Please see Appendix L for full details on the survey method and results.

16.2.1 Delphi statements where consensus was reached

Table 138: Patients: ASA1

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	86.45 strongly disagree (round 1)
Intermediate surgery	77.41 strongly disagree (round 1)

Table 139: Patients: ASA2 with cardiovascular, respiratory, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	75.32 strongly disagree (round 1)

Table 140: Patients: ASA3 or ASA4 with cardiovascular, respiratory, renal or obesity comorbidities

Surgery grade	Results % (round in which consensus was achieved)
Minor surgery	70.20 strongly disagree (round 1)

16.2.2 Delphi statements where consensus was not reached

Table 141: Patients: ASA1

Surgery grade		Results %

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Major or complex	1	66.87	26.11
	2	59.26	28.4
	3	63.76	23.2

Table 142: Patients: ASA2 with diabetes

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Minor surgery	1	69.62	22.78
Intermediate surgery		63.05	29.75
Major or complex surgery		61.54	34.62
Minor surgery	2	54.43	21.52
Intermediate surgery		45.57	35.44
Major or complex surgery		42.86	30.01
Minor surgery	3	55.88	25.0
Intermediate surgery		44.12	38.24
Major or complex surgery		32.84	50.75

Table 143: Patients: ASA2 with cardiovascular, respiratory, renal or obesity comorbidities

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Intermediate surgery	1	68.35	23.42
Major or complex surgery		63.69	33.75
Intermediate surgery	2	50.63	24.04
Major or complex surgery		39.44	28.17
Intermediate surgery	3	52.18	23.19
Major or complex surgery		40.58	42.03

Table 144: Patients: ASA3 or ASA4 with diabetes

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Minor surgery	1	64.56	27.22
Intermediate surgery		61.15	31.85
Major or complex surgery		60.65	34.19
Minor surgery	2	43.75	33.75
Intermediate surgery		37.5	43.75
Major or complex surgery		32.4	42.25
Minor surgery	3	43.29	32.84

Surgery grade		Results %	
Intermediate surgery		36.23	43.47
Major or complex surgery		31.34	58.21

Table 145: Patients: ASA3 or ASA4 with cardiovascular, respiratory, renal or obesity comorbidities

Surgery grade		Results %	
	Round of Delphi	Strongly disagree	Strongly agree
Intermediate surgery	1	65.19	27.21
Major or complex surgery		62.42	35.03
Intermediate surgery	2	41.78	36.7
Major or complex surgery		35.75	41.43
Intermediate surgery	3	42.03	30.44
Major or complex surgery		36.23	42.03

16.3 Economic evidence

Unit costs were provided for consideration alongside the Delphi survey results. Please see Appendix M for details. These are reported in Table 146 and Table 147 below.

Table 146: Unit cost of urine test (using dipstick)

Equipment/staff	Quantity (units/minutes)	Cost	Source
Strip[(a)]	1	£0.22	NHS supply chain catalogue 2014[3]
Container/collector[(a)]	1	£0.67	NHS supply chain catalogue 2014[3]
Gloves[(a)]	1	£0.03	NHS supply chain catalogue 2014[3]
Apron[(a)]	1	£0.09	NHS supply chain catalogue 2014[3]
Nurse time[(b)]	5	£2.83	PSSRU 13/14[23]
	Total per patient	£3.85	

(a) Taken as an average price from available equipment on the NHS supply chain catalogue
(b) Nurse time based on day ward nurse costing £34 per hour

Table 147: Unit cost of urine test (using urinalysis analyser)

Equipment/staff	Quantity (units/minutes)	Cost	Source
Urinalysis analyser[(a)]	1	£0.45	NHS supply chain catalogue 2014[3]
Container/collector[(b)]	1	£0.67	NHS supply chain catalogue 2014[3]
Gloves[(b)]	1	£0.03	NHS supply chain catalogue 2014[3]
Apron[(b)]	1	£0.09	NHS supply chain

Equipment/staff	Quantity (units/minutes)	Cost	Source
			catalogue 2014[3]
Nurse time[(c)]	5	£2.83	PSSRU 13/14[23]
	Total per patient	£4.08	

(a) The machine costs £44,500 and if we assume it is used 100,000 times before it is replaced the marginal cost of using this machine equates to £0.45 per patient.
(b) Taken as an average price from available equipment on the NHS supply chain catalogue
(c) Nurse time based on day ward nurse costing £34 per hour

16.4 Recommendations and link to evidence

Recommendations	**20. Do not routinely offer urine dipstick tests before surgery.** **21. Consider microscopy and culture of midstream urine sample before surgery if the presence of a urinary tract infection would influence the decision to operate.**
Delphi	The GDG discussed the results of the Delphi survey and felt they were more conservative than the consensus of the GDG. While consensus was not reached to 'offer' the test to any group, the majority opinion from the Delphi survey was to consider the test. The GDG believed that this was probably due to historical use of urine dipstick tests to pick up UTIs. However the GDG noted that urine dipstick tests are not sensitive or specific in the diagnosis of UTIs. A midstream urine sample (MSU) is considered to be the definitive diagnostic test for UTI. If a UTI would influence surgical decision-making, then the GDG suggests performing an MSU within an appropriate timeframe for the surgery. The GDG also discussed the value of the urine dipstick test as a quick and cheap screen for diabetes, however in keeping with other guidance, agreed that urine dipsticks should not be used routinely for screening or diagnosis of diabetes mellitus.
Economic considerations	The cost of performing a urine dipstick test was found to be £3.85, including staff time and equipment. The cost is based on a dipstick test measuring 10 parameters (urine protein, glucose, nitrite, haemoglobin, ketones, bilirubin, urobilinogen, leukocytes, pH and specific gravity). The cost of using a urinalysis analyser was found to be £4.07. The GDG considered the cost of potential further investigations that could arise due to an abnormal result including ultrasounds (cost at £59) or a urology outpatient visit (cost at £99). Based on the results of the Delphi and GDG opinion, it is unlikely to be cost-effective to carry out a urine test in the majority of patients prior to elective surgery. The GDG felt that urine dipstick tests had poor diagnostic accuracy for identifying complications, and in the majority of cases identifying complications would not lead to a change in management that would improve health outcomes. The GDG recognised that urine testing potentially holds screening benefits, such as diagnosing early diabetes. However, this is not a gold standard diagnostic tool and there are already screening pathways in place to pick up such diseases. The GDG noted however that identifying a urinary tract infection (UTI) could have an impact on some individuals undergoing certain types of surgery, so narrowing down testing to this subset of patients could be a cost-effective use of NHS resources.

17 Pregnancy testing

17.1 Introduction

Pregnancy testing may be offered to women of child-bearing age in the preoperative setting, as elective surgery is generally avoided unless absolutely necessary in pregnant women due to the risks of teratogenicity and miscarriage. Pregnancy testing is usually performed using a urine test, but a venous blood test may also be used. The tests are considered safe, but false negative results may occur especially in the first few weeks of pregnancy, and false positive results may also occur rarely. There is uncertainty about whether preoperative pregnancy testing should be performed on all women of child-bearing age, or only on those individuals who are unsure if they may be pregnant.

See section 4.4 for a summary of the methodological approach taken for this preoperative test.

17.2 Delphi survey results

Pregnancy testing was included in the consensus survey conducted for the original 2003 guideline. For this update, pregnancy testing was included in the modified Delphi survey to re-evaluate the consensus held amongst health professionals.

The survey participants were asked what criteria should be used to determine whether or not to offer pregnancy testing as a routine preoperative test for women undergoing elective surgery. Participants were asked to respond narratively, outlining in what circumstances they would offer the test. Please see Appendix L for full details on the survey method and results.

Table 148: Narrative summary

Round	Common themes
1	Common themes included: patient consent should be considered, but the importance of carrying out the test should be stressed to women of child bearing potential due to the increased risk of X-ray to the foetus.
2	Generally carried out on women of child bearing potential, unless unequivocal evidence to suggest the impossibility of pregnancy, for example women who have undergone hysterectomy. Patient consent should be sought.

17.3 Economic evidence

Unit costs were provided for consideration alongside the Delphi survey results. Please see Appendix M for details. These are reported in Table 149 below.

Table 149: Unit cost of pregnancy test

Equipment/staff	Quantity (units/minutes)	Cost	Source
Pregnancy test kit[(a)]	1	£0.69	NHS supply chain catalogue 2014[3]
Nurse[(b)]	5	£2.83	PSSRU 13/14[23]
	Total per patient	£3.52	

(a) Average cost of pregnancy test kits available through NHS supply chain catalogue
(b) Nurse time based on day ward nurse costing £34 per hour

17.4 Recommendations and link to evidence

Recommendations	**22. On the day of surgery, sensitively ask all women of childbearing potential whether there is any possibility they could be pregnant.** **23. Make sure women who could possibly be pregnant are aware of the risks of the anaesthetic and the procedure to the fetus.** **24. Document all discussions with women about whether or not to carry out a pregnancy test.** **25. Carry out a pregnancy test with the woman's consent if there is any doubt about whether she could be pregnant.** **26. Develop locally agreed protocols for checking pregnancy status before surgery.** **27. Make sure protocols are documented and audited, and in line with statutory and professional guidance.**
Delphi	The need to test for pregnancy depends on the risk presented to the woman and the fetus by the anaesthetic and the procedure.[79] The GDG felt it was necessary to define relevant female patients for pregnancy testing, as in some cases (for example women who have had a hysterectomy) the test would have no value. The GDG agreed that women of 'childbearing potential' rather than 'childbearing age', is a better description. For this guideline, the lower age boundary is 16 years (this guideline does not cover people under 16 years) while the upper boundary needs to be consistent with the NICE menopause guideline (NG23), which defines menopause as having occurred in healthy women over 45 years of age who have not had a period for at least 12 consecutive months . The mean age of natural menopause is 51 years.[75] The GDG debated the pros and cons of blanket pregnancy testing. The GDG noted that blanket testing avoided misinterpretation and awkwardness when asking the pregnancy question, but could be construed as offensive if the woman had already stated she was not pregnant. The GDG agreed that in most cases (excluding specific circumstances such as women who have had a hysterectomy) it is important that the patient is asked if there was a chance they could be pregnant before undergoing certain procedures. It was noted that certain tests (for example radiography) cannot be requested without a specific response regarding pregnancy. The GDG commented that the way patients are asked about pregnancy is important. They agreed that sensitivity and consideration of the individual's circumstances should be employed whenever the question was asked. A good approach would be to set a local protocol on pregnancy testing which can then be referred to when a woman is asked about her pregnancy status, although the decision on whether to undertake the test must always be dependent on valid consent being obtained.

	The GDG also expressed concern about whether there was a possibility of missing some pregnancies due to misinterpretation of the pregnancy question. The GDG agreed that women should be made aware of the risks and potential consequences if they say they are not pregnant but they are. The GDG did not deem that asking a female patient for the date of their last menstrual period was an effective means of establishing pregnancy status, due to the varying reliability of responses. The GDG noted the specific issues that may arise in relation to checking pregnancy status of young women aged 16-17, as outlined in 'Pre-procedure Pregnancy Checking in Under 16s: Guidance for Clinicians'.[84] This guidance has an upper age limit of 16 as this is the legal age of consensual sexual activity in the UK; however the report states that many of the references will be relevant to all children, including those up to 18 years. The GDG discussed the issues and agreed to it was important that local protocols are developed to ensure documented and audited compliance with professional guidance on checking pregnancy status. The local protocol should set out, for example, the criteria for enquiry or consented testing, what information is provided to patients, how pregnancy status is recorded and the procedures for management of consent and disclosure, particularly for groups who may find discussion of pregnancy a sensitive issue.
Economic considerations	The cost of performing a pregnancy test was found to be £3.52 including staff time and equipment. Identifying a pregnancy prior to surgery will have significant outcomes even for minor surgeries, depending on the anaesthetic used. However, the GDG felt that appropriate screening through questions and medical history should limit the number of individuals offered this test. Where there is any doubt, offering a pregnancy test could avoid serious complications and therefore represents a cost-effective use of NHS resources.

18 Reference list

1 International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. Diabetes Care. 2009; 32(7):1327-1334

2 Pre-operative assessment and patient preparation: the role of the anaesthetist 2. Anaesthetists of Great Britain and Ireland, 2010

3 NHS Supply Chain Catalogue. NHS Supply Chain, 2013. Available from: http://www.supplychain.nhs.uk/

4 Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen. The Nordic Cochrane Centre, The Cochrane Collaboration, 2014. Available from: http://tech.cochrane.org/Revman

5 AbuRahma AF, Srivastava M, Chong B, Dean LS, Stone PA, Koszewski A. Impact of chronic renal insufficiency using serum creatinine vs glomerular filtration rate on perioperative clinical outcomes of carotid endarterectomy. Journal of the American College of Surgeons. 2013; 216(4):525-3

6 Afsar B, Elsurer R. The primary arteriovenous fistula failure-a comparison between diabetic and non-diabetic patients: glycemic control matters. International Urology and Nephrology. 2012; 44(2):575-581

7 Amaranto DJ, Wang EC, Eskandari MK, Morasch MD, Rodriguez HE, Pearce WH et al. Normal preoperative white blood cell count is predictive of outcomes for endovascular procedures. Journal of Vascular Surgery. 2011; 54(5):1395-1403

8 ASA. ASA Physical Status Classification System. 2014. Available from: https://www.asahq.org/resources/clinical-information/asa-physical-status-classification-system [Last accessed: 17 December 2015]

9 Asua J, Lopez-Argumedo M. Preoperative evaluation in elective surgery. INAHTA synthesis report. International Journal of Technology Assessment in Health Care. 2000; 16(2):673-683

10 Ausania F, Snowden CP, Prentis JM, Holmes LR, Jaques BC, White SA et al. Effects of low cardiopulmonary reserve on pancreatic leak following pancreaticoduodenectomy. British Journal of Surgery. 2012; 99(9):1290-1294

11 Barakat HM, Shahin Y, McCollum PT, Chetter IC. Prediction of organ-specific complications following abdominal aortic aneurysm repair using cardiopulmonary exercise testing. Anaesthesia. 2015; 70(6):679-685

12 Beattie WS, Karkouti K, Wijeysundera DN, Tait G. Risk associated with preoperative anemia in noncardiac surgery: a single-center cohort study. Anesthesiology. 2009; 110(3):574-581

13 Bedke J, Chun FKH, Merseburger A, Scharpf M, Kasprzyk K, Schilling D et al. Inflammatory prognostic markers in clear cell renal cell carcinoma - preoperative C-reactive protein does not improve predictive accuracy. BJU International. 2012; 110(11 Pt B):E771-E777

14 Biteker M, Duman D, Tekkesin AI. Predictive value of preoperative electrocardiography for perioperative cardiovascular outcomes in patients undergoing noncardiac, nonvascular surgery. Clinical Cardiology. 2012; 35(8):494-499

15 Brunelli A, Belardinelli R, Pompili C, Xiume F, Refai M, Salati M et al. Minute ventilation-to-carbon dioxide output (VE/VCO2) slope is the strongest predictor of respiratory complications and death after pulmonary resection. Annals of Thoracic Surgery. 2012; 93(6):1802-1806

16 Brunelli A, Belardinelli R, Refai M, Salati M, Socci L, Pompili C et al. Peak oxygen consumption during cardiopulmonary exercise test improves risk stratification in candidates to major lung resection. Chest. 2009; 135(5):1260-1267

17 Carlisle J, Swart M. Mid-term survival after abdominal aortic aneurysm surgery predicted by cardiopulmonary exercise testing. British Journal of Surgery. 2007; 94(8):966-969

18 Carlisle JB, Stocker ME. Preoperative assessment. In: Smith I, McWhinnie D, Jackson I (eds), Day case surgery, London: Oxford University Press, 2012: 51-62

19 Chassot PG, Delabays A, Spahn DR. Preoperative evaluation of patients with, or at risk of, coronary artery disease undergoing non-cardiac surgery. British Journal of Anaesthesia. 2002; 89(5):747-759

20 Chrastil J, Anderson MB, Stevens V, Anand R, Peters CL, Pelt CE. Is Hemoglobin A1c or Perioperative Hyperglycemia Predictive of Periprosthetic Joint Infection or Death Following Primary Total Joint Arthroplasty? Journal of Arthroplasty. 2015; 30(7):1197-1202

21 Chung F, Yuan H, Yin L, Vairavanathan S, Wong DT. Elimination of preoperative testing in ambulatory surgery. Anesthesia and Analgesia. 2009; 108(2):467-475

22 Chung SA, Yuan H, Chung F. A systemic review of obstructive sleep apnea and its implications for anesthesiologists. Anesthesia and Analgesia. 2008; 107(5):1543-1563

23 Curtis L. Unit costs of health and social care 2012. Canterbury: Personal Social Services Research Unit, University of Kent; 2012. Available from: http://www.pssru.ac.uk/archive/pdf/uc/uc2012/full-with-covers.pdf

24 Czoski-Murray C, Lloyd Jones M, McCabe C, Claxton K, Oluboyede Y, Roberts J et al. What is the value of routinely testing full blood count, electrolytes and urea, and pulmonary function tests before elective surgery in patients with no apparent clinical indication and in subgroups of patients with common comorbidities: a systematic review of the clinical and cost-effective literature. Health Technology Assessment. 2012; 16(50):i-159

25 De Hert S, Imberger G, Carlisle J, Diemunsch P, Fritsch G, Moppett I et al. Preoperative evaluation of the adult patient undergoing non-cardiac surgery: Guidelines from the European Society of Anaesthesiology. European Journal of Anaesthesiology. 2011; 28(10):684-722

26 Department of Health. Good Practice in Consent Implementation Guide: Consent to Examination or Treatment, 2001

27 Department of Health. Reference guide to consent for examination or treatment (second edition), 2009

28 Department of Health. NHS reference costs 2010-11. 2012. Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_131140 [Last accessed: 27 March 2012]

29 Department of Health. NHS reference costs 2012-13. 2013. Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2012-to-2013 [Last accessed: 21 July 2014]

30 Department of Health. NHS reference costs 2013-14. 2014. Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2013-to-2014 [Last accessed: 3 September 2015]

31 Dhatariya K, Flanagan D, Hilton L, Kilvert A, Levy N, Rayman G et al. Management of adults with diabetes undergoing surgery and elective procedures: Improving standards. Joint British Diabetes Societies for Inpatient Care, 2015

32 Diabetes UK. Diabetes prevalence 2013 (February 2014). 2015. Available from: http://www.diabetes.org.uk/About_us/What-we-say/Statistics/Diabetes-prevalence-2013/ [Last accessed: 2 September 2015]

33 Diabetes UK. Diabetes: Facts and stats (version 4), 2015

34 Dormandy E, Bryan S, Gulliford MC, Roberts TE, Ades AE, Calnan M et al. Antenatal screening for haemoglobinopathies in primary care: a cohort study and cluster randomised trial to inform a simulation model. The Screening for Haemoglobinopathies in First Trimester (SHIFT) trial. Health Technology Assessment (Winchester, England). 2010; 14(20):1-160

35 Dronge AS, Perkal MF, Kancir S, Concato J, Aslan M, Rosenthal RA. Long-term glycemic control and postoperative infectious complications. Archives of Surgery. 2006; 141(4):375-380

36 Dunkelgrun M, Hoeks SE, Welten GMJM, Vidakovic R, Winkel TA, Schouten O et al. Anemia as an independent predictor of perioperative and long-term cardiovascular outcome in patients scheduled for elective vascular surgery. American Journal of Cardiology. 2008; 101(8):1196-1200

37 Farina A, Crimi E, Accogli S, Camerini G, Adami GF. Preoperative assessment of respiratory function in severely obese patients undergoing biliopancreatic diversion. European Surgical Research Europaische Chirurgische Forschung Recherches Chirurgicales Europeennes. 2012; 48(2):106-110

38 Fritsch G, Flamm M, Hepner DL, Panisch S, Seer J, Soennichsen A. Abnormal pre-operative tests, pathologic findings of medical history, and their predictive value for perioperative complications. Acta Anaesthesiologica Scandinavica. 2012; 56(3):339-350

39 Glance LG, Blumberg N, Eaton MP, Lustik SJ, Osler TM, Wissler R et al. Preoperative thrombocytopenia and postoperative outcomes after noncardiac surgery. Anesthesiology. 2014; 120(1):62-75

40 Goodnough LT, Maniatis A, Earnshaw P, Benoni G, Beris P, Bisbe E et al. Detection, evaluation, and management of preoperative anaemia in the elective orthopaedic surgical patient: NATA guidelines. British Journal of Anaesthesia. 2011; 106(1):13-22

41 Goodyear SJ, Yow H, Saedon M, Shakespeare J, Hill CE, Watson D et al. Risk stratification by pre-operative cardiopulmonary exercise testing improves outcomes following elective abdominal aortic aneurysm surgery: a cohort study. Perioperative Medicine. 2013; 2(1):10

42 GRADE Working Group. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group website. 2011. Available from: http://www.gradeworkinggroup.org/ [Last accessed: 1 October 2011]

43 Grant SW, Hickey GL, Wisely NA, Carlson ED, Hartley RA, Pichel AC et al. Cardiopulmonary exercise testing and survival after elective abdominal aortic aneurysm repair+. British Journal of Anaesthesia. 2015; 114(3):430-436

44 Greenky M, Gandhi K, Pulido L, Restrepo C, Parvizi J. Preoperative anemia in total joint arthroplasty: is it associated with periprosthetic joint infection? Clinical Orthopaedics and Related Research. 2012; 470(10):2695-2701

45 Gustafsson UO, Thorell A, Soop M, Ljungqvist O, Nygren J. Haemoglobin A1c as a predictor of postoperative hyperglycaemia and complications after major colorectal surgery. British Journal of Surgery. 2009; 96(11):1358-1364

46 Hamoui N, Anthone G, Crookes PF. The value of pulmonary function testing prior to bariatric surgery. Obesity Surgery. 2006; 16(12):1570-1573

47 Harris AHS, Bowe TR, Gupta S, Ellerbe LS, Giori NJ. Hemoglobin A1C as a marker for surgical risk in diabetic patients undergoing total joint arthroplasty. Journal of Arthroplasty. 2013; 28(8 Suppl):25-29

48 Harris DA, Al-Allak A, Thomas J, Hedges AR. Influence of presentation on outcome in abdominal aortic aneurysm repair. European Journal of Vascular and Endovascular Surgery. 2006; 32(2):140-145

49 Hartley RA, Pichel AC, Grant SW, Hickey GL, Lancaster PS, Wisely NA et al. Preoperative cardiopulmonary exercise testing and risk of early mortality following abdominal aortic aneurysm repair. British Journal of Surgery. 2012; 99(11):1539-1546

50 Huh J, Sohn TS, Kim JK, Yoo YK, Kim DK. Is routine preoperative spirometry necessary in elderly patients undergoing laparoscopy-assisted gastrectomy? Journal of International Medical Research. 2013; 41(4):1301-1309

51 Jamsen E, Nevalainen PI, Eskelinen A, Kalliovalkama J, Moilanen T. Risk factors for perioperative hyperglycemia in primary hip and knee replacements. Acta Orthopaedica. 2015; 86(2):175-182

52 Jans O, Jorgensen C, Kehlet H, Johansson PI, Lundbeck Foundation Centre for Fast-track Hip and Knee Replacement Collaborative Group. Role of preoperative anemia for risk of transfusion and postoperative morbidity in fast-track hip and knee arthroplasty. Transfusion. 2014; 54(3):717-726

53 Jeong O, Ryu SY, Park YK. The value of preoperative lung spirometry test for predicting the operative risk in patients undergoing gastric cancer surgery. Journal of the Korean Surgical Society. 2013; 84(1):18-26

54 Johansson T, Fritsch G, Flamm M, Hansbauer B, Bachofner N, Mann E et al. Effectiveness of non-cardiac preoperative testing in non-cardiac elective surgery: a systematic review. British Journal of Anaesthesia. 2013; 110(6):926-939

55 Junejo MA, Mason JM, Sheen AJ, Bryan A, Moore J, Foster P et al. Cardiopulmonary Exercise Testing for Preoperative Risk Assessment before Pancreaticoduodenectomy for Cancer. Annals of Surgical Oncology. 2014; 21(6):1929-1936

56 Junejo MA, Mason JM, Sheen AJ, Moore J, Foster P, Atkinson D et al. Cardiopulmonary exercise testing for preoperative risk assessment before hepatic resection. British Journal of Surgery. 2012; 99(8):1097-1104

57 Kaag M, Trost L, Thompson RH, Favaretto R, Elliott V, Shariat SF et al. Preoperative predictors of renal function decline after radical nephroureterectomy for upper tract urothelial carcinoma. BJU International. 2014; 114(5):674-679

58 Kinley H, Czoski-Murray C, George S, McCabe C, Primrose J, Reilly C et al. Effectiveness of appropriately trained nurses in preoperative assessment: randomised controlled equivalence/non-inferiority trial. BMJ. 2002; 325(7376):1323

59 Klein AA, Arrowsmith JE. Should routine pre-operative testing be abandoned? Anaesthesia. 2010; 65(10):974-976

60 Koike Y, Imaizumi H, Takahashi E, Matsubara Y, Komatsu H. Determining factors of mortality in the elderly with hip fractures. The Tohoku Journal of Experimental Medicine. 1999; 188(2):139-142

61 Kyo T, Takaoka K, Ono K. Femoral neck fracture. Factors related to ambulation and prognosis. Clin Orthop Relat Res. 1993;(292):215-222

62 Landesberg G, Berlatzky Y, Bocher M, Alcalai R, Anner H, Ganon-Rozental T et al. A clinical survival score predicts the likelihood to benefit from preoperative thallium scanning and coronary revascularization before major vascular surgery. European Heart Journal. 2007; 28(5):533-539

63 Licker M, Schnyder JM, Frey JG, Diaper J, Cartier V, Inan C et al. Impact of aerobic exercise capacity and procedure-related factors in lung cancer surgery. European Respiratory Journal. 2011; 37(5):1189-1198

64 Litton E, Xiao J, Ho KM. Safety and efficacy of intravenous iron therapy in reducing requirement for allogeneic blood transfusion: systematic review and meta-analysis of randomised clinical trials. BMJ. 2013; 347:f4822

65 Liu LL, Dzankic S, Leung JM. Preoperative electrocardiogram abnormalities do not predict postoperative cardiac complications in geriatric surgical patients. Journal of the American Geriatrics Society. 2002; 50(7):1186-1191

66 Mases A, Sabate S, Guilera N, Sadurni M, Arroyo R, Fau M et al. Preoperative estimated glomerular filtration rate and the risk of major adverse cardiovascular and cerebrovascular events in non-cardiac surgery. British Journal of Anaesthesia. 2014; 113(4):644-651

67 McCullough PA, Gallagher MJ, Dejong AT, Sandberg KR, Trivax JE, Alexander D et al. Cardiorespiratory fitness and short-term complications after bariatric surgery. Chest. 2006; 130(2):517-525

68 Mental Capacity Act. Codes of practice. London: Department for Constitutional Affairs, 2015

69 Munro J, Booth A, Nicholl J. Routine preoperative testing: a systematic review of the evidence. Health Technology Assessment. England 1997; 1(12):1-63

70 Mutter TC, Chateau D, Moffatt M, Ramsey C, Roos LL, Kryger M. A matched cohort study of postoperative outcomes in obstructive sleep apnea: could preoperative diagnosis and treatment prevent complications? Anesthesiology. 2014; 121(4):707-718

71 National Institute for Clinical Excellence. Preoperative testing: the use of routine preoperative tests for elective surgery. London: National Institute for Clinical Excellence (NICE). England 2003;30

72 National Institute for Health and Care Excellence. Patient experience in adult NHS services: improving the experience of care for people using adult NHS services , 2012

73 National Institute for Health and Care Excellence. Acute kidney injury: Prevention, detection and management of acute kidney injury up to the point of renal replacement therapy. London. National Institute for Health and Clinical Excellence, 2013. Available from: https://www.nice.org.uk/guidance/cg169

74 National Institute for Health and Care Excellence. Detecting, managing and monitoring haemostasis: viscoelastometric point-of-care testing (ROTEM, TEG and Sonoclot systems), 2014. Available from: https://www.nice.org.uk/guidance/dg13

75 National Institute for Health and Care Excellence. Menopause: Diagnosis and Management, 2015. Available from: https://www.nice.org.uk/guidance/ng23

76 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: http://www.nice.org.uk/media/C18/30/SVJ2PUBLICATION2008.pdf

77 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/

78 National Institute for Health and Clinical Excellence. Chronic kidney disease: early identification and management of chronic kidney disease in adults in primary and secondary care, 2014. Available from: https://www.nice.org.uk/guidance/cg182

79 National Patient Safety Agency. Rapid response report: checking pregnancy beforer surgery, 2010. Available from: www.nrls.npsa.nhs.uk/alerts

80 Organisation for Economic Co-operation and Development (OECD). Purchasing power parities (PPP). 2015. Available from: http://www.oecd.org/std/ppp [Last accessed: 7 June 2012]

81 Poso T, Winso O, Aroch R, Kesek D. Perioperative fluid guidance with transthoracic echocardiography and pulse-contour device in morbidly obese patients. Obesity Surgery. 2014; 24(12):2117-2125

82 Prentis JM, Trenell MI, Jones DJ, Lees T, Clarke M, Snowden CP. Submaximal exercise testing predicts perioperative hospitalization after aortic aneurysm repair. Journal of Vascular Surgery. 2012; 56(6):1564-1570

83 Prentis JM, Trenell MI, Vasdev N, French R, Dines G, Thorpe A et al. Impaired cardiopulmonary reserve in an elderly population is related to postoperative morbidity and length of hospital stay after radical cystectomy. BJU International. 2013; 112(2):E13-E19

84 Royal College of Paediatrics and Child Health. Pre-procedure Pregnancy Checking in Under 16s: Guidance for Clinicians. London. The Royal College of Paediatrics and Child Health, 2012. Available from:

http://www.rcpch.ac.uk/system/files/protected/page/pregnancy%20checking%20guidance%20final.pdf

85 Smetana GW, Macpherson DS. The case against routine preoperative laboratory testing. Med Clin North Am. 2003; 87(1):7-40

86 Snowden CP, Prentis JM, Anderson HL, Roberts DR, Randles D, Renton M et al. Submaximal cardiopulmonary exercise testing predicts complications and hospital length of stay in patients undergoing major elective surgery. Annals of Surgery. 2010; 251(3):535-541

87 Soong CV, Makar RR, O'Donnell ME, Badger SA, Lee B, Sharif MA. Effect of preoperative renal dysfunction on mortality and postoperative renal failure following endovascular abdominal aortic aneurysm repair. Vascular and Endovascular Surgery. 2008; 42(5):427-432

88 Torchio R, Guglielmo M, Giardino R, Ardissone F, Ciacco C, Gulotta C et al. Exercise ventilatory inefficiency and mortality in patients with chronic obstructive pulmonary disease undergoing surgery for non-small-cell lung cancer. European Journal of Cardio-Thoracic Surgery. 2010; 38(1):14-19

89 Van Klei WA, Bryson GL, Yang H, Kalkman CJ, Wells GA, Beattie WS. The value of routine preoperative electrocardiography in predicting myocardial infarction after noncardiac surgery. Annals of Surgery. 2007; 246(2):165-170

90 Wang RT, Zhang LQ, Mu YP, Li JB, Xu XS, Pang Q et al. Prognostic significance of preoperative platelet count in patients with gallbladder cancer. World Journal of Gastroenterology. 2015; 21(17):5303-5310

91 Waugh NR, Shyangdan D, Taylor-Phillips S, Suri G, Hall B. Screening for type 2 diabetes: a short report for the National Screening Committee. Health Technology Assessment. 2013; 17(35)

92 Weingarten TN, Flores AS, McKenzie JA, Nguyen LT, Robinson WB, Kinney TM et al. Obstructive sleep apnoea and perioperative complications in bariatric patients. British Journal of Anaesthesia. 2011; 106(1):131-139

93 West MA, Lythgoe D, Barben CP, Noble L, Kemp GJ, Jack S et al. Cardiopulmonary exercise variables are associated with postoperative morbidity after major colonic surgery: a prospective blinded observational study. British Journal of Anaesthesia. 2014; 112(4):665-671

94 Wijeysundera DN, Beattie WS, Karkouti K, Neuman MD, Austin PC, Laupacis A. Association of echocardiography before major elective non-cardiac surgery with postoperative survival and length of hospital stay: Population based cohort study. BMJ. 2011; 343(7816)

95 Yoshihara H, Yoneoka D. Predictors of allogeneic blood transfusion in total hip and knee arthroplasty in the United States, 2000-2009. Journal of Arthroplasty. 2014; 29(9):1736-1740

19 Acronyms and abbreviations

Phrase	Abbreviation
Abdominal Aortic Aneurysm	AAA
Absolute risk difference	ARD
Angiotensin-converting-enzyme inhibitor	ACE inhibitor
Activities of daily living assessment	ADL assessment
Apnoea-hypopnoea index	AHI
Analysis of variance	ANOVA
Acute Physiology and Chronic Health Evaluation II scoring system	APACHE II
American Society of Anesthesiologists	ASA
Anaerobic threshold	AT
American Thoracic Society-European Respiratory Society standards	ATS-ERS standards
Biphasic positive airway pressure	biPAP
Body mass index	BMI
British National Formulary	BNF
Carotid artery stenting	CAS
Carotid endarterectomy	CEA
Cost-effectiveness acceptability curve	CEAC
Congestive heart failure	CHF
Confidence interval	CI
Central nervous system	CNS
Chronic obstructive pulmonary disease	COPD
Continuous positive airway pressure	CPAP
Cardiopulmonary exercise test	CPET
C-reactive protein	CRP
Diffusing capacity of the lung for carbon monoxide	DLCO
Deep vein thrombosis	DVT
Direct oral anticoagulant	DOAC
Electrocardiography	ECG
Electroencephalography	EEG
Estimated glomerular filtration rate	eGFR
EuroQol 5-dimension	EQ-5D
Endovascular aneurysm repair	EVAR
Forced expiratory volume in 1 second	FEV1
Functional residual capacity	FRC
Forced vital capacity	FVC
Guideline Development Group	GDG
Glomerular filtration rate	GFR
Haemoglobin	Hb
Glycated haemoglobin	HbA1c
High dependency unit	HDU
Hazard ratio	HR

Phrase	Abbreviation
Health related quality of life	HRQoL
Health technology assessment	HTA
Incremental cost-effectiveness ratio	ICER
Intensive care unit	ICU
Intensive treatment unit	ITU
Lower extremity stenting	LES
Major adverse cardiovascular events	MACE
Metabolic equivalent score	MET
Minimal important difference	MID
Myocardial infarction	MI
Maximal voluntary ventilation	MVV
National Confidential Enquiry into Perioperative Deaths	NCEPOD
Non-steroidal anti-inflammatory drugs	NSAID
Non ST elevation myocardial infarction	NSTEMI
New York heart association class 3b	NYHA IIIb
Odds ratio	OR
Obstructive sleep apnoea	OSA
Highest oxygen uptake value during test	Peak VO2
Pulmonary embolism	PE
End-tidal oxygen value	$PETO_2$
Pulmonary function test	PFT
Partial pressure of oxygen	pO2
Predicted postoperative diffusing capacity of the lung for carbon monoxide	ppoDLCO
Predicted postoperative forced expiratory volume in one second.	ppoFEVI
Haemoglobin	Hb
Quality-adjusted life-years	QALYS
Corrected QT interval	QTc
Red blood cell	RBC
Renal cell carcinoma	RCC
Randomised controlled trial	RCT
Revised cardiac risk index score	RCRI
Respiratory exchange ratio	RER
Risk ratio (relative risk)	RR
Short form	SF-(36)
Serum glutamic-oxaloacetic transaminase/aspartate aminotransferase	SGOT/AST
Serum glutamic pyruvic transaminase/alanine aminotransferase	SGPT/ALT
ST elevated myocardial infarction	STEMI
Thoracoabdominal aortic aneurysm	TAAA
Thoracoabdominal aneurysm	TEVAR
Total hip arthroplasty	THA
Transient ischaemic attack	TIA
Total knee arthroplasty	TKA

Phrase	Abbreviation
Total lung capacity	TLC
Tumour, nodes and metastases classification system	TNM
Tidal volume	TV
Urea, estimated glomerular filtration rate and electrolyte tests	U&Es
Veterans activity score index	VASI
Veterans Affairs Surgery Quality Improvement Programme functional health status score	VASQIP
Vital capacity	VC
Carbon dioxide exhaled	VCO2
Ventilation efficiency	VE
Oxygen uptake	VO2
Maximal oxygen uptake	VO2 max
Vascular physiological and operative severity score for the enumeration of mortality and morbidity	VPOSSUM
White blood cell	WBC

20 Glossary

The NICE Glossary can be found at www.nice.org.uk/glossary.

Term	Definition
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.
Base case analysis	In an economic evaluation, this is the main analysis based on the most plausible estimate of each input. In contrast, see Sensitivity analysis.
Baseline	The initial set of measurements at the beginning of a study (after run-in period where applicable), with which subsequent results are compared.
Bayesian analysis	A method of statistics, where a statistic is estimated by combining established information or belief (the 'prior') with new evidence (the 'likelihood') to give a revised estimate (the 'posterior').
Before-and-after study	A study that investigates the effects of an intervention by measuring particular characteristics of a population both before and after taking the intervention, and assessing any change that occurs.
Beta coefficient	Beta coefficients are a statistical measure used to define the individual contribution of an independent variable to an outcome of interest. They are obtained following multiple regression analysis.
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 and 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.

Term	Definition
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 that 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).
Concordance	This is a recent term whose meaning has changed. It was initially applied to the consultation process in which doctor and patient agree therapeutic decisions that incorporate their respective views, but now includes patient support in medicine taking as well as prescribing communication. Concordance reflects social values but does not address medicine-taking and may not lead to improved adherence.
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

Term	Definition
	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–consequences 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.
Credible interval (CrI)	The Bayesian equivalent of a confidence interval.
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.
Deterministic analysis	In economic evaluation, this is an analysis that uses a point estimate for each input. In contrast, see Probabilistic analysis

Term	Definition
Diagnostic odds ratio	The diagnostic odds ratio is a measure of the effectiveness of a diagnostic test. It is defined as the ratio of the odds of the test being positive if the subject has a disease relative to the odds of the test being positive if the subject does not have the disease.
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.
Disutility	The loss of quality of life associated with having a disease or condition. See Utility
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–consequences 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

Term	Definition
	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.
Hazard ratio	For some outcomes, the time elapsed before an event occurs is important. This type of evidence is known as time-to-event data or survival data, the outcome of which is expressed as a *hazard ratio*: How many times more (or less) likely a participant is to suffer the event at a particular point in time if they receive the experimental intervention rather than control.
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

Term	Definition
	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.
Intraoperative	The period of time during a surgical procedure.
Kappa statistic	A statistical measure of inter-rater agreement that takes into account the agreement occurring by chance.
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.
Logistic regression or Logit model	In statistics, logistic regression is a type of analysis used for predicting the outcome of a binary dependent variable based on one or more predictor variables. It can be used to estimate the log of the odds (known as the 'logit').
Loss to follow-up	A patient, or the proportion of patients, actively participating in a clinical trial at the beginning, but whom the researchers were unable to trace or contact by the point of follow-up in the trial
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.
Multivariate model	A statistical model for analysis of the relationship between 2 or more predictor (independent) variables and the outcome (dependent) variable.
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: <Insert formula>
Net monetary benefit (NMB)	The value in monetary terms of an intervention net of its cost. The NMB can be calculated for a given cost-effectiveness threshold. If the threshold is £20,000 per QALY gained then the NMB for an intervention is calculated as: (£20,000 × mean QALYs) − mean cost.

Term	Definition
	The most preferable option (that is, the most clinically effective option to have an ICER below the threshold selected) will be the treatment with the highest NMB.
Number needed to treat (NNT)	The average number of patients who need to be treated to get a positive outcome. For example, if the NNT is 4, then 4 patients would have to be treated to ensure 1 of them gets better. The closer the NNT is to 1, the better the treatment. For example, if you give a stroke prevention drug to 20 people before 1 stroke is prevented, the number needed to treat is 20. See also number needed to harm, absolute risk reduction.
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 healthcare 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

Term	Definition
	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.
Perioperative	The period from admission through surgery until discharge, encompassing the preoperative and postoperative periods.
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.
Posterior distribution	In Bayesian statistics this is the probability distribution for a statistic based after combining established information or belief (the prior) with new evidence (the likelihood).
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: <Insert formula>
Postoperative	Pertaining to the period after patients leave the operating theatre, following surgery.
Post-test probability	In diagnostic tests: The proportion of patients with that particular test result who have the target disorder (post-test odds/[1 plus post-test odds]). <Adjust formula>
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.
Pre-test probability	In diagnostic tests: The proportion of people with the target disorder in the population at risk at a specific time point or time interval. Prevalence may depend on how a disorder is diagnosed.
Prevalence	See Pre-test probability.
Prior distribution	In Bayesian statistics this is the probability distribution for a statistic based on previous evidence or belief.
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.
Probabilistic analysis	In economic evaluation, this is an analysis that uses a probability distribution for each input. In contrast, see Deterministic analysis.
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

Term	Definition
	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.
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.
Reporting bias	See 'Publication bias'.
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.

Term	Definition
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

Term	Definition
	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.
State transition model	See Markov model
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.
Transition probability	In a state transition model (Markov model), this is the probability of moving from one health state to another over a specific period of time.
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).

Guideline-specific terms used

Term	Definition
Arterial blood gases	Analysis of an arterial blood sample that determines the pH, partial pressure of oxygen and carbon dioxide and bicarbonate level. Some analysers also report concentrations of lactate, glucose, electrolytes and haemoglobin.
Cardiac surgery	Surgery on the heart and/or surrounding great vessels.
Cardiopulmonary exercise test (CPET)	A non-invasive simultaneous measurement of the cardiovascular and respiratory system during exercise, typically performed using a cycle ergometer.
Chest X-ray	Formally known as a chest radiograph. Ionising radiation is used to generate images of the chest structures including the lungs, pleura, heart, major vasculature, mediastinum, chest wall and diaphragm.
Comorbidity	Having 2 or more diagnosable conditions at the same time.
Elective surgery	Scheduled procedure: that is, not an urgent or emergency procedure.
Full blood count	Analysis of a venous blood sample that measures red blood cell, white blood cell, platelet and haemoglobin concentrations.
Haemostasis tests	Analysis of a venous blood sample to evaluate blood clotting parameters. Reported values include the prothrombin time (PT), activated partial thromboplastin time (APTT) and international normalised ratio (INR; derived from the patient's PT and normative data). This test is also commonly known as a clotting screen.
HbA1c (glycated haemoglobin)	Analysis of a venous blood sample to determine the amount of glucose bound to red blood cells, a measure of average plasma blood glucose

Term	Definition
	concentration.
Lung function tests	Non-invasive tests of lung function. The simplest of these is spirometry, which measures values including the peak expiratory flow rate (PEFR), forced vital capacity (FVC) and forced expiratory volume (FEV). More complex tests can determine lung volumes and the diffusing capacity of the lungs.
Perioperative	The period from admission through surgery until discharge, encompassing the preoperative and postoperative periods.
Polysomnography	A test that electronically records specific body functions, including cardiovascular, respiratory, neurological and musculoskeletal parameters, during sleep.
Postoperative	The period after patients leave the operating theatre, following surgery.
Pregnancy test	Biochemical testing for pregnancy. In most cases this would be a urine test.
Preoperative	The period from the time of preparation for surgery/administration of premedication to the time of first anaesthetic intervention.
Renal (kidney) function tests	Analysis of a venous blood sample to determine creatinine, electrolyte and sometimes urea concentrations. A calculated estimated glomerular concentration rate (eGFR) is also frequently reported.
Resting electrocardiography (ECG)	A non-invasive test that records the electrical activity of the heart.
Resting echocardiography	Echocardiography uses ultrasonography to image the heart. Transthoracic echocardiography (TTE) is non-invasive and images the heart using an ultrasound device placed on the chest, whereas transoesophageal echocardiography (TOE) is more invasive and uses an ultrasound probe inserted into the oesophagus to image the heart from behind.
Sickle cell anaemia test	Analysis of a venous blood sample to test for sickle cell anaemia or trait usually takes place in 2 stages. First a test is used to detect haemoglobin S, which is present in both sickle cell anaemia and trait (for example the Sickledex test). If this test is positive, haemoglobin electrophoresis is performed to determine which condition is present, sickle cell anaemia or trait.
Surgery grades	An operation represents a physiological stress. The magnitude of the physiological stress increases with the 'invasiveness' of the procedure. There is no widely accepted and validated system for classifying the stressfulness of operative procedures, so this guideline adopted a simple graded scale, illustrated with examples. **Minor surgery** • excising skin lesion • draining breast abscess **Intermediate surgery** • primary repair of inguinal hernia • excising varicose veins in the leg • tonsillectomy or adenotonsillectomy • knee arthroscopy **Major or complex surgery** • total abdominal hysterectomy • endoscopic resection of prostate • lumbar discectomy • thyroidectomy

Term	Definition
	• total joint replacement • lung operations • colonic resection • radical neck dissection
The ASA Physical Status Classification System	ASA stands for American Society of Anesthesiologists. The ASA Physical Status Classification system is a simple scale describing fitness to undergo an anaesthetic. The American Society of Anesthesiologists clearly states that it does not endorse any elaboration of these definitions. However, anaesthetists in the UK often qualify (or interpret) ASA grades as relating to functional capacity: that is, comorbidity that does (ASA3) or that does not (ASA2) limit a patient's activity. These definitions appear in each annual edition of the ASA Relative Value Guide.
Thoracic surgery	Surgery on organs located in the thorax or chest cavity.
Urine analysis test	Urine analysis tests are manufactured to test for different conditions separately and together. Urine analysis tests are for pH, protein, glucose, ketones and blood/haemoglobin.
Vascular surgery	Surgery on blood vessels excluding those intimately associated with the heart.