

Acute Heart Failure

Acute heart failure: diagnosing and managing acute heart failure in adults

Clinical guideline 187

Methods, evidence and recommendations

October 2014

Final version

*Commissioned by the National Institute for
Health and Care Excellence*

Disclaimer

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

Copyright

National Clinical Guideline Centre, 2014.

Funding

National Institute for Health and Care Excellence

Contents

Guideline Development Group members	8
NCGC technical team members	8
Co-optees	9
Acknowledgements	10
1 Introduction	11
2 Development of the guideline	13
2.1 What is a NICE clinical guideline?	13
2.2 Remit	13
2.3 Who developed this guideline?	14
3 Methods.....	18
3.1 Developing the review questions and outcomes.....	18
3.2 Searching for evidence.....	23
3.2.1 Clinical literature search.....	23
3.2.2 Health economic literature search.....	23
3.3 Evidence of effectiveness.....	24
3.3.1 Inclusion and exclusion criteria	25
3.3.2 Methods of combining clinical studies.....	26
3.3.3 Type of studies	28
3.3.4 Appraising the quality of evidence by outcomes	29
3.3.5 Grading the quality of clinical evidence	30
3.3.6 Risk of bias.....	31
3.3.7 Inconsistency	33
3.3.8 Indirectness	33
3.3.9 Imprecision	33
3.3.10 Publication bias	34
3.3.11 Assessing clinical importance (benefit, harm or no difference)	35
3.3.12 Evidence statements	35
3.4 Evidence of cost effectiveness	35
3.4.1 Literature review	35
3.4.2 Undertaking new health economic analysis	37
3.4.3 Cost-effectiveness criteria.....	38
3.4.4 In the absence of economic evidence	38
3.5 Developing recommendations.....	38
3.5.1 Research recommendations	39
3.5.2 Validation process	39
3.5.3 Updating the guideline	39

3.5.4	Disclaimer	40
3.5.5	Funding.....	40
4	Guideline summary.....	41
4.1	Algorithm	41
4.2	Key priorities for implementation.....	42
4.3	Full list of recommendations	43
4.4	Key research recommendations	46
5	Organisation of care	47
5.1	Specialist management units	47
5.1.1	Clinical evidence.....	47
5.1.2	Economic evidence.....	61
5.1.3	Evidence statements	65
5.1.4	Recommendations and link to evidence.....	68
6	Diagnosis, assessment and monitoring	71
6.1	Natriuretic peptides	71
6.1.1	Clinical evidence.....	72
6.1.2	Economic evidence.....	83
6.1.3	Evidence statements	86
6.1.4	Recommendations and link to evidence.....	88
6.2	Echocardiography	92
6.2.1	Clinical evidence.....	93
6.2.2	Economic evidence.....	93
6.2.3	Evidence statements	93
6.2.4	Recommendations and link to evidence.....	93
6.3	Invasive monitoring.....	94
6.3.1	Clinical evidence.....	95
6.3.2	Economic evidence.....	101
6.3.3	Evidence statements	101
6.3.4	Recommendations and link to evidence.....	102
7	Initial pharmacological treatment	104
7.1	General considerations	104
7.2	Opiates	104
7.2.1	Clinical evidence.....	105
7.2.2	Economic evidence.....	112
7.2.3	Evidence statements	112
7.2.4	Recommendations and link to evidence.....	113
7.3	Diuretic administration strategies	114
7.3.1	Clinical evidence.....	115

7.3.2	Economic evidence.....	125
7.3.3	Evidence statements	125
7.3.4	Recommendations and link to evidence	128
7.4	Vasodilators.....	130
7.4.1	Clinical evidence.....	131
7.4.2	Economic evidence.....	138
7.4.3	Evidence statements	138
7.4.4	Recommendations and link to evidence	140
7.5	Inotropes and vasopressors	141
7.5.1	Clinical evidence.....	142
7.5.2	Economic evidence.....	155
7.5.3	Evidence statements	155
7.5.4	Recommendations and link to evidence	158
8	Initial non-pharmacological treatment	161
8.1	Ventilatory support	161
8.1.1	Clinical evidence.....	161
8.1.2	Economic evidence.....	175
8.1.3	Evidence statements	177
8.1.4	Recommendations and link to evidence	179
8.2	Invasive ventilation	180
8.2.1	Clinical evidence.....	181
8.2.2	Economic evidence.....	185
8.2.3	Evidence statements	185
8.2.4	Recommendations and link to evidence	186
8.3	Ultrafiltration	187
8.3.1	Clinical evidence.....	187
8.3.2	Economic evidence.....	203
8.3.3	Evidence statements	205
8.3.4	Recommendations and link to evidence	208
9	Treatment after stabilisation.....	210
9.1	Timing of beta blocker therapy	210
9.1.1	Clinical evidence.....	210
9.1.2	Economic evidence.....	220
9.1.3	Evidence statements	220
9.1.4	Recommendations and link to evidence	222
9.2	Timing of ACE inhibitor and aldosterone antagonist therapy	224
9.2.1	Clinical evidence.....	225
9.2.2	Economic evidence.....	232

9.2.3	Evidence statements	233
9.2.4	Recommendations and link to evidence	234
10	Valvular surgery and percutaneous intervention	235
10.1	Aortic stenosis	235
10.1.1	Clinical evidence	235
10.1.2	Economic evidence	250
10.1.3	Evidence statements	251
10.1.4	Recommendations and links to evidence	255
10.2	Mitral regurgitation	257
10.2.1	Clinical evidence	258
10.2.2	Economic evidence	268
10.2.3	Evidence statements	269
10.2.4	Recommendations and links to evidence	270
11	Mechanical assist devices	272
11.1.1	Clinical evidence	272
11.1.2	Economic evidence	285
11.1.3	Evidence statements	287
11.1.4	Recommendations and link to evidence	289
12	Reference list	292
13	Glossary	311

Guideline Development Group members

Name	Role
Professor Jonathan Mant (GDG Chair)	Professor of Primary Care Research and Honorary Consultant, University of Cambridge
Dr Abdallah Al-Mohammad	Consultant Cardiologist and Honorary Senior Clinical Lecturer, Sheffield Teaching Hospitals NHS Trust
Mr Peter Bolton	Patient member
Ms Jane Butler	Heart Failure Specialist Consultant Nurse, Bart's Health NHS Trust
Professor Martin Cowie	Professor of Cardiology and Honorary Consultant Cardiologist, Imperial College London
Dr Suzanna Hardman	Consultant Cardiologist Whittington Health and Honorary Senior Lecturer, University College London.
Dr Nicholas Ioannou	Consultant in Intensivist and Anaesthetist, Guy's and St Thomas' NHS Foundation Trust
Mr Christopher Jones ^a	Patient member
Dr Jason Kendall	Consultant in Emergency Medicine, North Bristol NHS Trust
Ms Jayne Masters	Lead Heart Failure Specialist Nurse, University Hospitals Southampton NHS Foundation Trust
Professor John McMurray	Professor of Medical Cardiology and Honorary Consultant Cardiologist, University of Glasgow
Dr Tanzeem Raza	Consultant Physician in Acute Medicine, Royal Bournemouth Hospital

NCGC technical team members

Name	Role
Liz Avital	Associate Director (from January - July 2013)
Amelia Ch'ng	Project Manager (until September 2013)
Saskia Cheyne	Project Manager (from July 2013)
Jill Cobb	Information Scientist (from January 2013)
Katharina Dworzynski	Senior Research Fellow
Lina Gulhane	Joint Head of Information Science (until January 2013)
Dr Andrew Ludman (Specialist trainee adviser)	Specialist Trainee in Cardiology, Royal Brompton and Harefield NHS Foundation Trust
Su Park	Research Fellow (from August 2013)
Elisabetta Fenu	Senior Health Economist (until March 2013)
Edward Griffin	Health Economist (from March 2013)
Gill Ritchie	Associate Director (until December 2012 and from August 2013)
Juan Carlos Rejon	Health Economist (until January 2013)
Emmert Roberts	Research Fellow (until July 2013)
Giulia Zuodar	Document Editor/Process Assistant (from June 2014)

^a Oct 2012-RIP February 2014

Co-optees

Name	Role
Professor Paul Collinson (expert adviser)	Consultant Chemical Pathologist and Professor of Cardiovascular Biomarkers, St George's Healthcare NHS Trust
Dr Mark Devonald (expert adviser)	Consultant Nephrologist, Nottingham University Hospitals
Professor Ahmet Fuat (expert adviser)	Professor of Primary Care Cardiology, Durham University
Dr Marie McKenniff (expert adviser)	Group General Manager, Sheffield Teaching Hospitals NHS Trust
Ms Alison Warren (expert adviser)	Lead Pharmacist in Cardiology, Brighton and Sussex University Hospitals NHS Trust
Professor Stephen Westaby (expert adviser)	Consultant Cardiac Surgeon and Professor of Biomedical Science, John Radcliffe Hospital, Oxford

Acknowledgements

In memory of Christopher Jones, patient member of the GDG who ensured that the patient voice was heard during the development of this guideline.

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

- Alex Allen, Information Scientist, NCGC
- Mayret Castillo, Research Fellow, NCGC
- Tamara Diaz, Project Manager, NCGC
- Lina Gulhane, Joint Head of Information Science, NCGC
- Sarah King
- Fatema Limbada, Project Coordinator, NCGC
- Paul Miller, Senior Information Scientist, NCGC
- Polly Mitchell, National Heart Failure Audit Project Manager, National Institute for Cardiovascular Outcomes Research
- Jacoby Patterson
- Ben Pordes, Project Manager, NCGC
- Elizabeth Rouncefield, Information Scientist, NCGC
- Aminat Shote, Statistician, National Institute for Cardiovascular Outcomes Research
- Maggie Westby, Clinical Effectiveness Lead, NCGC
- David Wonderling, Head of Health Economics

1 Introduction

The need for this guideline was identified as the NICE guidelines on chronic heart failure were being updated. We recognised at this time that there were important aspects of the diagnosis and management of acute heart failure that were not being addressed by the chronic heart failure guideline, which focussed on long term management rather than the immediate care of someone who is acutely unwell as a result of heart failure. The aim of this guideline is to provide guidance to the NHS on the diagnosis and management of acute heart failure.

Heart failure is a condition in which the heart does not pump enough blood to meet all the needs of the body. It is caused by heart muscle damage or dysfunction, valve problems, heart rhythm disturbances and other rarer causes. Acute heart failure can present as new-onset heart failure in people without known cardiac dysfunction, or as acute decompensation of chronic heart failure.

Acute heart failure is a common cause of admission to hospital (over 67,000 admissions in England and Wales per year) and is the leading cause of hospital admission in people 65 years or older in the UK.

This guideline includes important aspects of the diagnosis and management of acute heart failure that are not addressed by the NICE guideline on chronic heart failure (NICE clinical guideline 108). The guideline on chronic heart failure focused on long-term management rather than the immediate care of someone who is acutely unwell as a result of heart failure.

This guideline covers the care of adults (aged 18 years or older) who have a diagnosis of acute heart failure, have possible acute heart failure, or are being investigated for acute heart failure. It includes the following key clinical areas:

- the role of early natriuretic peptide testing and echocardiography
- the role of specialist management units
- the use of ventilatory support, pharmacological therapy and ultrafiltration
- treatment after stabilisation, including selected surgical interventions and start of the pharmacological therapies that are used in the management of chronic heart failure.

Patient centred care

Patients and healthcare professionals have rights and responsibilities as set out in the NHS Constitution for England – all NICE guidance is written to reflect these. Treatment and care should take into account individual needs and preferences. Patients should have the opportunity to make

informed decisions about their care and treatment, in partnership with their healthcare professionals. If the patient is under 16, their family or carers should also be given information and support to help the child or young person to make decisions about their treatment. Healthcare professionals should follow the Department of Health's advice on consent. If someone does not have capacity to make decisions, healthcare professionals should follow the code of practice that accompanies the Mental Capacity Act and the supplementary code of practice on deprivation of liberty safeguards. In Wales, healthcare professionals should follow advice on consent from the Welsh Government.

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

2 Development of the guideline

2.1 What is a NICE clinical guideline?

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

NICE clinical guidelines can:

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

2.2 Remit

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

The remit for this guideline is: 'To prepare a guideline on the diagnosis and management of acute heart failure'.

2.3 Who developed this guideline?

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

A multidisciplinary Guideline Development Group (GDG) comprising health professionals and researchers as well as lay members developed this guideline (see the list of Guideline Development Group members and the acknowledgements).

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

The group met every 5-6 weeks during the development of the guideline. At the start of the guideline development process all 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.

(a) What this guideline covers

This guideline covers the following populations:

- a) Adults (aged 18 years or older) who have a diagnosis of acute heart failure, or have possible acute heart failure, or are being investigated for acute heart failure.
- b) Specific consideration will be given to subgroups with pulmonary oedema, cardiogenic shock, acute right-sided heart failure or acute decompensated heart failure.

The following clinical issues are covered:

Diagnosis, assessment and monitoring

- a) In addition to the standard investigations (such as ECG, chest X-ray and blood tests), the added benefit of using natriuretic peptides or echocardiography.
- b) Indications for, and types of, invasive (arterial lines, central venous pressure lines and pulmonary artery catheters) monitoring when non-invasive monitoring alone is no longer appropriate.

Management of acute heart failure

- c) Specialist management units.

Initial treatment

Oxygen and ventilatory support

- d) The use of supplementary oxygen, ventilatory support (CPAP), non-invasive (NIPPV) or invasive ventilation to maximise oxygen delivery to the tissues to prevent multiple organ failure.

Pharmacological therapy

- e) Management with drug therapy, including diuretics, opiates, vasodilators, inotropic agents and vasopressors.
- f) Discontinuing beta-blockers.

Ultrafiltration

- g) Timing (initiation and duration) of ultrafiltration.

Mechanical cardiac support

- h) Mechanical circulatory assistance with intra-aortic balloon counterpulsation or ventricular assist devices.

Treatment after stabilisation

Pharmacological therapy

- i) Starting or re-instating treatment for new-onset acute heart failure with angiotensin-converting enzyme (ACE) inhibitors, beta-blockers and/or aldosterone antagonists.

Surgical or percutaneous treatment

- j) The use of coronary revascularisation and valvular surgery when acute heart failure is a severe complication of other cardiac disorders.

Organisation of care

- k) Transition from hospital to primary care after the acute phase.

Note that guideline recommendations will normally fall within licensed indications; exceptionally, and only if clearly supported by evidence, use outside a licensed indication may be recommended. The guideline will assume that prescribers will use a drug's summary of product characteristics to inform decisions made with individual patients. For further details please refer to the scope in Appendix A and review questions in Section 0.

(b) What this guideline does not cover

This guideline does not cover the following populations:

- a) Children and young people under 18 years.

The following clinical issues are not covered:

- a) The long-term management of underlying diseases (such as congenital heart disease) and comorbidities of acute heart failure.
- b) The management of perioperative acute heart failure.
- c) The long-term management of acute heart failure in pregnant women.

(c) Relationships between the guideline and other NICE guidance

Related NICE Technology appraisals:

- Bivalirudin for the treatment of ST-segment elevation myocardial infarction (STEMI). NICE technology appraisal guidance 230 (2011).
- Prasugrel for the treatment of acute coronary syndromes with percutaneous coronary intervention. NICE technology appraisal guidance 182 (2009).
- Varenicline for smoking cessation. NICE technology appraisal guidance 123 (2007).
- Cardiac resynchronisation therapy for the treatment of heart failure. NICE technology appraisal guidance 120 (2007).
- Coronary imaging: myocardial perfusion scintigraphy for the diagnosis and management of angina and myocardial infarction. NICE technology appraisal guidance 73 (2003).
- Ivabradine for the treatment of chronic heart failure. NICE technology appraisal guidance 267 (2012)

Related NICE Interventional procedures guidance:

- Short-term circulatory support with left ventricular assist devices as a bridge to cardiac transplantation or recovery. NICE interventional procedure guidance 177 (2006).

Related NICE Clinical guidelines:

- Myocardial infarction with ST-segment-elevation. NICE clinical guidance 167 (2013).
- MI: secondary prevention (update). NICE clinical guideline 172 (2013).
- Patient experience in adult NHS services. NICE clinical guidance 138 (2012).
- Hypertension. NICE clinical guideline 127 (2011).
- Stable angina. NICE clinical guideline 126 (2011).
- Chronic heart failure. NICE clinical guideline 108 (2010).
- Chest pain of recent onset. NICE clinical guideline 95 (2010).
- Unstable angina and NSTEMI. NICE clinical guideline 94 (2010).
- Type 2 diabetes – newer agents. NICE clinical guideline 87 (2009).
- Chronic kidney disease. NICE clinical guideline 73 (2008).
- Lipid modification. NICE clinical guideline 67 (2008).
- Atrial fibrillation. NICE clinical guideline 36 (2006).

Related NICE Public health guidance:

- Smoking cessation services. NICE public health guidance 10 (2008).
- Brief interventions and referral for smoking cessation. NICE public health guidance 1 (2006).

Related NICE guidance currently in development:

Lipid modification (update). NICE clinical guideline. Publication July 2014

Atrial fibrillation (update). NICE clinical guideline. Publication June 2014

Implantable cardioverter defibrillators for the treatment of arrhythmias and cardiac resynchronisation therapy for the treatment of heart failure (review of TA95 and TA120) Publication June 2014.

3 Methods

This chapter sets out in detail the methods used to review the evidence and to generate the recommendations that are presented in subsequent chapters. This guidance was developed in accordance with the methods outlined in the NICE guidelines manual 2012¹⁵⁵.

3.1 Developing the review questions and outcomes

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

This use of a framework guided the literature searching process, critical appraisal and synthesis of evidence, and facilitated the development of recommendations by the Guideline Development Group (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 25 review questions were originally identified. Some of these were later combined into one review when the protocols were agreed, resulting in an overall list of 18 review topics (see Table 1 below).

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

Table 1: Review questions

Chapter	Type of review	Review questions	Outcomes / risk factors / diagnostic measures
5	Intervention	In adults with suspected acute heart failure does early echocardiography compared to later echocardiography in addition to standard investigations (using ECG, chest x-ray and blood tests) improve outcome?	Critical outcomes: <ul style="list-style-type: none"> • Mortality • Serious adverse events • Quality of life Important outcomes: <ul style="list-style-type: none"> • Length of hospital stay • Readmission rates
5	Diagnostic	In adults with suspected acute heart failure does the addition of natriuretic peptides to the standard initial investigations (using ECG, chest x-ray and blood tests) improve speed and accuracy of diagnosis?	Diagnostic measures: <ul style="list-style-type: none"> • Sensitivity was identified as the most critical statistical measure as natriuretic peptides are used as 'rule out' tests • Specificity, Area under the Curve were also important measures
5	Intervention	Is the addition of invasive monitoring more clinically/cost-effective over and above non-invasive monitoring to improve outcome	Critical outcomes: <ul style="list-style-type: none"> • Mortality • Serious adverse events • Quality of life

Chapter	Type of review	Review questions	Outcomes / risk factors / diagnostic measures
			<p>Important outcomes:</p> <ul style="list-style-type: none"> • Length of hospital stay • Measures of renal function • Number of patients proceeding to invasive ventilation
6	Intervention	In patients with acute heart failure which diuretic administration strategy is the most clinically/cost-effective to improve outcome?	<p>Critical outcomes:</p> <ul style="list-style-type: none"> • Mortality • Urine Output • Weight Loss • Quality of life • Serum creatinine level (or other measure of kidney function such as eGFR) • Serious adverse events (particularly renal adverse events and ototoxicity) <p>Important outcomes:</p> <ul style="list-style-type: none"> • Dyspnoea • Length of hospital stay and re-admission rates
6	Intervention	In patients with acute heart failure are opiates as an adjunct to other first line therapies safe and clinically / cost effective compared with placebo and to other treatments alone?	<p>Critical outcomes:</p> <ul style="list-style-type: none"> • Mortality • Major cardiovascular events • Quality of life (as well as reported anxiety and pain) <p>Important outcomes:</p> <ul style="list-style-type: none"> • Length of hospital stay and re-admission rates • Number of patients proceeding to invasive ventilation • Measures of dyspnoea (breathing rate or breathlessness scales) • Adverse events (particularly respiratory arrest and nausea)
6	Intervention	In patients with acute heart failure are vasodilators more clinically or cost effective than placebo to improve clinical outcomes?	<p>Critical outcomes:</p> <p>Mortality</p> <ul style="list-style-type: none"> • Major cardiovascular events • Quality of life <p>Important outcomes:</p> <ul style="list-style-type: none"> • Length of hospital stay

Chapter	Type of review	Review questions	Outcomes / risk factors / diagnostic measures
			<p>and re-admission rates</p> <ul style="list-style-type: none"> • Adverse events (particularly: hypotension, headache) • Discontinuation of therapy due to adverse events
6	Intervention	In patients with acute heart failure are inotropes or vasopressors safe and clinically / cost effective compared to medical care or each other to improve outcome?	<p>Critical outcomes:</p> <ul style="list-style-type: none"> • Mortality • Major cardiovascular events • Quality of life • Renal outcomes (dopamine) <p>Important outcomes:</p> <ul style="list-style-type: none"> • Length of hospital stay and re-admission rates • Serious adverse events
7	Intervention	In people with confirmed acute heart failure and cardiogenic pulmonary oedema is non-invasive positive pressure ventilation (CPAP and/or bilevel NPPV) more clinically/cost effective than standard medical care alone to improve outcome?	<p>Critical outcomes:</p> <ul style="list-style-type: none"> • Mortality • Myocardial infarction • Intubation rate • Quality of life <p>Important outcomes:</p> <ul style="list-style-type: none"> • Length of hospital stay
7	Prognostic	What are the predictors of outcome in invasively ventilated acute heart failure patients?	<p>Prognostic Risk factors:</p> <ul style="list-style-type: none"> • Age • Aetiology of heart failure • BNP • Blood pressure • Killip Class • LV ejection fraction • Hyponatraemia • Renal disease • Body mass index • Inotropic / vasopressor support • Urinary output • Infection (particularly ventilator associated pneumonia) • APACHE score • Organ failure score

Chapter	Type of review	Review questions	Outcomes / risk factors / diagnostic measures
7	Intervention	In patients with acute heart failure is ultrafiltration more clinically/cost-effective than diuretic therapy alone or in addition to diuretic therapy to improve outcome?	<p>Critical outcomes:</p> <ul style="list-style-type: none"> • Mortality • Urine output • Weight loss • Quality of life • Serum creatinine level (or other measure of kidney function such as eGFR) • Serious adverse events <p>Important outcomes:</p> <ul style="list-style-type: none"> • Dyspnoea • Length of hospital stay and re-admission rates
8	Intervention	In people with acute heart failure already on beta-blocker therapy should beta-blockers be reduced or discontinued, and if so should they be reinstated in hospital after stabilisation?	<p>Critical outcomes:</p> <ul style="list-style-type: none"> • Mortality • Major cardiovascular events • Quality of life <p>Important outcomes:</p> <ul style="list-style-type: none"> • Length of hospital stay and re-admission rates • Beta-blocker prescriptions at follow-up • Serious adverse events
8	Intervention	For people with confirmed acute heart failure not already on beta-blocker therapy should beta-blocker treatment commence in hospital after stabilisation or following discharge?	<p>Critical outcomes:</p> <ul style="list-style-type: none"> • Mortality • Major cardiovascular events • Quality of life <p>Important outcomes:</p> <ul style="list-style-type: none"> • Length of hospital stay and re-admission rates
8	Intervention	For people with confirmed acute heart failure not already on angiotensin converting enzyme (ACE)-inhibitor therapy should ACEi therapy commence in hospital or following discharge?	<p>Critical outcomes:</p> <ul style="list-style-type: none"> • Mortality • Major cardiovascular events • Quality of life <p>Important outcomes:</p> <ul style="list-style-type: none"> • Length of hospital stay and re-admission rates • Change in renal function • Serious adverse events (hyperkalaemia, cough, symptomatic hypotension)

Chapter	Type of review	Review questions	Outcomes / risk factors / diagnostic measures
8	Intervention	For people with confirmed acute heart failure not already on aldosterone antagonists should aldosterone antagonist therapy commence in hospital or following discharge?	<p>Critical outcomes:</p> <ul style="list-style-type: none"> • Mortality • Major cardiovascular events • Quality of life <p>Important outcomes:</p> <ul style="list-style-type: none"> • Length of hospital stay and re-admission rates
9	Intervention	For people with heart failure with mitral regurgitation are surgical valvular or percutaneous interventions more clinically or cost effective compared to medical care or each other?	<p>Critical outcomes:</p> <ul style="list-style-type: none"> • Mortality • Major cardiovascular events • Quality of life <p>Important outcomes:</p> <ul style="list-style-type: none"> • Length of hospital stay and re-admission rates • Adverse events (perioperative vascular)
9	Intervention	For people with heart failure secondary to aortic stenosis are surgical valvular or percutaneous interventions more clinically or cost effective compared to medical care or each other?	<p>Critical outcomes:</p> <ul style="list-style-type: none"> • Mortality • Major cardiovascular events • Quality of life <p>Important outcomes:</p> <ul style="list-style-type: none"> • Length of hospital stay and re-admission rates <p>Adverse events (perioperative vascular)</p>
10	Intervention	For people with acute heart failure is intra-aortic balloon counterpulsation more clinically / cost effective compared to left ventricular assist devices, medical therapy alone or with each other?	<p>Critical outcomes:</p> <ul style="list-style-type: none"> • Mortality • Major cardiovascular events • Quality of life <p>Important outcomes:</p> <ul style="list-style-type: none"> • Admission to critical care • Length of hospital stay and re-admission rates • Serious adverse events
11	Intervention	For people with suspected or confirmed acute heart failure is a specialist management unit more clinically / cost effective than general medical hospital care?	<p>Critical outcomes:</p> <ul style="list-style-type: none"> • Mortality • Serious adverse events • Quality of life / patient satisfaction <p>Important outcomes:</p> <ul style="list-style-type: none"> • Length of hospital stay • Readmission rates

The transition from hospital to primary care after the acute phase was another topic identified in the scope of the guideline. This was already included as part of the NICE CG108 Chronic Heart Failure guideline¹⁵² (<http://www.nice.org.uk/CG108>) hence it was not separately reviewed here. The guideline group agreed to cross-refer to the relevant recommendations.

3.2 Searching for evidence

3.2.1 Clinical literature search

Systematic literature searches were undertaken in accordance with the Guidelines Manual 2012¹⁵⁵ to identify evidence within published literature to answer the review questions. Clinical 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 on three core databases: Medline, Embase, and The Cochrane Library. An additional subject specific database (HMIC Health Management Information Consortium) was used for the question on specialist management units. All searches were updated on 28th January 2014. No papers published added to above databases after this date were considered.

Search strategies were checked by looking at reference lists of relevant key papers, checking search strategies in other systematic reviews and asking the GDG for known studies. The questions, the study type filters applied, the databases searched and the years covered can be found in Appendix F.

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

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

- Guidelines International Network database (www.g-i-n.net)
- National Guideline Clearing House (www.guideline.gov/)
- National Institute for Health and Care Excellence (NICE) (www.nice.org.uk)
- National Institutes of Health Consensus Development Program (consensus.nih.gov/)
- National Library for Health (www.library.nhs.uk/).
- Trip Database (www.tripdatabase.com/)

3.2.2 Health economic literature search

Systematic literature searches were also undertaken to identify health economic evidence within published literature relevant to the review questions. The evidence was identified by conducting a broad search relating to acute heart failure 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 2010, 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 valvular surgery on Medline, EMBASE, NHS EED, HTA and HEED 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 F. All searches were updated on 28th January 2014. No papers published after this date were considered.

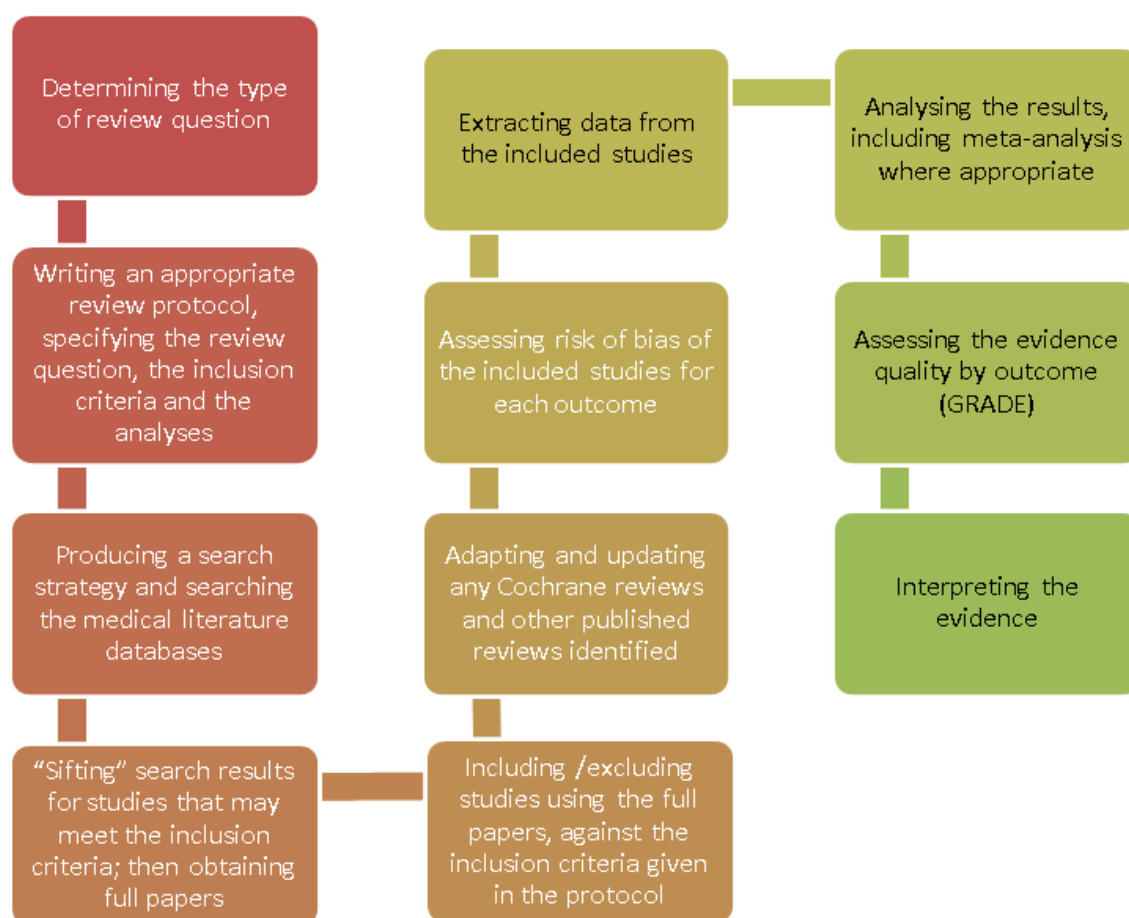
3.3 Evidence of effectiveness

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

- Potentially relevant studies were identified for each review question from the relevant search results by reviewing titles and abstracts. Full papers were then obtained.
- Full papers were reviewed against pre-specified inclusion and exclusion criteria to identify studies that addressed the review question in the appropriate population (review protocols are included in Appendix C).
- Relevant studies were critically appraised using the appropriate checklist as specified in The guidelines manual.¹⁵⁵ For diagnostic questions, the QUADAS-2 checklist was followed (<http://www.bris.ac.uk/quadas/quadas-2/>)
- Key information was extracted on the study's methods, PICO factors and results. These were presented in summary tables (in each review chapter) and evidence tables (in Appendix G).
- Summaries of evidence were generated by outcome (included in the relevant review protocols) and were presented in GDG meetings:
 - o Randomised studies: data were meta-analysed where appropriate and reported in GRADE profiles (for intervention reviews).
 - o Observational studies: data were presented as a range of values in GRADE profiles.
 - o Prognostic studies: data were presented as a range of values, usually in terms of the relative effect as reported by the authors.
 - o Diagnostic studies were presented as measures of diagnostic test accuracy (sensitivity, specificity, positive and negative predictive value). Coupled values of sensitivity and specificity were summarised in Receiver Operating Curves (ROC) to allow visual comparison between different index tests (plotting data at different thresholds) and to investigate heterogeneity more effectively (given data were reported at the same thresholds). Diagnostic meta-analyses were carried out whenever data from at least 5 studies were available. See chapter 1 and Appendix J for details.

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

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



3.3.1 Inclusion and exclusion criteria

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

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

The GDG agreed that randomised controlled trials were the most appropriate study design for intervention reviews. However, there were exceptions to this where observational studies were accepted as evidence since randomisation would be unethical or when GDG members were aware that little or no evidence was available for a particular topic. In most of these cases observational studies were only included if they were of sufficient size ($n \geq 2000$) and made multivariable adjustments for confounding baseline characteristics between treatment and control groups. This was the case for the following reviews:

- Beta-blocker reduction or discontinuation (Chapter 9.1)
- Commencing beta-blocker, angiotensin converting enzyme inhibitor or aldosterone antagonist therapy (Chapter 9.1)

Smaller observational studies ($n < 2000$) were accepted as evidence in the following reviews:

- Specialist management units (restricted to those with multivariable adjustments) (Chapter 5.1)
- Opiates (Chapter 7.2)

In the review on specialist acute heart failure management in chapter 5.1, the GDG decided to restrict evidence from the year 1999 onwards, because specialist services were considered to have gone through substantial changes and earlier evidence would no longer be applicable. Interventions focusing exclusively on specialist nursing services were also excluded from this review because the GDG considered this evidence not generalisable to an overall specialist management approach.

People with cardiogenic shock were included as a subgroup of people with acute heart failure. The GDG agreed to include evidence which described people with myocardial infarction complicated by cardiogenic shock, but considered this to be an indirect population.

In chapter 5.1, because there was a lack of evidence for a comparison between left ventricular assist devices and medical care, evidence from an indirect randomised trial was included. This study included people with 'chronic end stage heart failure' and used the procedure as destination therapy which is not current practice in the UK.

Conference abstracts were not automatically excluded from the review but were initially assessed against the inclusion criteria and then further processed only if no other full publication was available for that review question, in which case the authors of the selected abstracts were contacted for further information. The reviews that included abstracts were:

- In chapter 7: Inotropes (dobutamine) - for this review a trial which was not fully published (Zairis et al., 2004 CASINO trial as described in Cleland 2004^{42,43}) was included. The authors did not reply to a request for further information on methods and additional outcomes.
- Abstracts were also included in pharmacological treatment after stabilisation in chapter 9. For the timing of beta-blocker and aldosterone antagonist commencement reviews one abstract each was included as they described results of data from the OPTIMIZE registry. The design, rationale and methods of analysis of this registry were described in depth in other publications^{50,82,149}.

Composite outcomes were usually excluded. However, an exception was made for the composite endpoint of 'mortality or heart failure hospitalisations' for reviews where limited or no other evidence was available.

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

The review protocols are presented in Appendix C.

3.3.2 Methods of combining clinical studies

3.3.2.1 Data synthesis for intervention reviews

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

For continuous outcomes, measures of central tendency (mean) and variation (standard deviation) were required for meta-analysis. Data for continuous outcomes were analysed using an inverse variance method for pooling weighted mean differences and, where the studies had different scales, standardised mean differences were used. The generic inverse variance option in RevMan5 is used if any studies reported solely summary statistics and 95% confidence interval (95% CI) or standard error; this included any hazard ratios reported. When the only evidence available was based on studies that summarised results by presenting medians (and interquartile ranges), or only p values were given, this information was assessed in terms of the study's sample size and was included in the GRADE tables without calculating the relative or absolute effects. Consequently, aspects of quality

assessment such as imprecision of effect could not be assessed for evidence of this type. Where reported, time-to-event data was presented as a hazard ratio.

Stratified analyses were predefined for some review questions at the protocol stage when the GDG identified that these strata are different in terms of biological and clinical characteristics and the interventions were expected to have a different effect on subpopulations. It was decided at the outset that acute heart failure refers to distinct subpopulations, i.e. acute heart failure with pulmonary oedema, cardiogenic shock, acute right-sided heart failure, and acute decompensated chronic heart failure.

Statistical heterogeneity was assessed by visually examining the forest plots, and by considering the chi-squared test for significance at $p < 0.1$ or an I-squared inconsistency statistic (with an I-squared value of more than 50% indicating considerable heterogeneity). Where considerable heterogeneity was present, we carried out predefined subgroup analyses – see protocols in Appendix C.

Assessments of potential differences in effect between subgroups were based on the chi-squared tests for heterogeneity statistics between subgroups. If no sensitivity analysis was found to completely resolve statistical heterogeneity then a random-effects (DerSimonian and Laird) model was employed to provide a more conservative estimate of the effect.

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

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

3.3.2.2 Data synthesis for prognostic factor reviews

Odds ratios (ORs), risk ratios (RRs) or hazard ratios (HRs), with their 95% confidence intervals (95% CIs) for the effect of the pre-specified prognostic factors were extracted from the papers. Studies of at lower risk of bias were preferred, taking into account the analysis and the study design. In particular, prospective cohort studies were preferred that reported multivariable analyses, including key confounders as identified by the GDG at the protocol stage for that outcome. A narrative summary of results from univariate analyses was also given, highlighting the very high risk of bias as there was a high chance of unknown real effect due to lack of controlling for potential confounders. Data were not combined in meta-analyses for prognostic studies.

3.3.2.3 Data synthesis for diagnostic test accuracy review

Data and outcomes

For the reviews of diagnostic test accuracy, a positive result on the index test was found if the patient had values of the measured quantity above a threshold value, and different thresholds could be used. Diagnostic test accuracy measures used in the analysis were: sensitivity, specificity, positive and negative predictive value, area under the Receiver Operating Characteristics (ROC) curve. The threshold of a diagnostic test is defined as the value at which the test can best differentiate between those with and without the target condition and, in practice, the thresholds used varies amongst studies. In the one diagnostic review for this guideline, sensitivity was given more importance than specificity since natriuretic peptide testing is used as a 'rule out' test. This means that the test is carried out to minimise the false negative test results. The GDG defined the clinically relevant natriuretic thresholds to be used in the analysis based on the thresholds described in the current European heart failure guideline^{138,139} (see chapter 6.1 for details).

Data synthesis

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

To allow comparison between tests, summary ROC curves (by type of Natriuretic Peptide and by threshold level) were generated for each diagnostic test from the pairs of sensitivity and specificity calculated from the 2x2 tables, selecting 1 threshold per study. A ROC plot shows true positive rate (sensitivity) as a function of false positive rate (1 minus specificity). Data were entered into RevMan5 and ROC curves were fitted using the Moses Littenburg approach. In order to compare diagnostic tests, 2 or more tests were plotted on the same graph. The performance of the different diagnostic tests was then assessed by examining the summary ROC curves visually: the test that had a curve lying closest to the upper left corner (100% sensitivity and 100% specificity) was interpreted as the best test.

A second analysis was conducted on studies that used two types of natriuretic peptides in the same study population. Results were plotted on one graph indicating paired results for each study. Paired results could show whether one peptide performed consistently better within study populations.

For those studies that reported Area under the ROC curve (AUC) data, these were also plotted on a graph, for each diagnostic test. The AUC describes the overall diagnostic accuracy across the full range of thresholds. The GDG agreed on the following criteria for AUC:

- ≤ 0.50 : worse than chance
- 0.50–0.60: very poor
- 0.61–0.70: poor
- 0.71–0.80: moderate
- 0.81–0.90: good
- 0.91–1.00: excellent or perfect test.

Heterogeneity or inconsistency amongst studies was visually inspected in the forest plots where appropriate (only when there were similar thresholds).

When data from 5 or more studies were available, a diagnostic meta-analysis was carried out. Study results were pooled using the bivariate method for the direct estimation of summary sensitivity and specificity using a random effects approach (in WinBUGS® software –for the program code see Appendix J). This model also assesses the variability between studies by incorporating the precision by which sensitivity and specificity have been measured in each study. A confidence ellipse is shown in the graph that indicates the confidence region around the summary sensitivity / specificity point.

3.3.3 Type of studies

For most intervention reviews in this guideline, parallel randomised controlled trials (RCTs) were included because they are considered the most robust type of study design that could produce an unbiased estimate of the intervention effects. If the GDG believed RCT data were not appropriate or there was limited evidence from RCTs, well-conducted non-randomised studies were included. Please refer to Appendix C for full details on the study design of studies selected for each review question.

Where data from observational studies were included, the GDG decided that the results for each outcome should be presented separately for each study and meta-analysis not conducted.

For diagnostic reviews, cross-sectional and retrospective studies were included. For prognostic reviews, prospective and retrospective cohort studies were included. Case-control studies were not included.

3.3.4 Appraising the quality of evidence by outcomes

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

The evidence for each outcome was examined separately for the quality elements listed and defined in Table 2. Each element was graded using the quality levels listed in Table 3. For each of these quality elements evidence for each outcome is downgraded where applicable using the following levels.

The main criteria considered in the rating of these elements are discussed below (see Section 3.3.5 Grading of evidence). Footnotes were used to describe reasons for grading a quality element as having serious or very serious problems.

The ratings for each component are summed to obtain an overall assessment for each outcome. The grades described above lead to an overall quality rating as described in Table 4. For example, if the quality element 'risk of bias' is downgraded twice and 'imprecision' downgraded once, an overall rating of 'Very low' is given for this outcome and any further low or high risks in other quality elements will not change this rating.

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

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

Quality element	Description
Risk of bias ('Study limitations')	Limitations in the study design and implementation may bias the estimates of the treatment effect. High risk of bias for the majority of the evidence decreases confidence in the estimate of the effect.
Inconsistency	Inconsistency refers to an unexplained heterogeneity (as assessed by the I or Chi-squared statistic in intervention reviews or visual inspection of paired sensitivity / specificity forest plots in diagnostic reviews (i.e. when point estimates in sensitivity and specificity vary widely across studies).
Indirectness	Indirectness refers to differences in study population, intervention / diagnostic index test, comparator / diagnostic comparator test and outcomes between the available evidence and the review question, or recommendation made, such that the effect estimate is changed.

Quality element	Description
Imprecision	Intervention reviews: results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect. Imprecision results if the confidence interval includes the clinically important threshold but ranges from appreciable benefit to no effect or possible harm. Diagnostic reviews: results are considered to be imprecise if the confidence interval around either the pooled (or if not pooled the median) sensitivity / specificity ranges by between 10-20% (serious) and above 20% or more (very serious).
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. This aspect was not assessed in the diagnostic test accuracy review.

For each of these quality elements, evidence for each outcome is downgraded where applicable using the following levels.

Table 3: Levels of quality elements in GRADE

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

Table 4: Overall quality of outcome evidence in GRADE

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

3.3.5 Grading the quality of clinical evidence

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

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

The details of the criteria used for each of the main quality element are discussed further in the following Sections 3.3.6 to 3.3.10.

3.3.6 Risk of bias

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

The risk of bias for a given study and outcome is associated with the risk of over- or underestimation of the true effect. The common risks of bias are listed in Table 5.

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

The GDG accepted that investigator blinding in surgical intervention studies was impossible and participant blinding was also impossible to achieve in most of these situations. Nevertheless, open-label studies for surgery were downgraded when the outcomes were subjectively measured to maintain a consistent approach in quality rating across the guideline.

Table 5: Risk of bias in randomised controlled trials

Risk of bias	Explanation
Allocation concealment	Those enrolling patients are aware of the group to which the next enrolled patient will be allocated (this is a major problem in 'pseudo' or 'quasi' randomised trials with, for example, allocation by day of week, birth date, chart number)
Lack of blinding	Patient, caregivers, those recording outcomes, those adjudicating outcomes, or data analysts are aware of the arm to which patients are allocated
Incomplete accounting of patients and outcome events	Missing data not accounted for and failure of the trialists to adhere to the intention-to-treat principle when indicated
Selective outcome reporting	Reporting of some outcomes and not others on the basis of the results
Other risks of bias	For example: <ul style="list-style-type: none"> • Stopping early for benefit observed in randomised trials, in particular in the absence of adequate stopping rules • Use of unvalidated patient-reported outcomes • Recruitment bias in cluster-randomised trials

3.3.6.1 Diagnostic studies

For diagnostic accuracy studies, the Quality Assessment of Diagnostic Accuracy Studies version 2 (QUADAS-2) checklist was used (see Appendix F in The guidelines manual¹⁵⁵). Risk of bias and applicability in primary diagnostic accuracy studies in QUADAS-2 consists of 4 domains (see Figure 2):

- Patient selection
- Index test
- Reference standard
- Flow and timing

Figure 2: Summary of QUADAS-2 checklist

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

Source: QUADAS-2 website, University of Bristol²¹⁷

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

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

3.3.6.2 Prognostic studies

For prognostic studies, quality was assessed using the checklist for prognostic studies (Appendix I in The guidelines manual¹⁵⁵). The quality rating (Low, High, Unclear) was derived by assessing the risk of bias across 6 domains: selection bias, attrition bias, prognostic factor bias, outcome measurement bias, control for confounders and appropriate statistical analysis, with the last 4 domains being assessed for each outcome. A summary table on the quality of prognostic studies is presented at the beginning of each review to summarize the risk of bias across the 6 domains. More details about the quality assessment for prognostic studies are shown below:

- The study sample represents the population of interest with regard to key characteristics
- Missing data are unrelated to key characteristics, sufficient to limit potential bias – reasons for missing data are adequately described.
- The prognostic factor of interest is adequately measured in study participants.
- The outcome of interest is adequately measured in study participants.
- Important potential confounders are accounted for appropriately.
- The statistical analysis is appropriate for the design of the study, limiting potential for the presentation of valid results.

3.3.7 Inconsistency

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

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

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

3.3.8 Indirectness

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

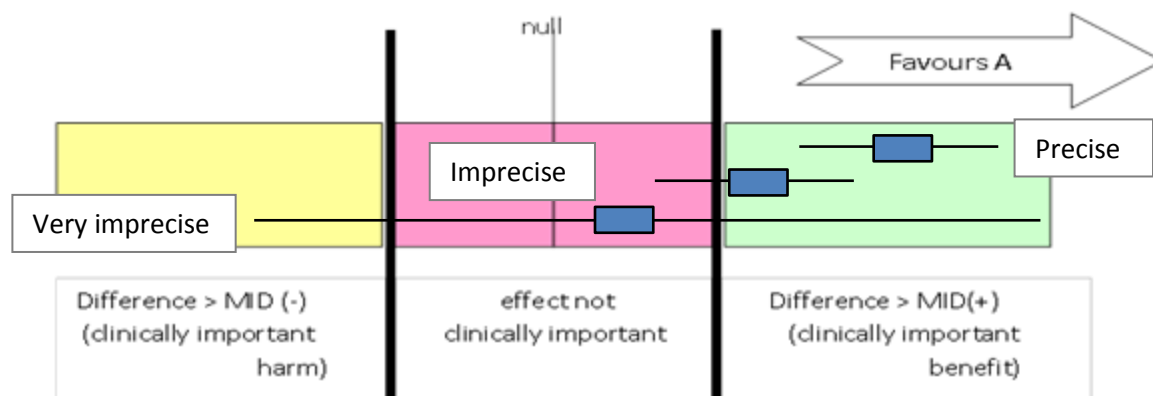
3.3.9 Imprecision

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

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

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

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



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

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

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

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

The default MID for binary outcomes is a 25% relative risk reduction or relative risk increase, and for continuous outcomes, half of the median control standard deviations. Unless there are established MIDs available in the literature, these default MIDs are used in the GRADE quality rating for imprecision.

The literature was searched for established MIDs for outcomes. In addition, the GDG was asked whether they were aware of any acceptable MIDs in the clinical community. The literature review did not identify any particular MIDs for the outcomes of interest. However, in the review in chapter 6.3 on the topic of invasive monitoring, authors of the ESCAPE trial⁶⁶ reported a minimal important difference for the Minnesota Living with Heart Failure scale. A difference of 5 points was reported to be the MID.

3.3.10 Publication bias

This was assessed when a minimum of 5 studies were available for a critical outcome. A funnel plot is a scatter plot of the treatment effects estimated from individual studies against a measure of study size (as indicated by study precision). This plot was visually inspected to assess the symmetry around the pooled estimate of the meta-analysis. When asymmetry is detected, i.e. a relative lack of studies with points in the lower left or right side of the funnel, it suggests that publication bias is present.

This translates into a risk that smaller studies (i.e. low on the vertical axis) with or without a particular intervention effect (i.e. towards the left or right of the pooled estimate line) were less likely to be published.

3.3.11 Assessing clinical importance (benefit, harm or no difference)

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.

The GDG considered an MID based on the point estimate of the absolute effect for intervention studies. For all outcomes, the GDG used the robustness of the evidence, i.e. GRADE rating, as well as the absolute effect (if positive) of the outcome of interest to decide whether the intervention could be considered beneficial for this outcome. The same point estimate but in the opposite direction would apply if the outcome was negative.

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.

3.3.12 Evidence statements

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

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

3.4 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.¹⁵⁵ 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 a new cost-effectiveness analysis to cover priority areas.

3.4.1 Literature review

The health economist:

- Identified potentially relevant studies for each review question from the economic search results by reviewing titles and abstracts. Full papers were then obtained.
- Reviewed full papers against pre-specified inclusion and exclusion criteria to identify relevant studies (see below for details).
- Critically appraised relevant studies using the economic evaluations checklist as specified in The guidelines manual.¹⁵⁵
- Extracted key information about the studies' methods and results into evidence tables (included in Appendix H).
- Generated summaries of the evidence in NICE economic evidence profiles (included in the relevant chapter for each review question) – see below for details.

3.4.1.1 Inclusion and exclusion criteria

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

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

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

For more details about the assessment of applicability and methodological quality see the economic evaluation checklist (Appendix F of The guidelines manual.¹⁵⁵ and the health economics review protocol in Appendix C).

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

3.4.1.2 NICE economic evidence profiles

The NICE economic evidence profile has been used to summarise cost and cost-effectiveness estimates. The economic evidence profile shows an assessment of applicability and methodological quality for each economic evaluation, with footnotes indicating the reasons for the assessment. These assessments were made by the health economist using the economic evaluation checklist from The guidelines manual.¹⁵⁵ 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 Appendix H 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.¹⁶³

Table 6: Content of NICE economic evidence profile

Item	Description
Study	First author name, reference, date of study publication and country perspective.

Item	Description
Applicability	<p>An assessment of applicability of the study to the clinical guideline, the current NHS situation and NICE decision-making^(a):</p> <ul style="list-style-type: none"> • 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	<p>An assessment of methodological quality of the study^(a):</p> <ul style="list-style-type: none"> • 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.
Incremental cost	The mean cost associated with one strategy minus the mean cost of a comparator strategy.
Incremental effects	The mean QALYs (or other selected measure of health outcome) associated with one strategy minus the mean QALYs of a comparator strategy.
Cost effectiveness	Incremental cost-effectiveness ratio (ICER): the incremental cost divided by the incremental effects.
Uncertainty	A summary of the extent of uncertainty about the ICER reflecting the results of deterministic or probabilistic sensitivity analyses, or stochastic analyses of trial data, as appropriate.

(a) *Applicability and limitations were assessed using the economic evaluation checklist in Appendix G of The guidelines manual (2012)*¹⁵⁵

3.4.2 Undertaking new health economic analysis

As well as reviewing the published economic literature for each review question, as described above, a new economic analysis was undertaken by the health economist to cover 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.

The GDG identified serum natriuretic peptide testing and specialist management units as the highest priority areas for original economic modelling. Early echocardiography was also prioritised but this was not subsequently modelled because there was not the data to quantify the incremental costs and benefits. These areas were prioritised because they potentially have a higher patient and cost impact than other areas of the guideline, and because of significant variation in clinical practice.

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

- Methods were consistent with the NICE reference case.¹⁵⁶
- The GDG was involved in the design of the model, selection of inputs and interpretation of the results.
- Model inputs were based on the systematic review of the clinical literature supplemented with other published data sources where possible.

- When published data was not available, GDG expert opinion was used to populate the model.
- Model inputs and assumptions were reported fully and transparently.
- The results were subject to sensitivity analysis, and limitations were discussed.
- The model was peer-reviewed by another health economist at the NCGC.

Full methods for the combined cost-effectiveness analysis of natriuretic peptide testing and specialist management units are described in Appendix M.

3.4.3 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.¹⁵⁴ 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'.¹⁵⁴

3.4.4 In the absence of economic evidence

When no relevant published studies were found, and a new 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 review of effectiveness evidence.

3.5 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 Appendices H and I.
- Summary of clinical and economic evidence and quality (as presented in Chapters [5-11]).
- Forest plots and summary ROC curves (Appendix [I-J]).
- A description of the methods and results of the cost-effectiveness analysis undertaken for the guideline (Appendix [M]).

Recommendations were drafted on the basis of the GDG 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, it was assessed whether the net benefit justified any differences in costs.

When clinical and economic evidence was of poor quality, conflicting or absent, the 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 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 3.5.1 on research recommendations below).

Members of the GDG reviewed all recommendations at the end of guideline development in a confidential online survey to gauge the level of support and provide space for free text comments. Recommendations where one or more members disagreed with the wording of a recommendation or where particular issues were raised in the free text comments were discussed again to resolve any particular concerns.

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

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

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

3.5.1 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.

3.5.2 Validation process

This guidance 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 when the pre-publication check of the full guideline occurs.

3.5.3 Updating the guideline

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

3.5.4 Disclaimer

Health care 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.

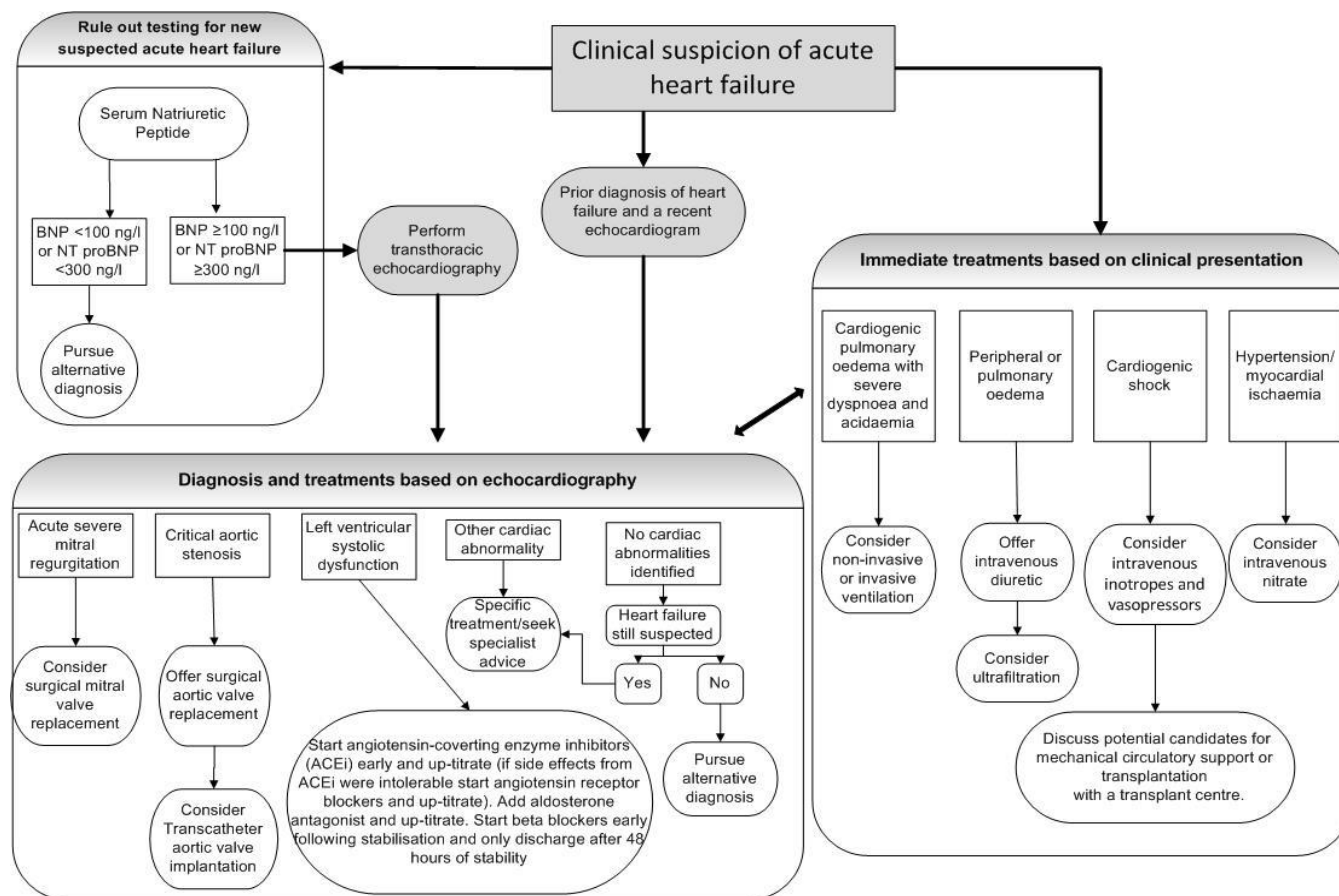
3.5.5 Funding

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

4 Guideline summary

4.1 Algorithm

Figure 4: Diagnostic and treatment algorithm for clinical suspicion of acute heart failure



4.2 Key priorities for implementation

From the full set of recommendations, the GDG selected ten key priorities for implementation. The criteria used for selecting these recommendations are listed in detail in The guidelines manual.¹⁵⁵ The reasons that each of these recommendations was chosen are shown in the table linking the evidence to the recommendation in the relevant chapter.

- All hospitals admitting people with suspected acute heart failure should provide a specialist heart failure team that is based on a cardiology ward and provides outreach services.
- Ensure that all people being admitted to hospital with suspected acute heart failure have early and continuing input from a dedicated specialist heart failure team.
- In people presenting with new suspected acute heart failure, use a single measurement of serum natriuretic peptides (B-type natriuretic peptide [BNP] or N-terminal pro-B-type natriuretic peptide [NT-proBNP]) and the following thresholds to rule out the diagnosis of heart failure.
 - o BNP less than 100 ng/litre
 - o NT-proBNP less than 300 ng/litre.
- In people presenting with new suspected acute heart failure with raised natriuretic peptide levels (see recommendation 6), perform transthoracic Doppler 2D echocardiography to establish the presence or absence of cardiac abnormalities.
- In people presenting with new suspected acute heart failure, consider performing transthoracic Doppler 2D echocardiography within 48 hours of admission to guide early specialist management.
- In a person presenting with acute heart failure who is already taking beta-blockers, continue the beta-blocker treatment unless they have a heart rate less than 50 beats per minute, second or third degree atrioventricular block, or shock.
- Start or restart beta-blocker treatment during hospital admission in people with acute heart failure due to left ventricular systolic dysfunction, once their condition has been stabilised – for example, when intravenous diuretics are no longer needed.
- Ensure that the person's condition is stable for typically 48 hours after starting or restarting beta-blockers and before discharging from hospital.
- Offer an angiotensin-converting enzyme inhibitor (or angiotensin receptor blocker if there are intolerable side effects) and an aldosterone antagonist during hospital admission to people with acute heart failure and reduced left ventricular ejection fraction. If the angiotensin-converting enzyme inhibitor (or angiotensin receptor blocker) is not tolerated an aldosterone antagonist should still be offered.

4.3 Full list of recommendations

Organisation of care

1. All hospitals admitting people with suspected acute heart failure should provide a specialist heart failure team that is based on a cardiology ward and provides outreach services.
2. Ensure that all people being admitted to hospital with suspected acute heart failure have early and continuing input from a dedicated specialist heart failure team.
3. Plan the following with people with acute heart failure in line with Chronic heart failure (NICE clinical guideline 108):
 - discharge from hospital after the acute phase and
 - subsequent management in primary care, including ongoing monitoring and care provided by the multidisciplinary team and
 - information and communication about their condition, its treatment and prognosis.
4. A follow-up clinical assessment should be undertaken by a member of the specialist heart failure team within 2 weeks of the person being discharged from hospital.

Diagnosis, assessment and monitoring

5. Take a history, perform a clinical examination and undertake standard investigations – for example, electrocardiography, chest X-ray and blood tests – in line with Chronic heart failure (NICE clinical guideline 108).
6. In people presenting with new suspected acute heart failure, use a single measurement of serum natriuretic peptides (B-type natriuretic peptide [BNP] or N-terminal pro-B-type natriuretic peptide [NT-proBNP]) and the following thresholds to rule out the diagnosis of heart failure.
 - BNP less than 100 ng/litre
 - NT-proBNP less than 300 ng/litre.
7. In people presenting with new suspected acute heart failure with raised natriuretic peptide levels (see recommendation 6), perform transthoracic Doppler 2D echocardiography to establish the presence or absence of cardiac abnormalities.
8. In people presenting with new suspected acute heart failure, consider performing transthoracic Doppler 2D echocardiography within 48 hours of admission to guide early specialist management.
9. Do not routinely offer pulmonary artery catheterisation to people with acute heart failure.

Initial pharmacological treatment

10. For guidance on patient consent and capacity follow recommendations 1.2.12 and 1.2.13 in Patient experience in adult NHS services (NICE clinical guideline 138).
11. Do not routinely offer opiates to people with acute heart failure.
12. Offer intravenous diuretic therapy to people with acute heart failure. Start treatment using either a bolus or infusion strategy.

13. For people already taking a diuretic, consider a higher dose of diuretic than that on which the person was admitted unless there are serious concerns with patient adherence to diuretic therapy before admission.
14. Closely monitor the person's renal function, weight and urine output during diuretic therapy.
15. Discuss with the person the best strategies of coping with an increased urine output.
16. Do not routinely offer nitrates to people with acute heart failure.
17. If intravenous nitrates are used in specific circumstances, such as for people with concomitant myocardial ischaemia, severe hypertension or regurgitant aortic or mitral valve disease, monitor blood pressure closely in a setting where at least level 2 care^b can be provided.
18. Do not offer sodium nitroprusside to people with acute heart failure.
19. Do not routinely offer inotropes or vasopressors to people with acute heart failure.
20. Consider inotropes or vasopressors in people with acute heart failure with potentially reversible cardiogenic shock. Administer these treatments in a cardiac care unit or high dependency unit or an alternative setting where at least level 2 care^b can be provided.

Initial non-pharmacological treatment

21. Do not routinely use non-invasive ventilation (continuous positive airways pressure [CPAP] or non-invasive positive pressure ventilation [NIPPV]) in people with acute heart failure and cardiogenic pulmonary oedema.
22. If a person has cardiogenic pulmonary oedema with severe dyspnoea and acidaemia consider starting non-invasive ventilation without delay:
 - at acute presentation or
 - as an adjunct to medical therapy if the person's condition has failed to respond.
23. Consider invasive ventilation in people with acute heart failure that, despite treatment, is leading to or is complicated by:
 - respiratory failure or
 - reduced consciousness or physical exhaustion.
24. Do not routinely offer ultrafiltration to people with acute heart failure.
25. Consider ultrafiltration for people with confirmed diuretic resistance^c.

Treatment after stabilisation

26. In a person presenting with acute heart failure who is already taking beta-blockers, continue the beta-blocker treatment unless they have a heart rate

^b Level 2 care is for people needing more detailed observation or intervention, including support for a single failing organ system or postoperative care and for those stepping down from higher levels of care. From Intensive Care Society, Levels of Critical Care for Adult Patients (2009).

^c Diuretic resistance is defined as dose escalation beyond a person's previously recognised dose ceiling or a dose approaching the maximum recommended daily dose without incremental improvement in diuresis. From Diuretics and ultrafiltration in acute decompensated heart failure.

less than 50 beats per minute, second or third degree atrioventricular block, or shock.

27. Start or restart beta-blocker treatment during hospital admission in people with acute heart failure due to left ventricular systolic dysfunction, once their condition has been stabilised – for example, when intravenous diuretics are no longer needed.
28. Ensure that the person's condition is stable for typically 48 hours after starting or restarting beta-blockers and before discharging from hospital.
29. Closely monitor the person's renal function, electrolytes, heart rate, blood pressure and overall clinical status during treatment with beta-blockers, aldosterone antagonists or angiotensin-converting enzyme inhibitors.
30. Offer an angiotensin-converting enzyme inhibitor (or angiotensin receptor blocker if there are intolerable side effects) and an aldosterone antagonist during hospital admission to people with acute heart failure and reduced left ventricular ejection fraction. If the angiotensin-converting enzyme inhibitor (or angiotensin receptor blocker) is not tolerated an aldosterone antagonist should still be offered.

Valvular surgery and percutaneous intervention

31. Offer surgical aortic valve replacement to people^d with heart failure due to severe aortic stenosis assessed as suitable for surgery.
32. Consider transcatheter aortic valve implantation (TAVI) in selected people^h, with heart failure caused by severe aortic stenosis, who are assessed as unsuitable for surgical aortic valve replacement. Details of all people undergoing TAVI should be entered into the UK Central Cardiac Audit database.
33. For guidance on coronary revascularisation see Chronic heart failure (NICE clinical guideline 108).
34. Consider surgical mitral valve repair or replacement for people with heart failure due to severe mitral regurgitation assessed as suitable for surgery.

Mechanical assist devices

35. At an early stage, the specialist should have a discussion with a centre providing mechanical circulatory support about:
 - people with potentially reversible severe acute heart failure or
 - people who are potential candidates for transplantation.

^d For information about patient selection, see Transcatheter aortic valve implantation for aortic stenosis (NICE interventional procedure guidance 421).

4.4 Key research recommendations

- 1. In people with acute heart failure, congestion and worsening renal function, does the addition of low-dose dopamine to standard therapy lead to greater diuresis and renal protection compared with adding placebo to standard therapy?**
- 2. In people with acute heart failure and persistent congestion, does the addition of a thiazide diuretic to standard therapy lead to greater diuresis compared with adding placebo to standard therapy?**
- 3. In people with acute heart failure and hypoperfusion syndrome, is the use of intra-aortic balloon counter-pulsation pump (IABP) better than the use of intravenous inotropes?**
- 4. In people with decompensated heart failure, fluid congestion and diuretic resistance, does ultrafiltration lead to more rapid and effective decongestion compared with continuing diuretic treatment?**

5 Organisation of care

5.1 Specialist management units

Patients with acute heart failure are usually admitted to secondary care facilities via the accident and emergency department. Frequently, patients with acute pulmonary oedema are admitted to intensive care units, high dependency units or the cardiac care units. The remaining patients with acute heart failure are admitted to the medical admission unit, from where they are triaged into either the general medical wards or to the cardiology wards. This practice is not standardised across hospitals and variable factors affect the decision of placing the patient with acute heart failure into different wards. These include age, co-morbidity, bed-availability or being recently cared for by a certain medical unit or firm. The National Heart Failure Audit in England⁴¹ reported that the management and outcomes of patients hospitalised with acute heart failure differ depending on the unit they were admitted to. It is therefore important to consider the best way to deliver optimal care to patients with acute heart failure. A review of the different modes of care currently delivered enables these services to be compared in order to propose the most cost effective model of care for acute heart failure patients.

Review question: For people with suspected or confirmed acute heart failure is a specialist management unit more clinically or cost effective than general medical hospital care?

For full details see review protocol in Appendix C.

Table 7: PICO characteristics of review question

Population	Adults with suspected or confirmed acute heart failure
Intervention/s	Treatment in a specialist management unit of non-geographical design (MDT, multi professional team, specialist care, specialist team) treatment in a specialist management unit of geographical design (heart failure unit, cardiology unit,
Comparison/s	Treatment in a general medical ward Other wards: e.g. surgical wards, care of the elderly wards, other wards
Outcomes	Mortality Major cardiovascular events Length of hospital stay and re-admission rates including length of stay and readmission to critical care to critical care units Quality of life / patient satisfaction Adverse events
Study design	Systematic reviews, RCTs and observational studies

5.1.1 Clinical evidence

We searched for systematic reviews, randomised controlled trials or observational studies addressing effectiveness of specialist management units in the care for patients with suspected or confirmed acute heart failure as compared to non-specialist attendance. Studies using a univariate analysis of factors associated with specialist management (i.e. without consideration of confounding factors) were excluded. Six observational studies were included in the review (Auerbach et al, 2000¹⁹, Boom et al, 2012³⁰, Cleland et al, 2012/2013 - National Heart Failure Audit^{40,42}, Howlett et al, 2003¹⁰³, Joynt et al, 2013¹⁰⁸, and Low et al, 2001¹²⁷). Auerbach and colleagues 2000¹⁹ used subgroups from a study on the improvement of care in hospitalised patients (conducted in the USA between 1989 and

1994)¹. Boom et al, 2012 conducted a study in Canada which compared clinical outcomes of patients newly hospitalised for heart failure attended to by cardiologists, generalists with cardiology consult and generalists without cardiology consult. None of the other studies made an explicit differentiation between generalists who did or did not consult a cardiologist. The National Heart Failure Audit of 142 NHS trusts in England and Health Boards in Wales identified patients that were treated in cardiology wards, general wards and other wards and conducted multivariable analyses of the effects of place of care on mortality rates. Howlett et al, 2003 compared internist with cardiologist care in a Canadian tertiary care facility (this study focused on an analysis of predictor for medication prescription at discharge) and Lowe et al, 2000 in Australia compared generalist with specialist care in hospital (as part of a study on selective admission policies). In another study from the USA (Joynt et al, 2013¹⁰⁸) all Medicare records of heart failure admissions to acute care hospitals in 2009 were used to examine the relationship between physician experience and speciality (cardiologists, internists and generalists) in relation to the overall patient volume in hospitals as well as the volume of heart failure cases seen by the physician. Generalist involvement from primary care is an issue particular to the USA setting and the focus in this review is on the comparison between cardiologists and internists. Results from studies were not pooled / synthesised due to the different definitions for both intervention and control, and different sets of factors used between studies in multivariable analyses.

The evidence was divided into three sections:

1. Specialist management compared to generalist management (where generalists could also include internists)
2. Generalists with or without cardiology consult compared to specialist management
3. Type of specialist (cardiologists or internists) compared to generalists taking into account physician and hospital volume

Evidence from these are summarised in the clinical GRADE evidence profile tables below. See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and lists of excluded studies in Appendix K.

Summary of included studies

The characteristics of included studies are briefly outlined in the following tables – for additional study details please see Appendix G.

Specialist management compared to generalist management (where generalists could also include internists)

Four studies compared specialist care to management by generalists (including internists). Each study is briefly summarised in Table 8 below.

Table 8: Summary of studies included in the review

Study	Intervention/comparison	Population	Outcomes	Comments
Auerbach et al, 2000 ¹⁹	Outcomes of patients managed by: cardiologists (42.5% stated that they would be providing care to their patient after discharge)	N=1298 patients hospitalised with an exacerbation of congestive heart failure. Patients of cardiologists were younger, more likely to be male, had a lower Acute Physiology Score	Discharge medication, admission to intensive care and mortality	Participants were recruited in teaching hospital in the USA between 1989 and 1994 as part of the SUPPORT study - Study to Understand Prognoses and Preferences for Outcomes and Risk of

Study	Intervention/comparison	Population	Outcomes	Comments
	vs. general internists (41.5% stated that they would be providing care to their patient after discharge)	(higher scores indicate increased risk of in-hospital death). They also had fewer comorbidities and were more independent in activities of daily living.		Treatments ¹ Not described whether or what proportion of general internists had consulted a specialist Clear description of multivariable model which also included a propensity score to adjust for the differences in the two groups
Cleland et al, 2012 / 2013 ^{40,41}	Longitudinal audit report of with the latest covering 145 out of 150 NHS Trusts in England and Health Boards in Wales	N=36,788 index admissions and N=7,106 readmissions 55% of men were treated on cardiology wards, compared with 44% of women. Women were more likely to be treated on general medical wards (46% vs. 36%) and other wards (11% vs. 9%). The likelihood of being treated on a cardiology ward decreased with age: 65% of patients who were 16-74% were treated on cardiology wards, compared with 43% of patients in the ≥ 75 years.	All cause and cardiovascular mortality	In the latest report of the audit a Cox proportional hazard model was employed which adjusted risks according to multiple variables selected from a literature review rather than based on statistical significance.
Howlett et al, 2003 ^{103,103}	Groups of patients managed by a cardiologist vs those seen by an internist	Total N =185 consecutive patients admitted to a tertiary care facility with a primary diagnosis of congestive heart failure Patients treated by a cardiologist were younger and had a larger left ventricular end-diastolic diameter by echocardiography.	Independent factors associated with medication prescription: ACE inhibitors and beta-blocker therapy	Indirect outcome
Lowe et al, 2000 ¹²⁷	Specialist vs. generalist care	N=256 patients admitted with congestive heart failure as defined by the Framingham criteria. Patients admitted under general physicians were older and more likely to	Multivariate results restricted to length of stay and mortality.	Aim of the study was to evaluate a new admitting policy in which patients with identifiable single system disorders were admitted to the relevant subspecialist

Study	Intervention/comparison	Population	Outcomes	Comments
		have impaired renal function and chest infections. The prevalence of diabetes mellitus was higher in the patients care for by generalists but this difference did not reach statistical significance. 47% of patients admitted under the cardiologists had a co-morbidity score greater than one, whereas 70% of patients admitted under the general physicians had an elevated score. 16% of patients admitted under the general physicians had either previously seen a cardiologist or were seen by one during or after admission.		<p>whereas patients with multiple medical problems were admitted under a general physician.</p> <p>16% of patients admitted under the general physicians had either previously seen a cardiologist or were seen by one during or after admission.</p> <p>Description of the multivariate results in the abstract contradicts the result section.</p>

Generalists with or without cardiology consult compared to specialist management

In one study generalist care was divided into generalists who consulted a cardiologist and those who did not. Both groups were then compared to specialist management.

Table 9: Summary of studies included in the review

Study	Intervention/comparison	Population	Outcomes	Comments
Boom et al, 2012 ³⁰	Outcomes of patients managed by a cardiologist vs. generalist with cardiology consult or generalist without cardiology consult	<p>N=7634 patients newly hospitalised for heart failure to acute care hospital corporations with congestive heart failure.</p> <p>Patients of cardiologists were younger more frequently male and were the least likely to have a do not resuscitate order. They were also more likely to have a cardiovascular related comorbidity including previous myocardial infarction, angina, arrhythmia and prior cardiac surgery.</p>	Mortality and readmission rates	<p>Post-hoc analysis of a randomised controlled trial (Tu et al, 2009 – EFFECT trial²¹⁶). This was an RCT in which hospital corporations were randomized to early or delayed feedback of a public report card. This was done in order to evaluate whether the public release of data on cardiac quality indicators effectively stimulates hospitals to undertake quality improvement activities that</p>

Study	Intervention/comparison	Population	Outcomes	Comments
				improve health care processes and patient outcomes.

Type of specialist (cardiologists or internists) compared to generalists taking into account physician and hospital volume

In one study physicians were divided into cardiologists, internists and generalists (these are particular. This study also looked at the relationship between the hospital size and physician experience (as measured in cases seen by the physician per year).

Table 10: Summary of studies included in the review

Study	Intervention/comparison	Population	Outcomes	Comments
Joynt et al, 2013 ¹⁰⁸	Outcomes of patients managed by a cardiologists, internists vs. generalist	<p>Medicare records of patients discharged from acute care hospital in the USA, with a primary discharge diagnosis of heart failure.</p> <p>For the purpose of the comparison in the current chapter the focus is on 81, 136 patients cared for by physicians with a 'medium volume' of HF patients. A narrative summary will also be provided on trends related to different physician volume categories (N=390,066 patients cared for by physicians with 'lowest', 'low', 'high' and 'highest' volumes of HF patients)</p>	Mortality and readmission rates	The study adjusted the rates to account for differences in patient characteristics (age, sex, race and 29 comorbid medical conditions) as well as hospital characteristics (volume, teaching status, hospital size, urban versus rural location, region of the country and non-profit vs for-profit ownership)

Table 11: GRADE clinical evidence profile: specialist vs. non-specialist management of patients with suspected or confirmed acute heart failure (multivariable analysis from observational study; data not pooled in an overall meta-analysis)

Quality assessment							Summary of Findings				Quality	Importance
							Specialist Number of event / Total N (%) or Mean (SD)	Generalist Number of event / Total N (%) or Mean (SD)	Effect Adjusted Odds Ratio (OR) (95% CI)	Absolute effect / Mean difference (MD)		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations						
Mortality ¹⁹ (HR adjusted for 12 different patient characteristics (including age, sex, comorbidities, medication at admission, ethnicity and others) as well as a propensity score which took into account life extending care and resuscitation preference and demographic information such as income and level of education)												
1 30 days	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ^(a)	none	75/743 (10.1%)	50/555 (9%)	HR 0.78 (0.48 to 1.28)	19 fewer per 1000 (from 46 fewer to 24 more)	VERY LOW	CRITICAL
1 180 days	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	201/743 (27.1%)	149/555 (26.8%)	HR 0.72 (0.54 to 0.96)	67 fewer per 1000 (from 9 fewer to 113 fewer)	LOW	
1 1 year	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ^(a)	none	281/743 (37.8%)	211/555 (38%)	HR 0.82 (0.65 to 1.04)	56 fewer per 1000 (from 113 fewer to 12 more)	VERY LOW	
1 Maximum	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	438/743 (59%)	351/555 (63.2%)	HR 0.80 (0.66 to 0.96)	81 fewer per 1000 (from 15 fewer to 149 fewer)	LOW	

Quality assessment							Summary of Findings				Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Specialist Number of event / Total N (%) or Mean (SD)	Generalist Number of event / Total N (%) or Mean (SD)	Effect Adjusted Odds Ratio (OR) (95% CI)	Absolute effect / Mean difference (MD)		
follow-up (median 4.6 years)												
Mortality ^{40,41} - HR in-hospital adjusted for age, sex, main place of care, NYHA class III/IV, systolic blood pressure, valve disease, sodium and urea level, heart rate, haemoglobin level, creatinine and potassium level; 30-day mortality and 4 year follow-up adjusted also for additional drug effects (ACE/ARB, loop diuretics, beta-blockers), cardiology follow-up as well as length of stay; 1 and 3 year follow-up HR data was from a previous audit report that adjusted for fewer characteristics.												
1 In-hospital	observational study / audit	no serious limitation	no serious inconsistency	no serious indirectness	no serious imprecision	none	1150/16514 (7%)	1499/13221 (1.1%)	HR 1.54 (1.38 to 1.72)	56 fewer per 1000 (from 74 fewer to 40 fewer)	LOW	CRITICAL
1 30-day	observational study / audit	no serious limitation	no serious inconsistency	no serious indirectness	no serious imprecision	none	767/15364 (5%)	812/11722 (6.9%)	HR 1.25 (1.05 to 1.5)	17 fewer per 1000 (from 33 fewer to 3 fewer)	LOW	
1 1 year	observational study / audit	serious ^(b)	no serious inconsistency	no serious indirectness	no serious imprecision	none	2944/13463 (21.9%)	4222/13834 (30.5%)	HR 1.10 (1.03 to 1.17)	25 fewer per 1000 (from 42 fewer to 8 fewer)	VERY LOW	
1 3	observational study /	serious ^(b)	no serious inconsistency	no serious indirectness	no serious imprecision	none	9971/32074	14581/33999	HR 1.11 (1.08 to	34 fewer per 1000 (from 46 fewer to	VERY LOW	

Quality assessment							Summary of Findings					
							Specialist Number of event / Total N (%) or Mean (SD)	Generalist Number of event / Total N (%) or Mean (SD)	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			Adjusted Odds Ratio (OR) (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
years	audit						(31.1%)	(42.9%)	1.15	25 fewer)		
14 years	observational study / audit	no serious limitation	no serious inconsistency	no serious indirectness	no serious imprecision	none	15812/46323 (34.1%)	17476/38193 (45.8%)	HR 1.14 (1.11 to 1.18)	45 fewer per 1000 (from 57 fewer to 35 fewer)	LOW	
Mortality ¹²⁷ - OR adjusted for co-morbidities, use of ACE-inhibitors, NYHA grade and whether or not the admission was the first with heart failure												
1 In hospital	observational studies	serious ^(c)	no serious inconsistency	no serious indirectness	no serious imprecision	none	12/102 (11.8%)	6/154 (3.9%)	OR 3.1 (1.1 to 8.74)	73 more per 1000 (from 4 more to 223 more)	VERY LOW	CRITICAL
1 28 days	observational studies	serious ^(c)	no serious inconsistency	no serious indirectness	no serious imprecision	strong association ^(d)	16/102 (15.7%)	8/154 (5.2%)	OR 4.3 (1.5 to 12.33)	139 more per 1000 (from 24 more to 351 more)	LOW	
1 1 year	observational studies	serious ^(c)	no serious inconsistency	no serious indirectness	serious ^(a)	none	38/102 (37.3%)	47/154 (30.5%)	OR 1.6 (0.85 to 3.01)	108 more per 1000 (from 33 fewer to 264 more)	VERY LOW	
Transfer to the intensive care unit ^{19,19} - adjusted for 12 different patient characteristics (including age, sex, comorbidities, medication at admission, ethnicity and others)												
1	observation	no	no serious	no serious	no serious	strong	27/74	8/555	OR 2.8	25 more per 1000	MOD	IMPORT

Quality assessment							Summary of Findings				Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Specialist Number of event / Total N (%) or Mean (SD)	Generalist Number of event / Total N (%) or Mean (SD)	Effect	Absolute effect / Mean difference (MD)		
	al studies	serious risk of bias	inconsistency	indirectness	imprecision	association ^(d)	3 (3.6%)	(1.4%)	Adjusted Odds Ratio (OR) (95% CI)	(from 8 more to 52 more)	ERATE	ANT
Beta-blocker prescription at discharge ^{19,103}												
2	observational studies	serious ^(e)	no serious inconsistency	no serious indirectness	serious ^(a)	none	6/743 (0.81%) 45/65 (69.2%)	5/555 (0.9%) 59/120 (49.2%)	OR 1.0 (0.49 to 2.1) OR 1.30 (1.07 to 1.64)	0 fewer per 1000 (from 5 fewer to 10 more) 65 more per 1000 (from 17 more to 122 more)	VERY LOW	IMPORTANT
ACE inhibitor prescription at discharge ¹⁸ adjusted for 12 different patient characteristics (including age, sex, comorbidities, medication at admission, ethnicity and others)												
1	observational studies	serious ^(e)	no serious inconsistency	no serious indirectness	serious ^(a)	none	63/743 (8.5%)	65/555 (11.7%)	OR 1.15 (0.82 to 1.6)	15 more per 1000 (from 19 fewer to 58 more) ^(e,f)	VERY LOW	IMPORTANT
Diuretic prescription at discharge ¹⁸ adjusted for 12 different patient characteristics (including age, sex, comorbidities, medication at admission, ethnicity and others)												
1	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ^(a)	none	80/743 (10.8%)	84/555 (15.1%)	OR 0.85 (0.6 to 1.3)	20 fewer per 1000 (from 55 fewer to 37 more)	VERY LOW	IMPORTANT

- (a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed one of the default MIDs (i.e. ranging from a noticeable effect to no effect). Outcomes were downgraded by two increments if 95% CI of the effect crossed both default MIDs (i.e. ranging from a noticeable improvement to a noticeable harm). Default MIDs are set at ORs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the median control group standard deviation either side of the null line for continuous variables.
- (b) The audit report does not clearly describe how the confounders in the multivariate analysis were selected.
- (c) For a study with N=256 participants a high number of variables (19) were used in the multivariable analysis. The reporting of which variables were used as confounders for the mortality analysis is a bit unclear. No adjustment for multiple comparisons were reported.
- (d) RRs of >2 are considered to be a large effect to upgrade quality of observational data by one increment and >5 upgraded by two increments.
- (e) This study uses a retrospective cohort design and is therefore more prone to selection bias and it does not include sufficient numbers of participants N=185 for the number of variable that are reported and included in the multivariate analysis
- (f) For the findings of Howlett et al, 2003 on ACE-inhibitor prescription please see narrative summary below

Narrative summary

Length of hospital stay - Lowe et al, 2000¹²⁷

It is stated in the report that Length of hospital stay was reduced by 5% (95% CI -23% to 17%) which was described as p=n.s. This percentage was adjusted for co-morbidities, use of ACE-inhibitors, NYHA grade and whether or not the admission was the first with heart failure (VERY LOW QUALITY)

ACE inhibitor and angiotensin II antagonist therapy – Howlett et al, 2003¹⁰³

Specialist care was not an independent predictor of ACE inhibitor or angiotensin II antagonist therapy prescription (i.e. p>0.05 since only independent predictors listed in the results tables) - (VERY LOW QUALITY)

Table 12: GRADE clinical evidence profile: specialist vs. non-specialist management of patients with suspected or confirmed acute heart failure (multivariable analysis from observational study; data not pooled in an overall meta-analysis)

Quality assessment	Summary of Findings			Quality	Importance
	Specia	Gener	Effect		

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	list Numb er of event / Total N (%) or Mean (SD)	alist Numb er of event / Total N (%) or Mean (SD)	Adjusted Odds Ratio (OR) (95% CI)	Absolute effect / Mean difference (MD)		
Mortality ³⁰ - Generalist with cardiology consult - OR adjusted for odds ratios adjusted age, sex, respiratory rate, systolic blood pressure, urea nitrogen, haemoglobin, serum sodium concentration, history of cerebrovascular disease, dementia, chronic obstructive pulmonary disease, cirrhosis, and cancer (patients without ‘do not resuscitate’ order)												
1 30 days	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ^(a)	none	91/1523 (6%)	102/1210 (8.4%)	OR 0.70 (0.42 to 1.18)	24 fewer per 1000 (from 47 fewer to 14 more)	VERY LOW	CRITICAL
1 1 year	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ^(a)	none	353/1523 (23.2%)	374/1210 (30.9%)	OR 1.03 (0.83 to 1.28)	6 more per 1000 (from 38 fewer to 55 more)		
Mortality ³⁰ - Generalist without cardiology consult - (rates adjusted as described see above)												
1 30 days	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ^(a)	none	91/1523 (6%)	564/4901 (11.5%)	OR 1.34 (0.94 to 1.91)	33 more per 1000 (from 6 fewer to 84 more)	VERY LOW	CRITICAL
1 1 year	observational studies	no serious risk of serious	no serious inconsistency	no serious indirectness	no serious imprecision	none	353/1523 (23.2%)	1676/4901 (34.2%)	OR 1.22 (1.02 to 1.44)	46 more per 1000 (from 4 more to 86 more)	LOW	
Readmission for heart failure ³⁰ - Generalist with cardiology consult (rates adjusted as described see above)												
1 30	observational studies	serious ^(b)	no serious inconsistency	no serious indirectness	very serious ^(a)	none	unclear	unclear	OR 0.95 (0.70 to	not calculable ^(c)	VERY LOW	IMPORTANT

Quality assessment							Summary of Findings				Quality	Importance
							Specialist Number of event / Total N (%) or Mean (SD)	Generalist Number of event / Total N (%) or Mean (SD)	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			Adjusted Odds Ratio (OR) (95% CI)	Absolute effect / Mean difference (MD)		
days									1.29)			
1 year	observational studies	serious(b)	no serious inconsistency	no serious indirectness	serious ^(a)	none			OR 1.00 (0.83 to 1.20)			
Readmission for heart failure ³⁰ - Generalist without cardiology consult (adjusted as described see above)												
1 30 days	observational studies	serious ^(b)	no serious inconsistency	no serious indirectness	serious ^(a)	none	unclear	unclear	OR 0.81 (0.64 to 1.03)	not calculable ^(c)	VERY LOW	IMPORTANT
1 year	observational studies	serious ^(b)	no serious inconsistency	no serious indirectness	no serious imprecision	none			OR 0.95 (0.70 to 1.29)			

(a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed one of the default MID (i.e. ranging from a noticeable effect to no effect). Outcomes were downgraded by two increments if 95% CI of the effect crossed both default MID (i.e. ranging from a noticeable improvement to a noticeable harm). Default MID are set at ORs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the median control group standard deviation either side of the null line for continuous variables.

(b) Overall readmission rate not reported

(c) Only adjusted ORs were reported but the exact rate of readmission events was unclear. Hence absolute effect could not be calculated

Table 13: GRADE clinical evidence profile: specialist (cardiologist / internist) vs. non-specialist management of patients with suspected or confirmed acute heart failure (multivariable analysis from observational study; data not pooled in an overall meta-analysis)

Quality assessment							Summary of Findings					
							Specialist Number of event / Total N (%) or Mean (SD)	Generalist Number of event / Total N (%) or Mean (SD)	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			Adjusted Odds Ratio (OR) (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
Cardiologist vs. generalist - Mortality ¹⁰⁸ (rates adjusted for patient characteristics (age, sex, race and 29 comorbid medical conditions) as well as hospital characteristics (volume, teaching status, hospital size, urban versus rural location, region of the country and non-profit vs for-profit ownership))												
1 30 day	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	891/14604 (6.1%)	2688/24665 (10.9%)	RR 0.56 (0.52 to 0.6)	48 fewer per 1000 (from 44 fewer to 52 fewer)	LOW	CRITICAL
Internist vs. generalist - Mortality ¹⁰⁸ (adjusted as described see above)												
1 30 day	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	4061/41866 (9.7%)	2688/24665 (10.9%)	RR 0.89 (0.85 to 0.93)	12 fewer per 1000 (from 8 fewer to 16 fewer)	LOW	CRITICAL
Cardiologist vs. generalist – Readmission rate ¹⁰⁸ (adjusted as described see above)												
1	observational	no	no serious	no serious	no serious	none	3271/	5648/2	RR 0.98	5 fewer per 1000	LOW	IMPORT

Quality assessment							Summary of Findings						
							Specialist Number of event / Total N (%) or Mean (SD)	Generalist Number of event / Total N (%) or Mean (SD)	Effect				
No of studies	Follow-up length	Design	Risk of bias	Inconsistency	Indirectness	Imprecision			Other considerations	Adjusted Odds Ratio (OR) (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
30 day		al studies	serious risk of bias	inconsistency	indirectness	imprecision		14604 (22.4%)	4665 (22.9%)	(0.94 to 1.02)	(from 14 fewer to 5 more)		ANT
Internist vs. generalist – Readmission rate ¹⁰⁸ (adjusted as described see above)													
1	30 day	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	9964/41866 (23.8%)	5648/24665 (22.9%)	RR 1.04 (1.01 to 1.07)	9 more per 1000 (from 2 more to 16 more)	LOW	IMPORTANT

Narrative summary – physician volume Joynet et al, 2013¹⁰⁸

Mortality taking into account physician volume

For cardiologists there were similar mortality rates regardless of the volume of patients that was seen by the physician. For internists and generalists the rates of mortality decreases with increasing patient volume, i.e. the more cases the internist or generalist was seeing the lower the mortality rate.

The authors concluded that physician volume is associated with lower rates of mortality, particularly among non-cardiologist physicians.

30 day readmission taking into account physician volume

For all cardiologist, internist and also generalist care there was a pattern that with increasing physician volume the 30 day readmission rates increased.

5.1.2 Economic evidence

Published literature

No relevant economic evaluations were identified. Evidence from the literature focussed on the cost-effectiveness of outpatients clinics.

See also the study selection flow chart in Appendix E.

5.1.2.1 New cost-effectiveness analysis – model specification

None of the included economic evaluations assessed the cost effectiveness of natriuretic peptide testing in a UK NHS setting, nor did they assess the economic impact of the test beyond one year. NP testing was prioritised for original economic analysis. A single cost-effectiveness model was constructed to evaluate both NP testing and specialist management [Section 5.1, 6.1]. Full details of the model can be found in Appendix M.

The analysis population was adults who present to the emergency department with acute dyspnoea and who are suspected of acute heart failure, that is, they have no clear alternative diagnosis. Four strategies were compared:

- | | |
|---|----------|
| 1. Standard management | (STM) |
| 2. Standard management with natriuretic peptide testing | (STM-NP) |
| 3. Specialist management | (SPM) |
| 4. Specialist management with natriuretic peptide testing | (SPM-NP) |

Strategies 1 and 3 include current standard clinical investigations, specifically the clinical history, physical examination, electrocardiography and chest radiography. Strategies 2 and 4 include these standard investigations plus serum natriuretic peptide, specifically the B-type natriuretic peptide (BNP) test using a rule-out threshold of 100ng/L (where results of less than 100ng/L indicate that the patient does not have acute heart failure). The NT-proBNP test, using a rule-out threshold of 300ng/L, is assessed in a sensitivity analysis.

In the base case analysis 'standard management' strategies placed 50% of the patients with a positive work-up for heart failure in specialist heart failure team care on a cardiology ward, and the other 50% in general medical team care on non-cardiology wards. Specialist management strategies also placed 50% of the patients with a positive work-up in specialist heart failure team care on the

cardiology ward, but places the remainder into joint team care, in non-cardiology wards. Therefore, standard and specialist strategies differed in the way that the non-cardiology ward patients were given care:

- Standard management: received care from non-cardiologists
- Specialist management: received care from non-cardiologists *and* a cardiology 'outreach' service (from a specialist heart failure team)

Both standard and specialist arrangements deal in the same way with patients whose work-up is negative for acute heart failure: they are admitted to non-cardiology wards and receive care from non-cardiologists only.

The model consists of a decision tree for each strategy which divides patients according to their underlying condition, diagnostic work-up and care pathway. At the end of each path in the decision tree is a cohort (Markov) model which estimates each cohort's survival, QALYS and costs over 4 years (the follow-up of the National Heart Failure audit).

5.1.2.2 Methods relating to specialist management

Differences in mortality and heart failure re-admission were mainly predicated on differences in the number of patients on effective LVSD drug therapy (beta blockers, ace inhibitors and aldosterone antagonists) at discharge.

For each drug class logistic regression modelling was conducted using the National Heart Failure Audit to estimate the propensity of being prescribed the drug with or without specialist care and controlling for potential confounders (systolic blood pressure; haemoglobin; NYHA class; urea; creatinine; serum sodium; serum potassium; age; gender; previous COPD, MI, ischemic heart disease, vascular disease). This analysis was conducted specifically for this model by NICOR so that potential confounders were controlled.

Table 14: Probability of receiving LVSD drug treatment, by class and type of care¹⁵³

LVSD drug class	Care from Specialist heart failure team ^(a)	Care without a Specialist heart failure team ^(a)
Angiotensin Converting Enzyme inhibitor (ACEi) or Angiotensin Receptor Antagonist (ARB)	78.0% (n=5493)	60.7% (n=1462)
Beta Blocker (BB)	86.6% (n=5399)	58.6% (n=1433)
Aldosterone antagonist (AA)	37.9% (n=5389)	17.9% (n=1444)

(a) Adjusted for confounders: Systolic blood pressure; haemoglobin; NYHA class; urea; creatinine; serum sodium; serum potassium; age; gender; previous COPD, MI, ischemic heart disease, vascular disease

In the absence of existing published reviews which include the major trials and report the relevant outcomes, the relative treatment effects (hazard ratios for cardiovascular mortality and heart failure readmission) were obtained by pooling randomised placebo-controlled trials of patients with chronic heart failure and reduced ejection fraction, with at least 1000 patients per arm.

Table 15: Risk ratios with 95% confidence intervals (compared with placebo)

LVSD drug class	Cardiovascular mortality	Heart failure readmission	Included trials
-----------------	--------------------------	---------------------------	-----------------

LVSD drug class	Cardiovascular mortality	Heart failure readmission	Included trials
Angiotensin Converting Enzyme inhibitor (ACEi) or Angiotensin Receptor Antagonist (ARB)	0.91 (0.85, 0.97)	0.80 (0.75, 0.86)	SOLVD-T, CHARM, Val-HeFT
Beta Blocker (BB)	0.78 (0.71, 0.86)	0.77 (0.72, 0.84)	BEST, MERIT-HF, CIBIS-2, COPERNICUS
Aldosterone antagonist (AA)	0.80 (0.65, 0.90)	0.65 (0.54, 0.78)	EMPHASIS

In addition there was assumed to be a benefit from specialist management for the in-hospital period. The national heart failure audit showed the mortality of patients during their index hospital stay is improved if care included input from a cardiologist, other physician with an interest in heart failure, or heart failure specialist nurse.⁴¹ An original analysis was conducted for this model by NICOR so that a separate effect could be applied for LVSD patients and so that the control group were patients who had not received any specialist heart failure or cardiologist care. After adjustment for the same confounders as above, using a multivariate cox regression model, a hazard ratio of 1.94 (95%CI: 1.62, 2.33) was calculated for LVSD patients for in-hospital mortality under a general medical team versus a cardiology team¹⁵³.

The base case model was conservative with respect to specialist management strategies in that:

- No benefits from specialist care were assumed for patients with acute heart failure of non-LVSD cause (even though costs of specialist management were attributed to these patients).
- Baseline readmission rate for patients on no drugs was taken from a study of patients who were receiving some medication. Hence the effect on re-admissions assumed was modest.
- The time horizon was only 4 years, possibly under-estimating the QALYs gained.

5.1.2.3 Methods relating to diagnostic accuracy

The model combined existing evidence on BNP test accuracy (from a diagnostic meta-analysis conducted for this guideline – see section 5.1.1) with clinical outcomes from the UK (the national heart failure audit), as described in Table 16.

Table 16: Sensitivity and specificity of the diagnostic work-up using the BNP test and the physician using standard clinical investigations

Method diagnostic work-up	Sensitivity	Specificity	Source
BNP Test	0.95	0.63	Guideline meta-analysis. See Table 22
Physician with access to standard clinical investigations (ECG, E-ray, clinical examination)	0.80	0.77	Breathing Not Properly Trial ¹³⁷

The consequences (mortality, risk of readmission, and length of index hospital stay) of false positive and false negative diagnostic work-ups were based on expert clinical opinion:

- A 2-day length of stay penalty for patients who are falsely assessed as being likely/ unlikely to have AHF work-up.

- A treatment delay of 3 months and a 1/3 probability of readmission during those 3 months for 20% of patients with heart failure with a false work-up is not corrected during index admission (i.e. 80% of the few people who have AHF, but are missed by the physician/NP test at work-up, will be identified during the admission - a conservative assumption with regard to the benefits of NP testing).

A point estimate of the diagnostic accuracy was not found for the physician in the absence of NP testing. Instead, there was a ROC curve (a set of alternative pairs of sensitivities and specificities) from the Breathing Not Properly Trial.¹³⁷ The point on the curve closest to the origin was taken as the base case and other points on the curve were looked at in sensitivity analyses.

5.1.2.4 Base case results

Standard management (strategy 1) had the lowest cost but specialist management with NP testing (strategy 4) had the highest QALYs (Table 17). Strategy 4 had the highest cost but was the most cost-effective strategy (Table 18). The increased cost of staff, tests and visits, and drugs was partly offset by the reduced cost of hospital stay (Table 19). Figure 5 shows the results illustrated on the cost-effectiveness plane, where incremental QALYs are plotted against incremental costs.

Table 17: Base case results: Life-years, QALYs and costs (probabilistic)

Strategy		Mean LYs	Mean QALYs	Mean costs
1	Standard Management	3.146	2.206	£2,625
2	Standard Management with NP	3.151	2.209	£2,669
3	Specialist Management	3.169	2.222	£2,673
4	Specialist Management with NP	3.178	2.228	£2,729

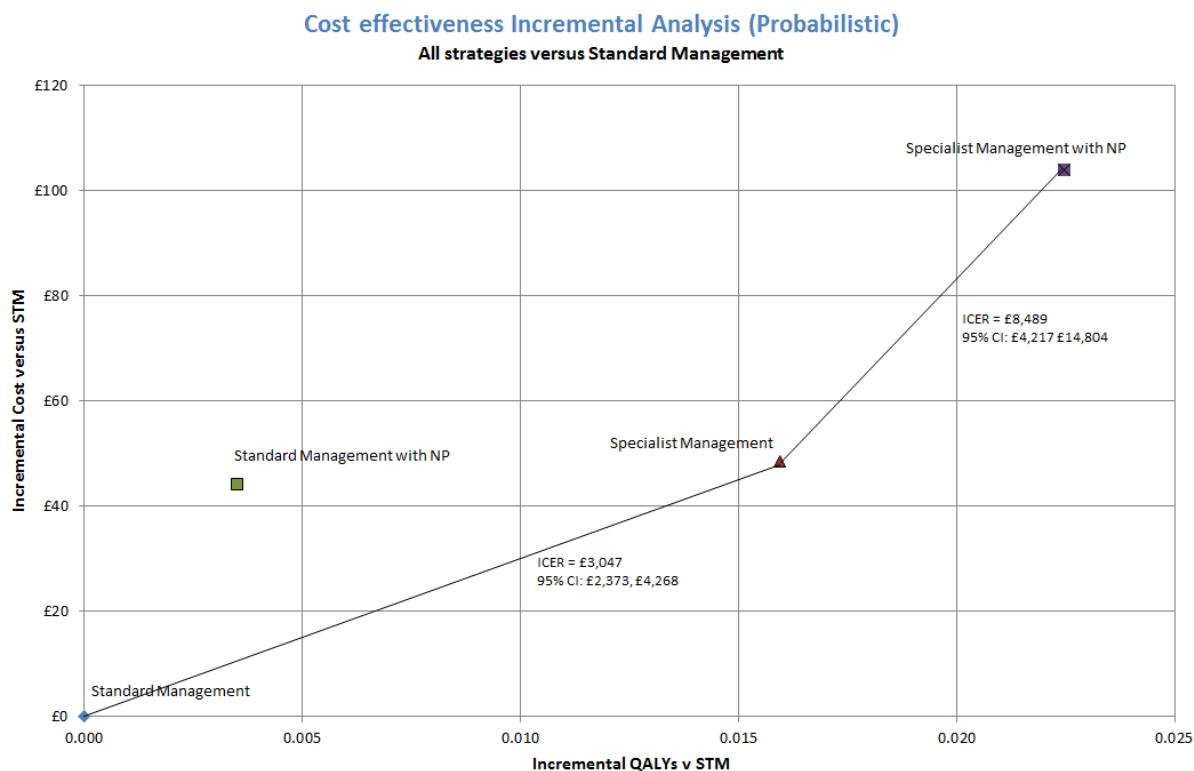
Table 18: Base case results: Cost-effectiveness (probabilistic)

Strategy	NMB (£20 000/QALY)	Rank (£20 000/QALY)	Probability the strategy is the most cost-effective at £20,000 per QALY
1	£41,495	4	0%
2	£41,521	3	0%
3	£41,765	2	0%
4	£41,840	1	>99%

Table 19: Base case results: Breakdown of mean costs (probabilistic) (£ per patient)

Strategy	NP and echocardiography	Index admission	Re-admissions	Drugs and follow-up visits	Total
1	£37	£2,003	£313	£272	£2,625
2	£70	£2,011	£305	£283	£2,669
3	£37	£2,022	£302	£312	£2,673
4	£70	£2,036	£293	£330	£2,729

Figure 5: Incremental costs plotted against incremental QALYs (the cost effectiveness plane)



5.1.2.5 Sensitivity analyses

Sensitivity analysis, in which the key variables (mortality, risk of re-admission, resource use, time-horizon, and proportion of patients receiving care in the specialist ward) were significantly inflated or deflated individually, all found specialist management with NP testing to be the most cost-effective strategy. For example, when the benefits of treatment were assumed to last only for one year, this strategy was still highly cost-effective.

Probabilistic sensitivity analysis in which all input variables were simultaneously varied confirmed this stability: greater than 99% of probabilistic simulations found specialist management with BNP testing to be the optimal strategy.

5.1.3 Evidence statements

5.1.3.1 Clinical

Specialist care vs. generalist (including internists)

Mortality

Auerbach et al, 2000¹⁹:

Low to very low quality evidence from one study comprising 1298 patients with exacerbations of heart failure showed no clear difference in length of survival at 30 day follow-up (very low quality). After half a year there was a lower rate of mortality and longer length of survival associated with management by a specialist (low quality). This was no longer a clear difference at 1 year follow-up (very low quality). However, at the maximum follow-up (with a median of 4.6 years) a lower rate of mortality and longer length of survival was associated with specialist management (low quality evidence). All results were adjusted for 12 different patient characteristics (including age, sex, comorbidities, medication at admission, ethnicity and others) as well as a propensity score which

took into account life extending care and resuscitation preference and demographic information such as income and level of education.

Cleland et al, 2013 National Heart Failure Audit^{40,41}

Low to very low quality evidence from one audit (N=36,788 index admissions and N=7,106 readmissions) showed that in-hospital mortality was lower for patients with heart failure admitted to a cardiology ward compared to patients admitted to a general or other ward. This was the case for in-hospital as well as 30-day at 1, 3 and 4 years follow-up. In-hospital results were adjusted for age, sex, main place of care, NYHA class III/IV, systolic blood pressure, valve disease, sodium and urea level, heart rate, haemoglobin level, creatinine level, and potassium level; in addition to these variable 30-day and 4 year mortality results were adjusted for drug effects (ACEi/ARB, loop diuretics, beta-blockers), cardiology follow-up as well as length of stay; 1 and 3 year follow-up HR data was from a previous audit report that adjusted for fewer characteristics..

Lowe et al, 2000¹²⁷

Very low quality evidence from one observational study (N=256) showed that in-hospital mortality was higher for those attended to by a specialist this was also shown at 28 days follow-up. At 1 year follow-up no clear difference was observed (very low quality evidence). All results were adjusted for differences in co-morbidities, use of ACEis, NYHA grade and whether or not the admission was the first with heart failure.

Transfer to the intensive care unit - Auerbach et al, 2000¹⁹

Very low quality evidence from one observational study (N=1,298) showed that more patients were admitted to the intensive care unit when they were managed by a cardiologist compared to those managed by generalists. This result was adjusted for differences in age, sex, respiratory rate, systolic blood pressure, urea nitrogen, haemoglobin, serum sodium concentration, history of cerebrovascular disease, dementia, chronic obstructive pulmonary disease, cirrhosis, and cancer

Length of stay - Lowe et al, 2000¹²⁷

Very low quality evidence from one study (N=256) reported no clear difference in length of stay between patients managed by a cardiologist and patients managed by generalists. This result was adjusted for group differences in co-morbidities, use of ACEis, NYHA grade and whether or not the admission was the first with heart failure.

Discharge medication - Auerbach et al, 2000¹⁹ and Howlett et al, 2003¹⁰³

Beta-blockers: Very low quality evidence from two observational study (N=1298 and N=185) showed inconsistent results. One of the studies showed no clear differences in beta blocker medication at discharge (adjusting for 12 different patient characteristics, including age, sex, comorbidities, medication at admission, ethnicity and others) whereas the other reported increased prescription rates associated with specialists (after accounting for differences in coexisting acetylsalicylic acid therapy and presence of oedema) but there is uncertainty around this effect.

ACE inhibitors: Very low quality evidence from two observational studies (N=1298 and N=185) showed no differences in the rate of ACE inhibitor medication at discharge depending on whether a generalist or a specialist managed the patient. One study adjusted for 12 different patient characteristics (including age, sex, comorbidities, medication at admission, ethnicity and others) whereas the other reported only the independent predictors and the speciality of the treating physician was not one of them.

Diuretics: Very low quality evidence from one observational study (N=1298) showed no differences in the rate of diuretic medication at discharge depending on whether a generalist or a specialist managed the patient. The study adjusted for 12 different patient characteristics (including age, sex, comorbidities, medication at admission, ethnicity and others).

Generalist with or without cardiology consult vs. specialist care

Mortality- Boom et al, 2012³⁰

Very low quality evidence from one observational study (N=7634) showed no clear differences in mortality for patients with acute heart failure treated in specialist care compared to patients managed by generalists with cardiology consult (follow-up 30 days and 1 year). When comparing specialist care to generalists without cardiology consult no clear difference was observed at 30 days (very low quality evidence), but at 1 year there was a higher rate of mortality in the generalist group who did not consult a cardiologist (low quality evidence). All results were adjusted for differences in age, sex, respiratory rate, systolic blood pressure, urea nitrogen, haemoglobin, serum sodium concentration, history of cerebrovascular disease, dementia, chronic obstructive pulmonary disease, cirrhosis, and cancer. Due to a baseline difference in the rate of patients with 'do not resuscitate' orders this was restricted to those without such an order.

Readmission rate for heart failure - Boom et al, 2012³⁰

Very low quality evidence from one observational study (N=7634) no clear differences in readmission rates for heart failure were seen in the specialist care group compared to generalists with cardiology consult with considerable uncertainty (both at follow-up 30 days and 1 year). This was also the case when specialist care was compared to generalists without cardiology consult. All results were adjusted for differences in age, sex, respiratory rate, systolic blood pressure, urea nitrogen, haemoglobin, serum sodium concentration, history of cerebrovascular disease, dementia, chronic obstructive pulmonary disease, cirrhosis, and cancer.

Type of specialist (cardiologists or internists) vs. generalist care

30 day mortality – Joynt et al, 2013¹⁰⁸

Low quality evidence from one observational study (N=81136 at a medium physician volume) indicated that both cardiologist and internist care was more effective in reducing mortality rates compared to generalist management. When taking physician volume into consideration cardiologists lower mortality rates were constant across different levels of volume whereas for non-cardiologists mortality rates decreased with increasing volume.

30 day readmission rate – Joynt et al, 2013¹⁰⁸

Low quality evidence from one observational study (N=81136 at a medium physician volume) indicated that physician specialty was not associated with a reduction in readmission rates at 30 days. When taking physician volume into consideration readmission rates all increased with increasing physician volume, but even more so for generalists and internists (rate increases of 2.6%, 3.8% and 5.9% for cardiologists, internists and generalists across volume categories).

Economic

- One original cost-utility analysis found that a specialist heart failure management service was cost-effective compared with standard management for patients presenting to the emergency

department with acute dyspnoea. This analysis was assessed as directly applicable with minor limitations.

- o In a context of NP testing: ICER= £3,159 per QALY gained
- o In a context of no NP testing: ICER= £3,047 per QALY gained

5.1.4 Recommendations and link to evidence

Recommendations	<ol style="list-style-type: none"> 1. All hospitals admitting people with suspected acute heart failure should provide a specialist heart failure team that is based on a cardiology ward and provides outreach services. 2. Ensure that all people being admitted to hospital with suspected acute heart failure have early and continuing input from a dedicated specialist heart failure team. 3. Plan the following with people with acute heart failure in line with Chronic heart failure (NICE clinical guideline 108): <ul style="list-style-type: none"> • discharge from hospital after the acute phase and • subsequent management in primary care, including ongoing monitoring and care provided by the multidisciplinary team and • information and communication about their condition, its treatment and prognosis. 4. A follow-up clinical assessment should be undertaken by a member of the specialist heart failure team within 2 weeks of the person being discharged from hospital.
Relative values of different outcomes	The GDG considered mortality and readmission rates to be the most important outcomes.
Trade-off between clinical benefits and harms	Overall, specialist management was associated with lower mortality, particularly in the most applicable audit report (the National Heart Failure Audit). One study found lower mortality in patients receiving specialist care as compared to generalist care, but this difference was not apparent if the generalists were receiving cardiologist support. One study found specialist management was associated with greater use of beta-blockers at discharge. .
Economic considerations	<p><i>Cost-effectiveness of a specialist heart failure management service:</i></p> <p>An original economic model with conservative assumptions found specialist heart failure management to be cost effective compared to less specialist approaches for incident patients (i.e. those without previously diagnosed heart failure) suspected of acute heart failure. The cost per QALY gained was £3,159 in the context of NP testing and £3,047 without NP testing.</p> <p>The finding that specialist heart failure management was cost effective was robust to sensitivity analysis in which the key variables (mortality, risk of re-admission, resource use, time-horizon, and proportion of patients receiving care in the specialist</p>

	<p>ward) were adjusted within plausible limits.</p> <p>Probabilistic sensitivity analysis in which all input variables were simultaneously varied confirmed this stability: all 1000 probabilistic simulations found specialist management (with NP testing) to be the optimal strategy</p> <p>The model:</p> <p>Impact of specialist management in terms of mortality and heart failure readmission was based on differences in the number of patients on effective LVSD drug therapy (beta blockers, ace inhibitors and aldosterone antagonists) at discharge.</p> <p>The proportion of patients prescribed drugs in cardiology ward and non-cardiology ward settings was taken from the national heart failure audit (controlling for 13 different patient characteristics including NYHA score).</p> <p>The relative treatment effects (hazard ratios for cardiovascular mortality and heart failure readmission) were obtained from RCT evidence in chronic heart failure populations.</p> <p>The base case model was conservative in that</p> <ul style="list-style-type: none"> • Only benefits from LVSD drug prescribing were attributed to specialist management. • No benefits from specialist care were assumed for patients with acute heart failure of non-LVSD cause (even though costs of specialist management were attributed to these patients). <p>External validity:</p> <p>Four-year survival in the model for patients with LVSD seen by specialists was 35%, compared to 30% for those not receiving specialist input. This survival benefit is similar to that found in randomised evidence included in the clinical review, albeit graded low quality. A US study showed a lower rate of mortality associated with specialist care at the maximum follow-up of 4.6 years (hazard ratio 0.8 [95% CI: 0.66-0.97]).</p> <p>Cardiology ward and outreach:</p> <p>Both the 'specialist' and 'standard' arrangements of care that were modelled assumed that 50% of patients admitted for heart failure are managed on cardiology wards and 50% on general medical wards. In the 'specialist arrangement', an 'outreach' team provides specialist input for the 50% of patients managed on general wards.</p> <p>The 'outreach' is provided by a specialist heart failure team including a cardiologist specialising in heart failure and a specialist nurse, and this is costed in addition to the input from general medicine.</p> <p>In sensitivity analyses the cost of specialist management was varied by changing the ratio of patients on cardiology wards to those being seen by outreach. Regardless of the ratio assumed, specialist management was always cost-effective compared to standard management.</p>
Quality of evidence	<p>The clinical evidence was rated as low to very low according to the GRADE criteria. The evidence was entirely from observational data. The GDG agreed that the results of such studies should be interpreted with caution, even though reasonable adjustments had been made for differences in baseline characteristics. In the GDG discussion, greater weight was given to the National Heart Failure audit as it was the largest and most applicable study to UK NHS practice, though it was given a very low</p>

	quality GRADE rating.
Other considerations	<p>The GDG discussed the various ways of providing specialist management care and focused on two main issues; team composition and whether or not there should be a discrete management unit. The GDG considered early identification by specialist nursing staff to be important, particularly in hospitals without a separate cardiology unit. The GDG proposed that hospitals should develop a system to alert the heart failure specialist team of new admissions. The results from the observational studies indicated that generalist care could be effective when combined with cardiology input. This suggests that a 'roaming' specialist might be an alternative model to a geographically discrete specialist management unit, particularly for those patients with multiple co-morbidities for whom acute heart failure was not the predominant issue. These deliberations informed the construction of the cost-effectiveness analysis that was carried out. GDG acknowledged the point that a sensitivity analysis in the economic model showed that a higher proportion of people cared in a cardiology ward led to greater cost effectiveness. However, in light of the lack of direct clinical evidence the GDG concluded that it was not possible to specify a minimal proportion of patients who should be cared for in a specialist unit.</p> <p>The GDG discussed this at length but it was not possible to specify a specific timeframe as it would be as early as possible and based on clinical judgement.</p> <p>The GDG noted the importance of the transition from hospital to primary care, including discharge planning from hospital and subsequent management in primary care. Ongoing monitoring of the patient post discharge, delivery of care by a community multidisciplinary team and ensuring patients receive the support and information they need are areas considered in the NICE Chronic Heart Failure guideline (CG108). The GDG agreed reference should be made to the recommendations within this guideline.</p>

6 Diagnosis, assessment and monitoring

6.1 Natriuretic peptides

The diagnosis of acute heart failure in people urgently admitted to hospital can sometimes be difficult due to the similarity of the presentation to people with acute respiratory distress from other causes such as pneumonia, infective exacerbation of chronic obstructive pulmonary disease, etc. Not infrequently, patients may be treated with diuretics, antibiotics and nebulised bronchodilators simultaneously until the diagnosis of acute heart failure has been positively confirmed by clinical assessment, initial investigations and echocardiography. In similarity to chronic heart failure, there has been considerable research into the role of natriuretic peptides in making the diagnosis of acute heart failure. This section aims to explore the role of natriuretic peptides in reaching an early diagnosis in people presenting to hospital with acute heart failure.

Review question: In people with suspected (or under investigation for) acute heart failure, is the addition of natriuretic peptides to the standard initial investigations (using ECG, chest x-ray and blood tests) more accurate compared to standard initial investigations, clinical judgement and each other?

This review focuses on the diagnostic accuracy of serum natriuretic peptides in patients presenting in an acute care setting with suspected acute heart failure. A brief description of the protocol is provided in Table 20 below.

For full details see review protocol in Appendix C.

Table 20: PICO characteristics of review question

Population	All adults with suspected (under investigation for) acute heart failure presenting in an acute care (i.e. non primary care) setting.
Index Tests	<p>Serum natriuretic peptides:</p> <p>BNP</p> <p>NT-proBNP</p> <p>ANP</p> <p>NT-proANP</p> <p>mid regional-proANP</p> <p>Data to be extracted for individual natriuretic peptides at the thresholds specified in the European (ESC) Guidelines for heart failure 2102:</p> <p>BNP ≤ 100 pg/mL, 100-500 pg/mL, >500pg/mL</p> <p>NTproBNP ≤ 300 pg/mL, 300-1800 pg/mL, >1800pg/mL</p> <p>MRproANP <120 pmol/L, ≥ 120 pmol/L</p>
Reference Standard	Clinical judgement (including use of ECG, chest x-ray and blood tests)
Outcomes	<p>2x2 tables</p> <p>Sensitivity</p> <p>Specificity</p> <p>PPV</p> <p>NPV</p> <p>Most accurate threshold</p> <p>ROC curve</p> <p>Destination of care</p>

	Consequences of false positive and false negative outcomes
Study design	Cross sectional studies, retrospective or prospective case reviews and cohort studies. Case-control studies will be excluded

Studies examining the use of urinary natriuretic peptides; studies screening for left or right ventricular dysfunction and studies concerning the diagnostic accuracy of natriuretic peptides in pleural effusion of unknown aetiology were excluded from this review.

6.1.1 Clinical evidence

This review focuses on the diagnostic accuracy of the serum natriuretic peptides BNP, NTproBNP and MRproANP in patients presenting in an acute care setting with suspected acute heart failure.

Diagnostic meta-analysis was conducted where appropriate, i.e. when 5 or more studies were available per threshold. Test accuracy for the studies was pooled using the bivariate method modelled in Winbugs®. The bivariate method uses logistic regression on the true positives, true negatives, false positives and false negatives reported in the studies. Summary ROC curves were constructed and confidence regions plotted (using methods outlined by Novielli et al. 2010¹⁶⁰). Forty nine studies were included in the review. Evidence from these are summarised in the clinical GRADE evidence profile below Table 22. See also the study selection flow chart in Appendix D, graphs in Appendix I-J, study evidence tables in Appendix G and exclusion list in Appendix K.

The following review strategy was employed. For each natriuretic peptide data were extracted from individual studies when they reported data from a threshold specified in the European (ESC) Guideline 2012, with some studies contributing data to more than one threshold analysis per peptide. Studies that used a different cut-off threshold (for instance, BNP <90pg/ml) were included for the closest threshold. The ranges used were as follows:

- BNP ≤ 100 pg/mL, 100-500 pg/mL, >500pg/mL
- NTproBNP ≤300 pg/mL, 300-1800 pg/mL, >1800pg/mL
- MRproANP <120 pmol/L, ≥120 pmol/L

The GRADE approach for evidence summaries was adapted to diagnostic test accuracy review (see Table 21), but incorporates the same assessments that are used for interventional reviews (risk of bias, inconsistency, indirectness and imprecision). See also the study selection flow chart in Appendix D, study evidence tables in Appendix G, forest plots in Appendix I, GRADE tables and excluded studies list in Appendix K.

Table 21: Summary of studies included in the review

Study	N	Design	Setting	Index test (Assay)	Reference standard	Comments
Afaq 2011 ⁷	502	Retrospective cohort	ED	NTproBNP Roche	Single physician using Framingham criteria	Results presented stratified by age related threshold. Cumulative data presented without specific threshold AUC only extracted

Study	N	Design	Setting	Index test (Assay)	Reference standard	Comments
Ailbay 2005 ¹¹	160	Cross sectional	ED	BNP Triage NTproBNP Roche	Retrospective review by two cardiologists	No exclusion criteria reported Elderly cohort mean age: 80.1 years
Arques 2005 ¹⁶	70	Prospective cohort	Acute referrals	BNP Triage	Retrospective review by two cardiologists and one pulmonologist	Excludes those with EF < 45%
Arques 2007 ^{16,17}	41	Prospective cohort	ED	BNP Triage	Retrospective review by two cardiologist and one respiratory physician	Study in AF population : Inclusion criteria ≥ 70 years of age permanent non valvular AF; normal LV EF
Barcase 2004 ²¹	98	Prospective cohort	ED	BNP Triage	Retrospective review by one cardiologist	Reference standard not blinded to BNP Male/Female(n):100/0
Behnes 2009: MANPRO study ²⁶	401	Prospective cohort	ED	NTproBNP Dimension Dade	Retrospective review by study physician	N/A
Berdague 2006 ²⁷	254	Prospective cohort	ED	NTproBNP Roche	Retrospective review by two cardiologists	Excludes patients younger than 70
Blonde-Cynober 2011 ²⁹	64	Prospective cohort	Inpatients	BNP Triage	Retrospective review by one cardiologist and one geriatrician	Geriatric inpatients: Elderly cohort mean age: 84.3 years
Chenevier-Gobeaux 2005 ³⁷ (Results without	378	Prospective cohort	ED	BNP Triage NTproBNP Roche MRproANP BRAHMS	Consensus of two senior ED physicians	No exclusion criteria reported

Study	N	Design	Setting	Index test (Assay)	Reference standard	Comments
renal function stratification from Chenevier-Gobeaux 2010 ³⁸⁾						
Chung 2006 ³⁹	143	Prospective cohort	ED	BNP Triage	Retrospective review by one cardiologist	Reference standard not blinded to BNP
Dao 2001 ⁵²	250	Cross sectional	ED	BNP Triage	Retrospective review by two cardiologists	Male/Female (n): 235/15
Davis 2004 ⁵³	52	Prospective cohort	ED	BNP in house assay	Retrospective review by committee of physicians	N/A
Defilippi 2007 ⁵⁵	831	Prospective cohort	ED	BNP Triage NTproBNP Roche	Case report forms reviewed by a cardiologist	Analysis conducted between groups with varying renal function stratified by eGFR Dade Behring Corporation, manufacturer of NTproBNP assay (but not the one evaluated in the study) sponsors
Dokanish 2004 ^{57,58}	122	Cross sectional	Inpatients	BNP Triage	Retrospective review by one cardiologist	Excludes valvular pathology
Eckstein 2012 ⁶²	632	Prospective cohort	ED	NTproBNP Roche	Retrospective review by two cardiologists	Overlap with BACH trial ^{128,129} data. Data only extracted for NTproBNP

Study	N	Design	Setting	Index test (Assay)	Reference standard	Comments
						to avoid double counting for MRproANP Male/Female ratio not reported for included patients
Fabian 2011 ⁶⁸	130	Prospective cohort	Acute admissions	NTproBNP Roche	Clinical diagnosis according to European society of cardiology guidelines	Reference standard not blinded to NTproBNP AUC only extracted
Fleischer 1997 ⁷⁸	123	Prospective cohort	Acute admissions	BNP in house assay	Clinical diagnosis based on intent to treat HF with diuretic therapy for 24 hours	No exclusion criteria reported Reference standard not blinded to BNP
Gargani 2008 ⁸⁴	149	Prospective cohort	Cardiology/Pulmonology admissions	NTproBNP Roche	Retrospective review by two cardiologists	No exclusion criteria reported
Gorissen 2007 ⁹²	80	Retrospective cohort	ED	BNP Triage NTproBNP Roche	Retrospective review by cardiologist and pulmonologist	Excluded patients with no consensus on clinical diagnosis
Gruson 2008 ⁹⁷	137	Prospective cohort	ED	BNP Access NTproBNP Roche	Based upon clinical signs, chest radiography, echocardiography and/or radionuclide angiography	NTproANP results extracted
Gruson 2012 ^{96,97}	153	Prospective cohort	ED	BNP Access NTproBNP Roche	On the basis of clinical signs, chest x-ray, echocardiography and/or radionuclide	Study primarily looking at diagnostic accuracy

Study	N	Design	Setting	Index test (Assay)	Reference standard	Comments
					angiography	of proBNP1- 108 non protocol NP AUC only extracted
Havelka 2011 ¹⁰¹	54	Prospective cohort	ED	BNP Triage	Discharge diagnosis electronic medical record for each patient	Reference standard not blinded to BNP AUC only extracted
Januzzi 2006 ^{105,106} : ICON study (comprising pooled data from Lainchbury 2003 ¹¹⁸ , Bayes Genis 2004 ²⁵ and Januzzi 2005 ¹⁰⁶ : PRIDE study, and unpublished registry data)	1256	Pooled prospective trial data	ED	NTproBNP Roche	Retrospective review utilising European society of cardiology guidelines. "Suitable for pooling across studies"	No exclusion criteria reported The study stratified by age
Karpalitis 2007 ¹⁰⁹	80	Prospective cohort	ICU	BNP Triage	Retrospective review by two intensivists	Patients undergoi g right- heart catheterisa tion in ICU for diagnostic uncertain ty. 6 'mixed diagnosis' patients excluded from analysis
Klemen 2009 ¹¹³	441	Prospective cohort	Pre hospita l emerge	NTproBNP Roche	Final hospital diagnosis confirmed by cardiologists and or intensivists	N/A

Study	N	Design	Setting	Index test (Assay)	Reference standard	Comments
			ncy			
Lainchbury 2003 ¹¹⁸	205	Prospective cohort	ED	BNP Triage (+3 in house assays not extracted)	Retrospective review by two cardiologists	No exclusion criteria reported Only BNP results extracted. NTproBNP from pooled results in ICON study ^{105,106}
Logeart 2002 ¹²⁵	163	Cross sectional	ICU	BNP Triage	Retrospective review by two cardiologists and one pulmonologist	N/A
Lokuge 2010 ¹²⁶	612	Retrospective cohort	ED	BNP Abbott	Retrospective review by one physician and one cardiologist	Excludes patients with cardiogenic shock
Maisel 2002: Breathing Not Properly Study ^{128,130}	1586	Prospective cohort	ED	BNP Triage	Retrospective review by two cardiologists	Financial support from industry
Maisel 2010 ^{128,129} : BACH trial	1641	Prospective cohort	ED	BNP Triage MRproANP BRAHMS	Retrospective review by two cardiologists	N/A
Moe 2007 ¹⁴⁵ data from IMPROVE-CHF study	500	Cohort data from RCT	ED	NTproBNP Roche	Retrospective review by cardiologists	Supported by Roche diagnostics
Mueller 2005: BASEL study ^{147,148} (+Gegenhuber 2006 ⁸⁷)	251	Prospective cohort	ED	BNP Abbot NTproBNP Roche MRproANP BRAHMS	Retrospective review by one study investigator	Retrospective samples MRproANP results measured 1 year after collection Male/Female (n): 234/17
Nazarian 2009 ¹⁵⁸	145	Prospective Cohort	ED	NTproBNP Roche	Retrospective review by two cardiologists and one respiratory physician	Reference standard not blinded to NTproBNP

Study	N	Design	Setting	Index test (Assay)	Reference standard	Comments
Parab 2005 ¹⁶⁷	70	Retrospective cohort	ED	BNP Triage	Retrospective chart review	No exclusion criteria reported Physicians not blinded to BNP
Potocki 2010 ¹⁷⁹	287	Prospective cohort	ED	NTproBNP Roche MRproANP BRAHMS	Retrospective review by two cardiologists	N/A
Prosen 2011 ¹⁸¹	218	Prospective cohort	Pre hospital emergency	NTproBNP Roche	Final hospital diagnosis confirmed by cardiologists and or intensivists	N/A
Ray2004 ¹⁸⁴	150	Cross sectional	ED	BNP Triage	Retrospective review by two experts	Elderly cohort mean age: 80 years
Ray2005 ^{184,185}	202	Cross sectional	ED	NTproBNP Roche	Retrospective review by two experts	Elderly cohort mean age: 80 years Overlap with Ray 2004 ¹⁸⁴ Only NTproBNP results extracted
Rogers 2009 ¹¹² : Substudy of HEARD-IT	740	Prospective cohort	ED	BNP (5 sites Triage; 2 site Abbott)	Retrospective review by two cardiologists	Reference standard not blinded to BNP results Authors received funding from Aboott and Biosite and had shares in company
Sanz 2006 ¹⁹⁵	75	Prospective cohort	ED	BNP Access (+ BNP Advia not extracted) NTproBNP Roche	Diagnosed according to symptoms and signs and ECG, CXR and in some cases	No exclusion criteria reported

Study	N	Design	Setting	Index test (Assay)	Reference standard	Comments
					echocardiography	
Shah 2012 ^{200,201} (from PRIDE study data Januzzi 2005 ¹⁰⁶)	560	Prospective cohort	ED	MRproANP BRAHMS	Retrospective review by one cardiologist	All data presented by age stratification AUC only extracted
Seronde 2013 ¹⁹⁸	336	Prospective cohort	ED	BNP, NT-proBNP, MRproANP	Cardiology discharge diagnosis	AUC only extracted
Shaikh 2011 ²⁰²	100	Cross sectional	ED	NTproBNP Roche	Cardiology discharge diagnosis	Reference standard not blinded to NTproBNP
Villacorta 2002 ²¹⁹	70	Cross sectional	ED	BNP Triage	Retrospective review by one cardiologist	N/A
Wang 2010 ^{222,223}	84	Prospective cohort	ED	BNP Abbott	Retrospective review by two cardiologists	N/A
Zaninotto 2005 ²³²	122	Prospective cohort	ED	NTproBNP Roche	Discharge diagnosis basis of clinical and instrumental investigations	Reference standard not blinded to NTproBNP

Table 22: Adapted GRADE profile for diagnostic test accuracy studies comparing Natriuretic Peptide tests (index test) to clinical diagnosis (reference standard) - (rows represent index tests which are divided by thresholds and the evidence is summarised across studies comparable to interventional GRADE profiles, summary statistics are provided for sensitivity and specificity from meta-analysis and area under curve as reported in studies)

Natriuretic peptide (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity % (95%CI/range)	Specificity % (95%CI/range)	Area Under Curve (range)	Quality
BNP										
BNP	22	7090	No serious risk of bias	Serious inconsistency ^(b)	No serious indirectness	N/A	-	-	0.77 [0.59,0.95] - 0.99 [NR] Median 0.91 [0.90, 0.93]	MODERATE
BNP ≤ 100 pg/mL	19	6950	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision ^(c)	0.95 (0.95-0.95) ^(d)	0.63 (0.62-0.63) ^(d)	-	HIGH
BNP 100-500 pg/mL	20	4543	No serious risk of bias	Serious inconsistency ^(b)	No serious indirectness	No serious imprecision ^(c)	0.85 (0.85-0.85) ^(d)	0.86 (0.86-0.86) ^(d)	-	MODERATE
BNP ≥ 500 pg/mL (not pooled)	4	283	Serious risk of bias ^(a)	Very serious inconsistency ^(b)	No serious indirectness	Very serious imprecision ^(c)	0.35 [0.17, 0.56] - 0.83 [0.69, 0.92] Median 0.61 [0.34, 0.78]	0.78 [0.56, 0.93] - 1.00 [0.91, 1.00] Median 0.79 ^(e) [0.57, 0.93]	-	VERY LOW
NTproBNP										
NTproBNP	21	6756	No serious risk of bias	Serious inconsistency ^(b)	No serious indirectness	N/A	-	-	0.576 [0.476, 0.676] - 0.99 [NR] Median 0.9 [0.84, 0.94]	MODERATE
NTproBNP ≤ 300 pg/mL	10	3349	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	0.99 (0.99-0.99) ^(d)	0.43 (0.43-0.43) ^(d)	-	HIGH
NTproBNP 300-1800 pg/mL	13	3223	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious imprecision ^(c)	0.90 (0.90-0.90) ^(d)	0.76 (0.76-0.77) ^(d)	-	MODERATE
NTproBNP ≥ 1800 pg/mL (not pooled)	3	840	Serious risk of	Serious inconsistency ^(b)	No serious indirectness	Serious imprecision ^(c)	0.67 [0.60, 0.73] - 0.87 [0.81, 0.92] Median 0.83	0.72 [0.63, 0.80] - 0.95 [0.91, 0.98] Median 0.7 ^(e)	-	VERY LOW

Natriuretic peptide (Threshold)	No of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity % (95%CI/range)	Specificity % (95%CI/range)	Area Under Curve (range)	Quality
			bias ^(a)				[0.71, 0.91]	[0.59, 0.8]		
MRproANP (not pooled)										
MRproANP	5	3117	No serious risk of bias	No serious inconsistency	No serious indirectness	N/A	-	-	0.81[0.76,0.84] - 0.90 [0.87, 0.93] Median 0.9 [0.87, 0.93]	HIGH
MRproANP < 120 pmol/L	1	251	No serious risk of bias	N/A	No serious indirectness	No serious imprecision	0.95 [0.90, 0.98]	0.56 [0.47, 0.65]	-	HIGH
MRproANP ≥ 120 pmol/L	4	2557	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	0.84 [0.77, 0.89] - 0.98 [0.94, 1.00] Median 0.93 [0.83, 0.98]	0.40 [0.34, 0.46] - 0.84 [0.77, 0.90] Median 0.68 ^(e) [0.57, 0.83]	-	HIGH

GRADE was conducted with emphasis on test sensitivity as this was the primary outcome for decision making.

- (a) Risk of bias was assessed using the QUADAS-II checklist. Outcomes were downgraded by one if the weighted (by sample size (n)) average number of QUADAS-II domains (patient selection, index test, reference standard and flow and timing) with methodological limitations was one. Outcomes were downgraded by two if the weighted average number of QUADAS-II domains with methodological limitations was more than one.
- (b) Inconsistency was assessed by inspection of the sensitivity / specificity RevMan 5 plots, or summary area under the curve (AUC) plots. Reasons for heterogeneity between studies include use of different assays, different reference standards, and differing settings.
- (c) The judgement of precision for sensitivity and specificity separately was based on visual inspection of the confidence region in the diagnostic meta-analysis; where diagnostic meta-analysis has not been conducted, imprecision was assessed using the confidence interval of the median sensitivity value (it was decided that a CI with a width of <0.2 was deemed to be precise, ≥0.2 – 0.3 was downgraded to serious imprecision and ≥0.3 downgraded to very serious imprecision).
- (d) Pooled sensitivity/specificity and (95%CI) from diagnostic meta-analysis.
- (e) When values are not pooled, the median specificity presented corresponds to the median sensitivity with the corresponding CI to maintain paired values as sensitivity was the primary outcome for decision making.

Narrative evidence

Defilippi et al. 2007⁵⁵ stratified their analysis between groups by varying renal function: eGFR (< or ≥ 60 ml/min/1.73m²). They demonstrated similar accuracy as no significant difference in AUC when split by renal function for NTproBNP versus BNP.

Seronde et al. 2013¹⁹⁸ reported that for all NPs that they tested (BNP, NTproBNP and MRproANP), areas under the curve were greater in the subgroup of people with acute decompensated heart failure compared to people with de novo acute heart failure (details described in supplementary material to the publication).

6.1.2 Economic evidence

6.1.2.1 Published literature

Six studies were included that addressed the relevant comparison^{6,32,145,146,192,204}. These are summarised in the economic evidence profile below (Table 23). See also the full study evidence tables in Appendix H.

Three included studies drew on the same randomised test and treat trial (BASEL)^{6,32,146}. One study was excluded²⁶. This is detailed in Appendix L.

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

Table 23: Summary of economic studies included in the review

First author	Title	Journal	Publication year
AHTA ⁶	B-type natriuretic peptide assays in the diagnosis of heart failure. Part A: in the hospital emergency setting. Part B: in the non-hospital setting	Medical Services Advisory Committee	2007
Breidthardt ³ ^{2(a)}	Medical and economic long-term effects of B-type natriuretic peptide testing in patients with acute dyspnea	Clinical Chemistry	2007
Moe ¹⁴⁵	N-terminal pro-B-type natriuretic peptide testing improves the management of patients with suspected acute heart failure: primary results of the Canadian prospective randomized multicenter IMPROVE-CHF study	Circulation	2007
Mueller ^{146,14} ⁷	Cost-effectiveness of B-Type Natriuretic Peptide Testing in Patients With Acute Dyspnea	Archives of Internal Medicine	2006
Rutten ¹⁹²	N-terminal pro-brain natriuretic peptide testing in the emergency department: beneficial effects on hospitalization, costs and outcome	American Heart Journal	2008
Siebert ²⁰⁴	Cost-effectiveness of using N-terminal pro-brain natriuretic peptide to guide the diagnostic assessment and management of dyspneic patients in the emergency department	American Journal of Cardiology	2006

(a) Breidthardt2007 is a follow-up report of Mueller2006

Table 24: Economic evidence profile: Supplementary natriuretic peptide testing versus Conventional diagnostic assessment – Economic study characteristics

Study	Applicability	Limitations	Other comments	Incremental Cost (per patient)	Incremental Effects (per patient)	Cost-effectiveness	Uncertainty
AHTA 2007 ⁶ Australia	Partly applicable ^(a)	Potentially serious limitations ^(b)	Intervention: BNP CEA of BASEL with Australian costs Analysis at 30-days follow-up Rule-in and -out thresholds	BNP saves £155 ^(c)	All-cause 30-day per patient deaths: -0.026 (CI: 0.032, -0.083)	BNP testing dominates	PSA: probability of BNP being less costly and more effective = 78.8%; less costly and less effective = 18.8%
Mueller 2006 ^{146,147} Switzerland	Partly applicable ^(a)	Minor limitations ^(d)	Intervention: BNP Within-trial CEA of BASEL Analysis at 180-days follow-up Rule-in and -out thresholds	BNP saves £1,650 ^(e)	All-cause 180-day per patient deaths: -0.034 (b)	BNP testing dominates	PSA: probability of BNP being less costly and more effective = 80.6%; less costly and less effective = 19.3%
Breidhardt 2007 ³² Switzerland	Partly applicable ^(a)	Potentially serious limitations ^(f)	Intervention: BNP Within-trial CEA of BASEL Analysis at 360-days economic and 720-days outcomes follow-up Rule-in and -out thresholds	BNP saves £1,669 ^(e)	All-cause 720-day per patient deaths: 0.01 (p=0.58) ^(g)	ICER not reported or calculated by NCGC ^(h)	PSA: probability of BNP being less costly and more effective = 39.5%; less costly and less effective = 59.1% ⁽ⁱ⁾
Moe 2007 ¹⁴⁵ Canada	Partly applicable ^(a)	Potentially serious limitations ^(j)	Intervention: NT-proBNP Within-trial CCA of IMPROVE-CHF Analysis at 60-days follow-up Rule-in and -out thresholds (age-adjusted)	BNP saves £604 (k)	All-cause 60-day per patient deaths: 0.01 (p=0.57)	ICER not reported or calculated by NCGC ^(h)	
Rutten 2008 ¹⁹² Netherlands	Partly applicable ^(a)	Potentially serious limitations ^(l)	Intervention: NT-proBNP Within-trial CEA of a single-centre RCT Analysis at 30-days follow-up Rule-in and -out thresholds (gender-adjusted)	NT-proBNP saves £870 ^(m)	All-cause 30-day per patient deaths: -0.01 (p=0.26)	NT-proBNP testing dominates	PSA: NT-proBNP being less costly and more effective was the most probable outcome.
Siebert	Partly	Potentially	Intervention: NT-proBNP	NT-proBNP	Per patient risk of	NT-proBNP	PSA and DSA showed the

Study	Applicability	Limitations	Other comments	Incremental Cost (per patient)	Incremental Effects (per patient)	Cost-effectiveness	Uncertainty
2006 ²⁰⁴ USA	applicable ^(a)	serious limitations ⁽ⁿ⁾	An economic model based on PRIDE (threshold analysis) Analysis at 60-days follow-up Single rule-in/out threshold	saves £301 (k)	serious adverse event s: -0.004 (CI NR) ^(o)	testing dominates	dominance of NT-proBNP to be robust

- (a) Costs and effects are not measured in a UK NHS context; no HR-QoL data to inform a cost per QALY
- (b) B-type NP thresholds not adjusted for gender, age, renal function or obesity ; short follow-up unlikely to reflect all differences in costs and outcomes
- (c) Currency converted from 2005 Australian dollars to 2005 UK pounds using purchasing power parities for 2005¹⁶³. Costs incorporated emergency care and admitted patient care including cardio-pulmonary investigations, outpatient care, and B-typeNP test
- (d) B-type NP thresholds not adjusted for gender, age, renal function or obesity ; not all relevant costs were included
- (e) Currency converted from 2003 US dollars to 2003 UK pounds using purchasing power parities for 2003¹⁶³. Costs incorporated emergency and admitted patient care (except medication for non-cardiac and non-pulmonary conditions) including cardio-pulmonary investigations, outpatient care, and B-typeNP test
- (f) B-type NP thresholds not adjusted for gender, age, renal function or obesity ; use of different follow-up periods for costs and outcomes would likely bias the cost-effectiveness finding; not all relevant costs were included; some mortality figures reported show contradictory findings which have not been clarified by the authors
- (g) There is an unresolved discrepancy between the reported mortality figures (Incremental 720-day all-cause deaths per patient is reported as both 0.01 and 0.027)
- (h) The mortality rate difference was not significant and the authors concluded that there was no difference in effect on mortality; we have not therefore estimated an ICER
- (i) The joint-probability distribution used different time horizons (360-day costs, 720-day mortality), which may underestimate the effect of incremental cost at 2-years
- (j) NT-proBNP thresholds not adjusted for gender, renal function or obesity ; short follow-up unlikely to reflect all differences in costs and outcomes
- (k) Converted from 2005 Canadian dollars to 2005 UK pounds using purchasing power parities for 2005¹⁶³. Costs incorporated emergency care and admitted patient care including cardio-pulmonary investigations, outpatient care, and B-typeNP test
- (l) NT-proBNP thresholds not adjusted for age, renal function or obesity; short follow-up unlikely to reflect all differences in costs and outcomes
- (m) Currency converted from 2005 US dollars to 2005 UK pounds using purchasing power parities for 2005¹⁶³. Costs incorporated emergency care and admitted patient care including cardio-pulmonary investigations, outpatient care, and BNP test
- (n) Adopted an NT-proBNP rule-out decision only (>900pg/mL); short follow-up unlikely to reflect all differences in costs and outcomes; not all relevant costs were included; use of modelling assumptions relating to the use of echo
- (o) Serious adverse events (SAEs): Urgent care visits, ED presentations, and re-hospitalisations

Abbreviations: CE = cost-effectiveness; CCA = cost-consequence analysis; CEA = cost effectiveness analysis; RCT = Randomised Clinical Trial; CI = 95% confidence interval; ICER = incremental cost-effectiveness ratio; NR = not reported; NCGC = National Clinical Guideline Centre; PSA = probabilistic sensitivity analysis; DSA = deterministic sensitivity analysis; LoS = length of hospital stay.

There was a consistent effect of reduced resource utilisation in patients managed with the supplement of both BNP and NT-pro BNP tests up to one year after presentation (all-studies). This reduction in resources - mainly driven by fewer admissions and reduced length of admission - results directly in a reduction in the cost of management of patients in the intervention groups versus standard assessment groups over the first-year (all-studies).

The effect of reduced all-cause deaths associated with BNP and NT-proBNP in the short-term (AHTA2007 and Rutten2008; 30-days) was consistent with the 180-day follow-up study (Mueller2006; BNP) but not the 60-day follow-up study (Moe; NT-proBNP). Indeed, whilst the long-term follow-up study (Bhreidhardt2007; B-typeNP; 720-days) reported two conflicting findings which require author clarification, neither support improved mortality at this time-point. Therefore the dominance of BNP and NT-proBNP versus standard assessment across shorter time-horizons is probably not maintained beyond the first-year, owing to a diminished outcome effect.

6.1.2.2 New cost-effectiveness analysis

Diagnosis, assessment and monitoring were prioritised for original economic analysis. A single cost-effectiveness model was constructed to evaluate both NP testing and specialist heart failure management. A summary of the model can be found in section 6.1.2 and full details of the model can be found in Appendix M.

6.1.3 Evidence statements

6.1.3.1 Clinical

BNP

- Moderate quality evidence from 22 studies with 7090 participants showed the range in area under the curve (AUC) for BNP was 0.77 [0.59-0.95] to 0.99 [CI NR]; however, we are unable to comment on the uncertainty surrounding these values.
- When diagnostic meta-analysis was conducted, high quality evidence from 19 studies with 6,950 participants showed the pooled sensitivity (SD) and specificity (SD) of BNP at a threshold of ≤ 100 pg/mL were 0.95 (0.01) and 0.63 (0.06) respectively.
- When diagnostic meta-analysis was conducted, moderate quality evidence from 20 studies with 4510 participants showed the pooled sensitivity (SD) and specificity (SD) of BNP at a threshold of 100-500 pg/mL were 0.85 (0.02) and 0.86 (0.03) respectively.
- Very low-quality evidence from 4 studies with 283 participants suggested the sensitivity of BNP at a threshold of ≥ 500 pg/mL ranged from 0.35 [0.17-0.56] to 0.83 [0.69-0.92] and the paired specificity ranged from 0.78 [0.56-0.93] to 1.0 [0.91-1.0]; however, we are unable to comment on the uncertainty surrounding these values.

NTproBNP

- Moderate quality evidence from 21 studies with 6756 participants showed the range in area under the curve (AUC) for NTproBNP was 0.576 [0.476-0.676] to 0.99 [CI NR]; however, we are unable to comment on the uncertainty surrounding these values.
- When diagnostic meta-analysis was conducted, high-quality evidence from 10 studies with 3349 participants showed the pooled sensitivity (SD) and specificity (SD) of NTproBNP at a threshold of ≤ 300 pg/mL were 0.99 (0.01) and 0.43 (0.10) respectively.
- When diagnostic meta-analysis was conducted, moderate-quality evidence from 13 studies with 3223 participants showed the pooled sensitivity (SD) and specificity (SD) of NTproBNP at a threshold of 300-1800 pg/mL were 0.90 (0.02) and 0.76 (0.04) respectively.

Low-quality evidence from 3 studies with 840 participants suggested the sensitivity of NTproBNP at a threshold of ≥ 1800 pg/mL ranged from 0.67 [0.60-0.73] to 0.87 [0.81-0.92] and the paired specificity ranged from 0.72 [0.63-0.80] to 0.95 [0.91-0.98]; however, we are unable to comment on the uncertainty surrounding these values.

MRproANP

- High quality evidence from 5 studies with 3117 participants showed the range in area under the curve (AUC) for MRproANP was 0.81 [0.76-0.84] to 0.90 [0.87-0.93] however we are unable to comment on the uncertainty surrounding these values.
- High quality evidence from 1 study with 251 participants suggested the sensitivity of MRproANP at a threshold of < 120 pmol/L was 0.95 [0.90-0.98] and the paired specificity was 0.56 [0.47-0.65] however we are unable to comment on the uncertainty surrounding these values.
- High quality evidence from 4 studies with 2557 participants suggested the sensitivity of MRproANP at a threshold of ≥ 120 pmol/L ranged from 0.84 [0.77-0.89] to 0.98 [0.94-1.0] and the paired specificity ranged from 0.40 [0.34-0.46] to 0.84 [0.77-0.90] however we are unable to comment on the uncertainty surrounding these values.

6.1.3.2 Economic

- One original cost-utility analysis found that BNP was cost-effective compared with no BNP testing for patients presenting to the emergency department with acute dyspnoea. This analysis was assessed as directly applicable with minor limitations.
 - In a context of specialist management: ICER = £8,489 per QALY gained
 - In a context of non-specialist management: ICER = £12,576 per QALY gained
- Three studies, which drew upon the same clinical trial, found that B-type natriuretic peptide testing dominated conventional diagnostic assessment (less costly and more effective) for patients presenting to the emergency department with acute dyspnoea. Two of these analyses were assessed as partially applicable with minor limitations; one was assessed as partially applicable with potentially serious limitations.
- Two studies found that NT-proBNP testing dominated conventional diagnostic assessment (less costly and more effective) for patients presenting to the emergency department with acute dyspnoea. These analyses were assessed as partially applicable with potentially serious limitations.
- One cost-consequence analysis found that NT-proBNP was less costly and less effective than conventional diagnostic assessment (£604 less per patient; 0.01 more deaths per patient over 60 days) for patients presenting to the emergency department with acute dyspnoea. This analysis was assessed as partially applicable with potentially serious limitations.

6.1.4 Recommendations and link to evidence

Recommendations	<p>5. Take a history, perform a clinical examination and undertake standard investigations – for example, electrocardiography, chest X-ray and blood tests – in line with Chronic heart failure (NICE clinical guideline 108).</p> <p>6. In people presenting with new suspected acute heart failure, use a single measurement of serum natriuretic peptides (B-type natriuretic peptide [BNP] or N-terminal pro-B-type natriuretic peptide [NT-proBNP]) and the following thresholds to rule out the diagnosis of heart failure.</p> <ul style="list-style-type: none"> • BNP less than 100 ng/litre • NT-proBNP less than 300 ng/litre. <p>7. In people presenting with new suspected acute heart failure with raised natriuretic peptide levels (see recommendation 6), perform transthoracic Doppler 2D echocardiography to establish the presence or absence of cardiac abnormalities.</p>
Relative values of different outcomes	The GDG focused on peptide sensitivity and specificity, with sensitivity the primary outcome for decision-making at 'rule-out' thresholds. For the economic analysis, the GDG placed most weight on clinical outcomes including mortality, length of stay and rehospitalisation.
Trade-off between clinical benefits and harms	The use of a natriuretic peptide test may facilitate earlier exclusion of the diagnosis of acute heart failure, and allow other conditions to be diagnosed and treated. As with any diagnostic test, there will be false negative and false positive results. The result must therefore be interpreted within the clinical context in order to guide patient management.
Economic considerations	<p><i>Cost-effectiveness of natriuretic peptide testing:</i></p> <p>NP testing represents an additional cost to standard diagnostic testing protocols but through the use of a rule-out cut-off it could potentially decrease the number of incorrect acute heart failure diagnoses or influence the number of patients going on to have an echocardiogram.</p> <p>BNP and NT-proBNP have the same unit cost (~£28).</p> <p><i>From the existing literature</i></p> <p>6 economic evaluations showed that the use of BNP and NT-proBNP tests to supplement diagnosis reduced the net cost of management in the first year. This was driven by reduced acute care burden due to fewer admissions, fewer re-admissions and a shorter length of stay, but no impact on mortality. None of the studies looked at outcomes beyond one year.</p> <p>The BASEL study, which contributed data to three of the included economic analyses, involved a strategy of BNP testing to inform a pre-specified treatment regimen. One further trial (Moe 2007) included a strategy of repeated natriuretic peptide monitoring. None of the economic analyses were testing a purely diagnostic natriuretic peptide strategy as the interventions being evaluated included pre-specified treatment regimens. Furthermore, the studies were conducted in non-UK healthcare settings. For these reasons an original economic model was developed.</p> <p><i>From the economic model</i></p>

<p>Recommendations</p>	<p>5. Take a history, perform a clinical examination and undertake standard investigations – for example, electrocardiography, chest X-ray and blood tests – in line with Chronic heart failure (NICE clinical guideline 108).</p> <p>6. In people presenting with new suspected acute heart failure, use a single measurement of serum natriuretic peptides (B-type natriuretic peptide [BNP] or N-terminal pro-B-type natriuretic peptide [NT-proBNP]) and the following thresholds to rule out the diagnosis of heart failure.</p> <ul style="list-style-type: none"> • BNP less than 100 ng/litre • NT-proBNP less than 300 ng/litre. <p>7. In people presenting with new suspected acute heart failure with raised natriuretic peptide levels (see recommendation 6), perform transthoracic Doppler 2D echocardiography to establish the presence or absence of cardiac abnormalities.</p>
	<p>The use of a serum BNP or NT-proBNP test was found to be cost effective compared to a physician using only standard clinical investigations. The cost per QALY gained in the context of specialist heart failure management was £8,489 for BNP (and in a sensitivity analysis £13,385 for NT-proBNP).</p> <p>The economic model compared a diagnostic strategy in which NP testing was available to guide use of echocardiography to one in which the decision was made by the physician using only standard investigations. The BNP test used a rule-out threshold of 100 ng/l and the NT-proBNP test used a threshold of 300 ng/l. Echocardiography was assumed to establish the presence or absence of cardiac abnormalities in all patients with natriuretic peptide levels higher than their rule-out thresholds. The impact of the diagnostic strategies was assessed in terms of mortality, readmission and resource utilisation.</p> <p>The model combined existing evidence on BNP test accuracy (from a diagnostic meta-analysis conducted for this guideline) with clinical outcomes from the UK (the national heart failure audit).</p> <p>The consequences (mortality and risk of readmission from heart failure causes) of false positive and false negative diagnostic work-ups were based on expert clinical opinion:</p> <ul style="list-style-type: none"> • A 2 day addition to length of stay for patients who are falsely assessed as being likely/unlikely to have AHF work-up • Of the few people who have AHF but are missed by the physician/NP test at work-up, 80% will be identified during the admission (a conservative assumption with regard to the benefits of NP testing). For the 20% that are not identified, it is assumed that they are correctly diagnosed at 3 months and at that point those with LVSD are put on to appropriate drugs (beta blockers, ace inhibitors and aldosterone antagonists). These 20% were assumed to have a 1/3 probability of re-admission during those 3 months. <p>To estimate the accuracy of physician diagnosis in the absence of NP testing, a ROC curve (a set of alternative pairs of sensitivities and specificities) from the Breathing Not Properly Trial was used. The point on the curve closest to the point of perfect accuracy was taken as the base case and other points on the curve were used in sensitivity analyses. In the development of the model the GDG noted the difficulty of estimating the sensitivity and specificity of the physician using standard clinical</p>

<p>Recommendations</p>	<p>5. Take a history, perform a clinical examination and undertake standard investigations – for example, electrocardiography, chest X-ray and blood tests – in line with Chronic heart failure (NICE clinical guideline 108).</p> <p>6. In people presenting with new suspected acute heart failure, use a single measurement of serum natriuretic peptides (B-type natriuretic peptide [BNP] or N-terminal pro-B-type natriuretic peptide [NT-proBNP]) and the following thresholds to rule out the diagnosis of heart failure.</p> <ul style="list-style-type: none"> • BNP less than 100 ng/litre • NT-proBNP less than 300 ng/litre. <p>7. In people presenting with new suspected acute heart failure with raised natriuretic peptide levels (see recommendation 6), perform transthoracic Doppler 2D echocardiography to establish the presence or absence of cardiac abnormalities.</p>
	<p>investigations. However, sensitivity analyses found that the results of the model were robust to different assumptions made about physician accuracy, unless physician sensitivity is greater than that 88% (in the context of specialist management).</p> <p>In the base case, where NP testing is assumed to be less specific (but more sensitive) than the physician not using NP testing, the model predicts an increase in the number of echocardiograms performed by NP testing, but this cost is justified by the QALY gains as represented by the low incremental cost effectiveness ratio. In a sensitivity analysis where a different point is taken from the ROC curve and the physician diagnosis is assumed to have greater sensitivity than in the base case, NP testing reduces the number of echocardiograms ordered.</p>
<p>Quality of evidence</p>	<p>At 'rule-out' thresholds for BNP and NT-proBNP, evidence was rated as high quality according to GRADE methodology. However, in some studies (9/45) the assessment of the reference standard was unblinded to the NP results. This would be considered very high risk of bias, but since the majority of evidence for index tests stemmed from high quality studies limitations for risk of bias were not downgraded in line with GRADE methodology.</p>
<p>Other considerations</p>	<p>The GDG discussed the meta-analysis of the diagnostic accuracy of the natriuretic peptides. The evidence was largely of high quality with a number of large-scale studies including patients of a wide age range. The GDG noted that studies excluded patients with a cause of dyspnoea with a clear alternative diagnosis, such as pneumothorax. The GDG also noted that the majority of studies were not conducted in the UK setting. Many were in settings similar to the UK, for example Western Europe. The studies consistently showed high sensitivity. The GDG were confident that the meta-analysis of over 7000 patients was applicable to the UK population (see Appendix J for further details).</p> <p>The GDG commented that both BNP and NT-proBNP had high values of sensitivity and are particularly useful in ruling out acute heart failure (threshold <100 ng/l for BNP and <300ng/l for NT-proBNP). Values above these rule-out levels require further investigation by echocardiography if acute heart failure is clinically suspected, as the diagnostic specificity is modest and variable, and give no indication of underlying cause of the heart failure.</p>

<p>Recommendations</p>	<p>5. Take a history, perform a clinical examination and undertake standard investigations – for example, electrocardiography, chest X-ray and blood tests – in line with Chronic heart failure (NICE clinical guideline 108).</p> <p>6. In people presenting with new suspected acute heart failure, use a single measurement of serum natriuretic peptides (B-type natriuretic peptide [BNP] or N-terminal pro-B-type natriuretic peptide [NT-proBNP]) and the following thresholds to rule out the diagnosis of heart failure.</p> <ul style="list-style-type: none"> • BNP less than 100 ng/litre • NT-proBNP less than 300 ng/litre. <p>7. In people presenting with new suspected acute heart failure with raised natriuretic peptide levels (see recommendation 6), perform transthoracic Doppler 2D echocardiography to establish the presence or absence of cardiac abnormalities.</p>
	<p>The GDG considered that in order to be useful in the diagnosis and management of patients with acute heart failure, the natriuretic peptide test result should be available rapidly (ideally within one hour) and the quality control should be rigorous to ensure consistent results. The natriuretic peptide blood test may be most efficiently processed if included with the usual batch of blood tests sent on initial patient assessment. The GDG discussed that ‘point-of-care’ tests may be subject to technical error if used by inexperienced operators. In a number of studies the ‘point-of-care’ system was actually used in the laboratory by trained technical staff. The GDG commented that repeated testing, particularly in patients with low initial values, would be expensive and unlikely to provide further benefit. This review did not examine the evidence regarding the use of natriuretic peptide testing to guide treatment or evaluate prognosis.</p> <p>The GDG noted that cut-offs may vary for different age groups and it was also acknowledged that, prior use of medication at the impact heart failure (in this context, diuretic in particular) may influence testing. The GDG considered these points, but studies included patients with a very narrow age range and did not divide patients by the type of medication they were receiving and therefore it was not possible to conduct analyses based on these characteristics. Therefore no specific recommendations were drafted.</p> <p>The GDG discussed the different types of natriuretic peptide test and were unable to recommend the use of MRproANP on the basis of this review, as there was no clear diagnostic advantage. The GDG also noted that MRproANP would require additional laboratory equipment in many hospitals as it is not widely available. The relative merits of BNP versus NT-proBNP were discussed; NT-proBNP has a marginally higher sensitivity but with reduced specificity. The GDG did not recommend the use of one natriuretic peptide test over the other, as there was no clear significant difference between them. The current laboratory system in place in each hospital may affect the choice of which natriuretic peptide to use.</p> <p>The GDG noted that other standard investigations performed at diagnosis, such as ECG, chest X-ray and blood tests, were in line with the recommendations already made in the chronic heart failure guideline and the GDG agreed a cross reference should be made to this.</p>

Recommendations	<p>5. Take a history, perform a clinical examination and undertake standard investigations – for example, electrocardiography, chest X-ray and blood tests – in line with Chronic heart failure (NICE clinical guideline 108).</p> <p>6. In people presenting with new suspected acute heart failure, use a single measurement of serum natriuretic peptides (B-type natriuretic peptide [BNP] or N-terminal pro-B-type natriuretic peptide [NT-proBNP]) and the following thresholds to rule out the diagnosis of heart failure.</p> <ul style="list-style-type: none"> • BNP less than 100 ng/litre • NT-proBNP less than 300 ng/litre. <p>7. In people presenting with new suspected acute heart failure with raised natriuretic peptide levels (see recommendation 6), perform transthoracic Doppler 2D echocardiography to establish the presence or absence of cardiac abnormalities.</p>
	The GDG noted that natriuretic peptide levels may be affected by factors such as use of ACEi or diuretics and in people with obesity.

6.2 Echocardiography

The constellation of symptoms and signs suggestive of heart failure are not always sufficient to make a definitive diagnosis and further investigation is commonly required to confirm cardiac dysfunction and to identify reversible structural causes. Echocardiography is the most common and readily available non-invasive test with which to evaluate the structural function of the heart and its role is well established. There is uncertainty however, around the optimal timing of performing echocardiography, particularly in the diagnosis of de novo acute heart failure, and whether this may impact on clinical outcomes.

Review question: In adults with suspected acute heart failure does early echocardiography compared to later echocardiography in addition to standard investigations (using ECG, chest x-ray and blood tests) improve outcome?

For full details see review protocol in Appendix C.

Table 25: PICO characteristics of review question

Population	Adults with suspected (or under investigation for) acute heart failure excluding primary care and community settings
Intervention/s	Early echocardiography
Comparison/s	Later echocardiography (any study that had a time comparison)*
Outcomes	Mortality Major adverse events Length of hospital stay and re-admission rates Quality of Life
Study design	Randomised controlled trials and non-randomised studies will be considered (no particular year or sample size restrictions)

* Definition of 'early' was discussed, but it was agreed not to put an exact timeframe to this since any comparison investigating earlier versus later timing of echocardiography was seen as relevant.

6.2.1 Clinical evidence

We searched for any study type comparing earlier versus later echocardiography in the diagnostic process of acute heart failure, or for studies that addressed timing of echocardiography more generally. The protocol was restricted to hospital settings, and studies set in primary care and in the community were excluded. No relevant studies were identified. See also the study selection flow chart in Appendix D and exclusion list in Appendix K.

6.2.2 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

New cost-effectiveness analysis

New analysis was not prioritised for this area.

6.2.3 Evidence statements

6.2.3.1 Clinical

No relevant evidence was identified.

6.2.3.2 Economic

No relevant economic evaluations were identified.

6.2.4 Recommendations and link to evidence

Recommendations	8. In people presenting with new suspected acute heart failure, consider performing transthoracic Doppler 2D echocardiography within 48 hours of admission to guide early specialist management.
Relative values of different outcomes	No relevant studies were identified which compared outcomes of different timing of echocardiography for patients admitted with suspected acute heart failure.
Trade-off between clinical benefits and harms	In the absence of empirical evidence, the GDG discussed the potential advantages and disadvantages of performing echocardiography earlier in the course of an admission with suspected acute heart failure. Earlier echocardiography allows earlier initiation of effective treatments for left ventricular systolic function. Immediate echocardiography is indicated in patients with haemodynamic instability (e.g. cardiogenic shock) where making specific diagnoses (e.g. pericardial tamponade, myocarditis, endocarditis, acute valve disease), will guide immediate management. The GDG also noted that echocardiography carried out early could lead to the process being rushed and the possibilities that errors in the interpretation of results may increase.
Economic considerations	No economic evidence was identified. Potential economic consequences of early echocardiography were discussed. In current practice, most patients with heart

Recommendations	8. In people presenting with new suspected acute heart failure, consider performing transthoracic Doppler 2D echocardiography within 48 hours of admission to guide early specialist management.
	failure will have echocardiography during the course of an admission, so specifying early investigation will not necessarily increase the amount of echocardiography that needs to be performed, but might increase costs through more use of 'out-of-hours' services. On the other hand, earlier diagnosis might lead to shorter lengths of stay.
Quality of evidence	Not applicable.
Other considerations	<p>Echocardiography is required to identify the cardiac abnormality that is underlying the clinical syndrome of heart failure. In particular, it is the investigation with which left ventricular systolic dysfunction is diagnosed. Therefore, the GDG considered that earlier echocardiography would allow earlier diagnosis and prompt treatment.</p> <p>The GDG discussed how urgently echocardiography needed to be performed. For some patients, where there is suspicion of an acute life threatening structural cardiac abnormality, the investigation should be performed immediately. For other patients (the majority of patients with suspected acute heart failure), the GDG considered that echocardiography should be performed within 48 hours of admission since this would allow appropriate management, both in terms of pharmacological treatment, site of care, and supervision from a heart failure team.</p> <p>Some patients would not require echocardiography. The evidence from the review of natriuretic peptides in acute heart failure demonstrates that natriuretic peptides have high sensitivity, thus, they are a useful 'rule-out' test if normal. Repeat echocardiography would not usually be needed for patients where the underlying diagnosis is already known, and there is little clinical suspicion of a change in pathology.</p>

6.3 Invasive monitoring

Patients with acute heart failure have high morbidity and mortality, and need close monitoring during their hospital admission. The majority of physical parameters can be monitored with non-invasive methods in addition to intermittent blood tests to evaluate the function of key organs (e.g. kidneys). However, some patients may be critically unwell and so more invasive methods may be used to provide constantly updated measures. An arterial line may be used to measure heart rate, blood pressure and to allow regular arterial blood sampling. A central venous line may be used to measure the central venous pressure, allow administration of multiple medications and regular venous blood sampling. A pulmonary artery catheter may be used to estimate pressures in the left side of the heart and to estimate different measures of cardiac function. Often, in patients admitted to a critical care area, invasive monitoring is mandated in order to provide instant accurate measures to allow rapid changes in treatment and to allow frequent blood sampling. However, the evidence for the benefit of routinely using invasive monitoring methods in some patients is unclear and particularly the use of pulmonary artery catheterisation.

Review question: Is the addition of invasive monitoring more clinically/cost-effective over and above non-invasive monitoring to improve outcome?

Table 26 summarises the main aspects of the protocol. For full details refer to Appendix C.

Table 26: PICO characteristics of review question

Population	Adults with acute heart failure
Intervention/s	Invasive monitoring with arterial lines, central venous pressure lines or pulmonary artery catheters
Comparison/s	All those who are not invasively monitored including those with non-invasive monitoring
Outcomes	Mortality Major cardiovascular events Length of hospital stay Re-admission rates Number of patients proceeding to invasive ventilation Measures of renal function (e.g. eGFR or serum creatinine) Quality of life (as well as reported anxiety and pain) Adverse events (cardiovascular)
Study design	Systematic reviews, randomised controlled trials and observational studies

Studies were excluded if they were observational and did not conduct multivariable adjustment for baseline differences between groups.

6.3.1 Clinical evidence

We searched for studies comparing the effectiveness of invasive monitoring (pulmonary artery catheter, arterial lines or central venous pressure lines) to standard medical care including non-invasive monitoring. Two randomised controlled trials (ESCAPE, 2005 trial with the published protocol^{66,200} and Harvey et al., 2005 PAC-Man trial¹⁰⁰) and two registry studies (Sotomi et al., 2012 ATTEND registry with the published registry description^{196,206} and Zion et al., 1990 SPRINT registry²³⁴) were included in this review.

All studies compared pulmonary artery catheter (PAC) to standard clinical care or clinical assessment. The PAC-Man trial investigated PAC use in the intensive care unit population, but presented findings relating to a pre-specified subgroup of patients with decompensated heart failure. The SPRINT registry studied patients with acute myocardial infarction and included an analysis of a subgroup of patients with cardiogenic shock. Patients with acute decompensated heart failure were the focus of the ATTEND registry, however, the study is currently available only as a conference abstract. The ESCAPE trial included patients with symptomatic congestive heart failure. The related ESCAPE registry^{12,13} was excluded as it was restricted to patients with PAC only and did not include a comparison group – see exclusion list in Appendix K.

For brief summaries of included studies see **Error! Reference source not found.** Table 27 with further details provided in study evidence tables in Appendix G. Evidence for protocol outcomes from these are summarised in the clinical GRADE evidence profile below (Table 28). See also the study selection flow chart in Appendix D and forest plots in Appendix I.

Summary of included studies

The characteristics of the included studies are summarised in the table below. Details of the studies can be found in Appendix G.

Table 27: Summary of studies included in the review

Study	Intervention / Comparison	Population	Protocol outcomes	Comments
The ESCAPE Investigators	PAC vs. Clinical	N=433 Patients with severe symptomatic	Mortality (and days alive out of	Quality of life data only available in a figure and

Study	Intervention / Comparison	Population	Protocol outcomes	Comments
2005 (ESCAPE trial) ⁶⁶	assessment	heart failure despite recommended therapies currently admitted for NYHA class IV heart failure symptoms. Country: USA	hospital); initial length of stay; adverse events; and health-related quality of life measured by the Minnesota Living with Heart Failure questionnaire	could not be clearly extracted. Low rate of hospital mortality. No definition of precise strategy in response to the hemodynamic information. Use of medication varied between sites.
Harvey et al., 2005 (PAC-Man trial) ¹⁰⁰	PAC vs. control (including less invasive cardiac output monitoring – i.e. indirect comparator)	Patients from intensive care units. A pre-specified subgroup (N=111) of patients with decompensated heart failure was separately analysed. Country: UK	Hospital mortality	Hospital mortality rate substantially higher than the rate in the ESCAPE trial.
Sotomi et al., 2012 (ATTEND registry) ²⁰⁶	PAC vs. control	Patients with acute decompensated heart failure (N=4796) Country: Japan	Hospital mortality with subgroup analyses (one of which was according to NYHA class)	Currently only available as a conference abstract. Used a propensity score analysis to account for differences between groups.
Zion et al., 1990 (SPRINT registry) ²³⁴	PAC vs. control	Hospitalised patients with acute myocardial infarction including a subgroup of patients with cardiogenic shock (N=581) Country: Israel	Hospital mortality	Approximately 40% of patients randomised to receive either nifedipine or placebo. This group was combined with a registry group. Mortality in the cardiogenic shock group ≥90%.

Table 28: Clinical evidence profile: Pulmonary Artery Catheter (PAC) vs. Control. Where cells in the control group are subdivided the second percentage refers to the median control event rate which is used to calculate the absolute effect.

Quality assessment							Summary of Finding				Quality	Importance
No of studies Follow-up time	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Invasive monitoring Number of event / Total N (%) or Mean (SD)	Control Number of event / Total N (%) or Mean (SD)	Effect			
									Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect or Mean Difference (95% CI)		
Mortality – In-hospital plus 30 days ^{66,100}												
2	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious ^(a)	None	49/264 (18.5%)	46/268 (17.2%) 33.8%	RR 1.08 (0.82 to 1.43)	27 more per 1000 (from 61 fewer to 145 more)	MODERATE	CRITICAL
Mortality - Follow-up 180 days ⁶⁶												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious ^(a)	None	43/209 (20.6%)	38/212 (17.9%)	RR 1.15 (0.78 to 1.7)	27 more per 1000 (from 39 fewer to 125 more)	MODERATE	CRITICAL
Mortality - PAC related deaths ⁶⁶												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	0/209 (0%)	0/212 (0%)	Not pooled	Not pooled	HIGH	CRITICAL
Days alive and out of hospital ⁶⁶												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	Mean 133 (sd not reported)	Mean 135 (sd not reported)	HR 1.00 (0.82 to 1.21)	0 fewer per 1000 (from 30 fewer to 34 more)	HIGH	CRITICAL
In-hospital mortality - ATTEND registry - Total ^{206,207}												
1	Observational	No serious	No serious inconsistency	No serious indirectness	Serious ^(a)	None	n = 806	n = 3990	OR 0.64 (0.37 to	450 fewer per 1000 (from	VERY LOW	CRITICAL

Quality assessment							Summary of Finding				Quality	Importance
No of studies Follow-up time	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Invasive monitoring Number of event / Total N (%) or Mean (SD)	Control Number of event / Total N (%) or Mean (SD)	Effect			
									Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect or Mean Difference (95% CI)		
	study	risk of bias							1.11)	990 fewer to 100 more) ^(b)		
In-hospital mortality - ATTEND registry - NYHA Class IV ^{206,207}												
1	Observational study	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious ^(a)	None	N/A ^(c)	N/A ^(c)	OR 0.43 (0.20 to 0.92)	N/A ^(c)	VERY LOW	CRITICAL
In-hospital mortality – SPRINT registry – Cardiogenic shock ²³⁴												
1	Observational study	Serious risk of bias ^(d)	No serious inconsistency	No serious indirectness	No serious imprecision	None	139/154 (90.3%)	388/427 (90.9%)	OR 0.99 (0.76 to 1.30)	1 fewer per 1000 (from 25 fewer to 20 more)	VERY LOW	CRITICAL
Health-related quality of life: Minnesota Living with Heart Failure questionnaire score (0 to 105; better indicated by lower values) – Follow-up 1 and 6 months ⁶⁶												
1	RCT	Very serious risk of bias ^(e)	No serious inconsistency	No serious indirectness	Not possible to assess ^(e)	None	N/A ^(e)	N/A ^(e)	N/A ^(e)	Improvement was greater by 5 points at 1 month (established MID) – at 6 months scores were similar	VERY LOW	CRITICAL
Health-related quality of life: Time trade-off score - Follow-up 1 and 6 months ⁶⁶												
1	RCT	Very serious risk of bias ^(e)	No serious inconsistency	No serious indirectness	Not possible to assess ^(e)	None	N/A ^(e)	N/A ^(e)	N/A ^(e)	The improvement in the PAC group was more than twice as great	VERY LOW	CRITICAL

Quality assessment							Summary of Finding				Quality	Importance
No of studies Follow-up time	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Invasive monitoring Number of event / Total N (%) or Mean (SD)	Control Number of event / Total N (%) or Mean (SD)	Effect			
									Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect or Mean Difference (95% CI)		
										at any point.		
Severe adverse events - Cardiogenic shock ⁶⁶												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	6/209 (2.9%)	2/212 (0.94%)	RR 3.04 (0.62 to 14.91)	19 more per 1000 (from 4 fewer to 131 more)	LOW	IMPORTANT
Severe adverse events - Ischemia/angina ⁶⁶												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	9/209 (4.3%)	4/212 (1.9%)	RR 2.28 (0.71 to 7.3)	24 more per 1000 (from 5 fewer to 119 more)	LOW	IMPORTANT
Severe adverse events - Myocardial Infarction ⁶⁶												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	0/209 (0%)	1/212 (0.47%)	Peto OR 0.14 (0.00 to 6.92)	0 fewer per 1000 (from 20 fewer to 10 more) ^(f)	LOW	CRITICAL
Severe adverse events - Stroke or transient ischaemic attack ⁶⁶												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	1/209 (0.48%)	0/212 (0%)	Peto OR 7.50 (0.15 to 377.81)	0 more per 1000 (from 10 fewer to 20 more) ^(f)	LOW	CRITICAL
Severe adverse events - Cardiac arrest ⁶⁶												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	9/209 (4.3%)	5/212 (2.4%)	RR 1.83 (0.62 to 5.36)	20 more per 1000 (from 9 fewer to 103 more)	LOW	CRITICAL
Severe adverse events - Infection ⁶⁶												

Quality assessment							Summary of Finding				Quality	Importance
No of studies Follow-up time	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Invasive monitoring Number of event / Total N (%) or Mean (SD)	Control Number of event / Total N (%) or Mean (SD)	Effect			
									Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect or Mean Difference (95% CI)		
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious ^(a)	None	27/209 (12.9%)	20/212 (9.4%)	RR 1.37 (0.79 to 2.36)	35 more per 1000 (from 20 fewer to 128 more)	MODERATE	IMPORTANT
Severe adverse events - Patients with ≥ 1 adverse event ⁶⁶												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious ^(a)	None	47/209 (22.5%)	25/212 (11.8%)	RR 1.91 (1.22 to 2.98)	107 more per 1000 (from 26 more to 233 more)	MODERATE	CRITICAL
Serum creatinine - Change from baseline (better indicated by lower values) ⁶⁶												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	n = 215 Mean 0 (0.4)	n = 218 Mean 0.1 (0.8)	N/A	MD 0.1 lower (0.22 lower to 0.02 higher)	HIGH	IMPORTANT
Days of initial hospitalisation ⁶⁶												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious ^(a)	None	Mean 8.7	Mean 8.3	HR 1.04 (0.86 to 1.27)	N/A ^(g)	MODERATE	IMPORTANT

(a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed one of the default MIDs (i.e. ranging from a noticeable effect to no effect). Outcomes were downgraded by two increments if 95% CI of the effect crossed both default MIDs (i.e. ranging from a noticeable improvement to a noticeable harm). Default MIDs are set at RRs / HRs of 0.75 and 1.25 for dichotomous outcomes and at 0.5 of the median control group standard deviation either side of the null line for continuous outcomes.

(b) It was unclear how many events there were in each group (data presented in a conference abstract). The absolute effect was calculated from the risk difference.

(c) Neither the number of events nor the total number of patients in this subgroup were provided (data presented in a conference abstract). The absolute effect could not be calculated.

(d) About half of the participants were part of a RCT investigating the effectiveness of nifedipine.

(e) Results were presented in a figure only and only a brief narrative was provided for the result at 1 month follow-up.

(f) Due to the low event rate Peto OR was calculated and the risk difference was used for the absolute effect.

(g) It was unclear how many events led to this Hazard Ratio. Hence absolute numbers could not be calculated.

6.3.2 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

New cost-effectiveness analysis

New analysis was not prioritised for this area.

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs were provided to aid consideration of cost-effectiveness. These are given in Appendix N.

6.3.3 Evidence statements

6.3.3.1 Clinical

Mortality from RCT data

Moderate quality evidence from two RCTs (n = 532) showed similar hospital mortality rates regardless of whether patients were invasively monitored with pulmonary artery catheter (PAC) or received standard medical assessment. One of the trials (n = 421) also provided moderate quality evidence for similar rates of mortality at 180-day follow-up, evidence of similar length of survival out of hospital (high quality evidence) and reported no PAC-related mortality in either group (high quality evidence).

Mortality from observational studies

Very low quality evidence from two observational registry studies (n = 5377) contributed to the outcome mortality. One of the registry studies (n = 4796) provided very low quality evidence for no clear group difference (between PAC and standard care) in overall hospital mortality, but included a subgroup of patients with NYHA Class IV in which PAC was effectively reducing mortality (very low quality evidence). Very low quality evidence for a subgroup of patients with cardiogenic shock (n = 581) indicated similar high rates of hospital mortality for patients regardless of whether they were invasively monitored or not.

Health-related quality of life as measured by the Minnesota Living with Heart Failure questionnaire and the time trade-off score.

Very low quality evidence from one RCT comprising 421 participants with congestive heart failure suggested that the difference in health-related quality of life at one month favouring PAC was clinically important. This improvement over and above medical assessment was no longer apparent at 6 month (very low quality evidence) from this questionnaire. However, the same RCT provided very low quality evidence suggesting that the patients awarded relatively more value to their lives when monitored by PAC at both 1 month and, in contrast to the Minnesota questionnaire result, also at 6 months follow-up compared to medical assessment.

Severe adverse events

Moderate quality evidence from one RCT (n = 421) showed that invasive monitoring was associated with a higher proportion of patients who experienced at least 1 adverse event.

Overall, only one event of myocardial infarction and stroke or transient ischaemic attack was reported, thus, it is unclear how this could be related to treatment.

Cardiac arrests, cardiogenic shock and ischaemia / angina were more common in the PAC group, but it was unclear whether there was a clear clinical difference (low quality evidence) between the two groups. Rates of infections were similar regardless of whether patients were invasively monitored or not (moderate quality evidence).

Serum creatinine (mg/dL) – change from baseline

High quality evidence from one RCT (n = 433) showed no clear difference in serum creatinine between patients receiving PAC and those receiving standard medical treatment.

Length of initial hospital stay

Moderate quality evidence from one RCT (n = 433) indicated no clear difference in length of hospital stay between patients monitored by PAC and those receiving standard medical assessments.

Other protocol outcomes such as re-admission rates and number of patients proceeding to invasive ventilation were not reported.

6.3.3.2 Economic

No relevant economic evaluations were included.

6.3.4 Recommendations and link to evidence

Recommendations	9. Do not routinely offer pulmonary artery catheterisation to people with acute heart failure.
Relative values of different outcomes	The GDG regarded mortality the most important outcome, but also considered health-related quality of life, severe adverse events and serum creatinine. Only one study, set in the USA, reported length of stay. Given healthcare differences and shorter lengths of stay in the USA, this was judged less applicable to UK clinical practice. A measure of quality of life, time-trade-off score, was presented and discussed. This score is not commonly used in heart failure studies, did not correlate with the more validated Minnesota Living with Heart Failure (MLHF) questionnaire and is conceptually a difficult measure for patients to assess. For these reasons, it was given less weight than the MLHF score in the discussion.
Trade-off between clinical benefits and harms	Routine pulmonary artery catheterisation did not reduce mortality in the randomised control trials but was associated with a higher rate of severe adverse events. Pulmonary artery catheterisation was associated with better health related quality of life at 1 month follow-up, but results for longer term quality of life were inconsistent, with differences maintained in the time-trade-off approach but not in the MLHF score.
Economic considerations	No relevant economic evaluations were identified for the specified population. Pulmonary artery catheterisation (PAC) did not reduce mortality or length of stay, and there was no clear difference in health-related quality of life beyond 1 month. For these reasons the GDG judged that PAC is not likely to be cost-effective.
Quality of evidence	Evidence was identified for pulmonary artery catheterisation (PAC), but not for

Recommendations	9. Do not routinely offer pulmonary artery catheterisation to people with acute heart failure.
	<p>central venous pressure or arterial lines.</p> <p>Two randomised controlled trials conducted in different settings were included, ESCAPE in the emergency department and PACMAN in intensive care. The evidence from these trials was rated as moderate to high, but both studies were unblinded, so the health-related quality of life measures were at high risk of bias. Two observational studies using registry data were considered to be of very low quality, so the results were not considered by the GDG in their decision making.</p>
Other considerations	<p>Pulmonary artery catheters are used infrequently in current practice. The GDG recognised that certain patients who might have benefited from invasive monitoring (e.g. candidates for mechanical circulatory support or cardiac transplantation) were not included in these trials. No relevant studies were identified for CVP or arterial lines.</p>

7 Initial pharmacological treatment

7.1 General considerations

In several of the GDG discussions on the specific treatments, the GDG noted that the majority of patients in the UK presenting with acute heart failure will have an acute decompensation of chronic heart failure. As such, the GDG raised concerns that it is important that any treatment should only be commenced with the consent of the person or, in the absence of capacity, if this is in accordance with any advance care plan the patient may have. This should be undertaken in line with the NICE guideline on Patient experience in adult NHS services (CG138). Thus, the GDG made the following cross referral recommendation concerning any treatments.

10. For guidance on patient consent and capacity follow recommendations 1.2.12 and 1.2.13 in Patient experience in adult NHS services (NICE Clinical Guideline 138).

7.2 Opiates

People presenting to hospital with acute heart failure may often be in considerable distress due to acute pulmonary oedema and respiratory failure. It is common practice to administer an opioid medication in these circumstances, although whether or not this improves clinical outcomes is unclear. This review aims to evaluate the use of opioid medications in acute heart failure.

Review question: In patients with acute heart failure are opiates as an adjunct to other first line pharmacological therapies more clinically and cost effective compared to other pharmacological treatments alone?

For full details see review protocol in Appendix C.

Table 29: PICO characteristics of the review question

Population	Adults with acute heart failure
Intervention(s)	Morphine or diamorphine
Comparison(s)	Standard medical care or placebo
Outcomes	<ul style="list-style-type: none"> • Mortality • Major cardiovascular events • Length of hospital stay and re-admission rates • Number of patients proceeding to invasive ventilation • Measures of dyspnoea (breathing rate or breathlessness scales) • Quality of life (as well as reported anxiety and pain) • Adverse events (particularly respiratory arrest and nausea)
Study design	Randomised controlled trials and observational studies

7.2.1 Clinical evidence

We searched for randomised controlled trials (RCTs) or observational studies. Five observational studies were included in this review^{94,102,104,174,193}. Evidence from these studies are summarised in Table 2. See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and exclusion list in Appendix K.

The observational studies that are included in this review generally presented univariate and multivariable results. This means that the former is an unadjusted effect and the latter takes into account baseline characteristics that may influence this effect.

The GRADE table and the forest plots for this review have been adapted to highlight the different variables that the studies accounted for and both unadjusted and adjusted effects are presented separately.

No evidence was found relating to major cardiovascular events, length of hospital stay or quality of life.

Table 30: Summary of included studies

Study	Study Design	Population	Intervention / Comparison	Outcomes
Gray 2010 ^{93,94}	Retrospective cohort n=1052	Clinical diagnosis of acute cardiogenic pulmonary oedema with pH <7.35; population from 3CPO trial in the UK	Those given IV opiates (n=541) compared with those not given IV opiates (n= 511)	<ul style="list-style-type: none"> • 7-day mortality • Change from baseline in dyspnoea
Hoffman 1987 ¹⁰²	Prospective non-randomised trial n=57	Presumed pre-hospital pulmonary oedema at UCLA hospital in the USA	Group A: Furosemide / Nitroglycerine (control group) Group B: Furosemide / IV Morphine Group C: Furosemide / Nitroglycerine / IV Morphine Group D: Nitroglycerine / IV Morphine	<ul style="list-style-type: none"> • Number of patients subjectively reporting improvement in symptoms • Number of patients objectively improved • Number of patients with possible serious adverse events (SAEs) within 1 hour
Iakobishvili 2011 ¹⁰⁴	Retrospective cohort n=2336	Clinical diagnosis of acute decompensated heart failure; population from the Heart Failure Survey in Israel (HFSIS)	Those given IV morphine (n=218) compared to those not given IV morphine (n=2118)	<ul style="list-style-type: none"> • In-hospital mortality • 30-day mortality
Peacock 2008 ¹⁷⁴	Retrospective cohort n=147, 362	Discharge diagnosis of acute decompensated heart failure from ADHERE registry in the USA	Those given IV morphine(n=20,782) compared with those not given IV morphine (n=126,580)	<ul style="list-style-type: none"> • Mortality • Number of patients progressing to invasive ventilation • Number of patients admitted to Intensive

Study	Study Design	Population	Intervention / Comparison	Outcomes
				care unit (ICU)
Sacchetti 1999 ¹⁹³	Retrospective cohort n=181	Acute pulmonary oedema from 'Our Lady of Lourdes Hospital' in the USA	Those given IV morphine(n=88) compared with those not given IV morphine (n=93)	<ul style="list-style-type: none"> • Number of patients progressing to invasive ventilation • Number of patients admitted to ICU

Mortality Peacock 2008 – in hospital

Table 31: Clinical evidence profile: Opiates in Acute Heart Failure

Quality assessment							No of patients		Effect	Quality	Importance
Study	Design	Risk of bias	Confounders adjusted for	Inconsistency	Indirectness	Imprecision (for adjusted results if reported)	Morphine	No Morphine	OR [95%CI]		
Mortality Iakobishvili (2011) – in hospital											
Iakobishvili 2011	Retrospective cohort	Very serious ^(a)	Serum urea >86 mg/dl; systolic blood pressure >115 mg Hg; serum creatinine >2.75 mg/dl; serum glucose; white blood cell count, Killip class on admission, the use of intravenous inotropes and vasodilators; dyslipidaemia and acute coronary syndrome	N/A	No serious indirectness	Serious imprecision ^(a)	n=218 n=218 n=218	n=2118 n=218 (propensity) n=2118 (multivariable analysis)	In hospital adjusted: 2.0 (1.1-3.5) In hospital with propensity score: 1.2 (0.6-2.5) 30 day adjusted: 1.5 (0.9-2.3)	VERY LOW	CRITICAL
Peacock 2008	Retrospective cohort	Serious risk of bias ^(a)	BUN; systolic BP; age; creatinine; dyspnoea at rest; chronic dialysis; heart rate; inotrope use and vasodilator use +/- serum troponin level	N/A	No serious indirectness	No serious imprecision	n=20,782 n=49,174	n=126,580	Unadjusted: 6.08 (5.76 to 6.41) Adjusted: 5.27 (4.96 to 5.6) Adjusted + Troponin: 4.84 (4.52 to 5.18) EF <40% Adjusted + Troponin:	VERY LOW	CRITICAL

									4.13 [3.73, 4.57]		
							n=48,969		EF >40% Adjusted + Troponin: 5.50 [4.88, 6.20]		
							n=15,014	n=97,841	Nil invasive ventilation Adjusted + Troponin: 5.75 [5.31, 6.23]		
							n=4,318	n=25,958	De novo patients Adjusted: 4.21 [3.59, 4.94]		
Mortality Gray 2010 – 7 day											
Gray 2010	Retrospective cohort	Very serious ^(a)	Age; systolic blood pressure; ability to obey commands	N/A	No serious indirectness	Serious ^(b)	n=541	n=511	7 day unadjusted: 1.2 (0.79 to 1.81) 7 day adjusted: 1.27 (0.8 to 2.02)	VERY LOW	CRITICAL
Number of patients progressing to invasive ventilation – Peacock 2008											
Peacock 2008	Retrospective cohort	Very serious ^(a)	BUN; systolic BP; age; creatinine; dyspnoea at rest; chronic dialysis; heart rate; inotrope use and vasodilator use +/- serum troponin	N/A	No serious indirectness	No serious imprecision	3200/20,782 (15.4%)	3544/126,580 (2.8%)	6.32 [6.01 to 6.64]	VERY LOW	CRITICAL

Number of patients admitted to ICU - Peacock 2008											
Peacock 2008	Retrospective cohort	Very serious ^(a)	BUN; systolic BP; age; creatinine; dyspnoea at rest; chronic dialysis; heart rate; inotrope use and vasodilator use +/- serum troponin level	N/A	No serious indirectness	No serious imprecision	8043/20,782 (38.7%)	18228/126,580 (14.4%)	3.75 [3.63, 3.88]	VERY LOW	IMPORTANT
			level								
Sacchetti 1999	Retrospective cohort	Very serious ^(a)	Age; MI; use of diuretic; captopril or nitroglycerine.	N/A	No serious indirectness	No serious imprecision	n=88	n=93	5.04 (2.31 - 11.76)	VERY LOW	CRITICAL
Number of patients admitted to ICU – Sacchetti 1999											
Sacchetti 1999	Retrospective cohort	Very serious ^(a)	Age; MI; use of diuretic; captopril or nitroglycerine.	N/A	No serious indirectness	No serious imprecision	n=88	n=93	3.08 (1.54 to 6.3)	VERY LOW	IMPORTANT
Number of patients subjectively reporting an improvement in symptoms											
Hoffman 1987	Prospective non-randomised trial	Very serious ^(a)	N/A	N/A	Serious ^(c)	No serious imprecision	22/42 (52.4%)	12/15 (80%)	All groups with morphine versus no morphine: RR 0.65 [0.45, 0.96]	VERY LOW	IMPORTANT
							8/15 (53.3%)	12/15 (80%)	Furosemide/ Nitroglycerine/ Morphine versus Furosemide/ Nitroglycerine: RR 0.67 [0.39, 1.14]		

Number of patients objectively improved											
Hoffman 1987	Prospective non-randomised trial	Very serious ^(a)	N/A	N/A	Serious ^(c)	N/A	20/42 (47.62%)	12/15 (80%)	All groups with morphine versus no morphine: RR 0.60 [0.40, 0.89]	VERY LOW	IMPORTANT
							9/15 (60%)	12/15 (80%)	Furosemide/ Nitroglycerine/ Morphine versus Furosemide/ Nitroglycerine: RR 0.75 [0.46, 1.22]		
Change in dyspnoea at 1 hour											
Gray 2010	Retrospective cohort	Very serious ^(a)	Baseline breathlessness; valvular heart disease, normal GCS verbal score; hypertension; baseline O2 saturation	N/A	No serious indirectness	N/A	n=541	n=511	Unadjusted: -0.1 (-0.6 to 0.4) Adjusted: 0.0009 (-0.025 to 0.043)	VERY LOW	IMPORTANT
Number of patients with possible SAEs at 1 hour											
Hoffman 1987	Prospective non-randomised trial	Very serious	N/A	N/A	Serious ^(c)	N/A	10/42 (23.8%)	0/15 (0%)	All groups with morphine versus no morphine: RR 7.81 [0.49, 125.75]	VERY LOW	IMPORTANT
							2/15 (13.3%)	0/15 (0%)	Furosemide/ Nitroglycerin		

									e/ Morphine versus Furosemide/ Nitroglycerin e: RR 5.00 [0.26, 96.13]		
--	--	--	--	--	--	--	--	--	--	--	--

- (a) *Observational studies begin at LOW GRADE rating. Outcomes were further downgraded by one increment if the number of serious methodological limitations across was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.*
- (b) *Outcomes were downgraded by one increment if the upper or lower 95% CI of the effects crossed (when unadjusted or adjusted were reported the adjusted effect was used to assess imprecision) the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID. Default MIDs were set at RRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the control group median standard deviation either side of the null line for continuous variables*
- (c) *Outcomes were downgraded as the study population contained 22% of cases which were not acute heart failure on further diagnosis.*

7.2.2 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

New cost-effectiveness analysis

New analysis was not prioritised for this area.

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs were provided to aid consideration of cost-effectiveness. These are given in Appendix N.

7.2.3 Evidence statements

7.2.3.1 Clinical

Mortality

Three studies^{94,104,174} provided mortality rates associated with opiates. The quality of evidence was very low. One study (Iakobishvili, 2011; N=2,336) suggested that once accounting for differences in patient characteristics, mortality rates were similar in patients who received morphine compared to those who did not. Another study (Peacock 2008; N=126,580 hospital admissions) showed that morphine was associated with a clearly higher mortality rate even when differences in patient characteristics were accounted for. In a third study (Gray 2010; N=1052) no clear differences were found in the mortality rate.

Number of patients proceeding to invasive ventilation

Very low quality evidence from two studies^{174,193} showed that the proportion of patients proceeding to invasive ventilation was clearly higher in those who received morphine compared to those who did not.

Number of patients admitted to Intensive Care Unit

Very low quality evidence from two studies^{174,193} showed that the proportion of patients admitted to ICU was clearly higher in those receiving morphine compared to those who did not.

Number of patients subjectively reporting an improvement in symptoms

Very low quality evidence from one study¹⁰² suggested that the proportion of patients who subjectively reported improvements was higher in those who did not receive morphine compared to those who did.

This effect decreased when morphine was given as an adjunct to furosemide and nitroglycerine compared to patients who were given furosemide and nitroglycerine without additional morphine. However, it was unclear whether there was a clear advantage associated with not receiving morphine.

Number of patients objectively improved

Very low quality evidence from one study¹⁰² suggested that the proportion of objectively reported improvements was higher in those who did not receive morphine compared to those who did.

This effect decreased when morphine was given as an adjunct to furosemide and nitroglycerine compared to patients who were given furosemide and nitroglycerine without additional morphine. However, it was unclear whether there was a clear advantage associated with not receiving morphine.

Change in dyspnoea

Very low quality evidence from one study^{93,94} showed no clear difference in the rate of dyspnoea between the two intervention groups even when accounting for differences in patient characteristics.

Number of patients with possible serious adverse events

Very low quality evidence from one study¹⁰² showed that possible serious adverse events were only reported in the morphine group. Fewer adverse events were reported when morphine was given as an adjunct to furosemide and nitroglycerine. However, it is difficult to draw a clear conclusion regarding harmful effects of morphine due to the uncertainty around the size of the effect.

7.2.3.2 Economic

No relevant economic evaluations were identified.

7.2.4 Recommendations and link to evidence

Recommendations	11. Do not routinely offer opiates to people with acute heart failure.
Relative values of different outcomes	The GDG considered mortality the most important outcome, but looked at other outcomes: the number of patients ventilated and/or being moved to Intensive Care, and subjective or objective symptomatic improvement.
Trade-off between clinical benefits and harms	The evidence from the quasi-randomised trial and observational studies of opiates suggest no evidence of benefit, but some evidence of harm.
Economic considerations	No economic studies were identified for inclusion in the review. The use of morphine represents additional cost, but this is not large.
Quality of evidence	All evidence was rated as very low according to the GRADE criteria and there was no randomised controlled trial evidence. The GDG expressed concern about making recommendations based solely on observational data which are not pooled due to the different confounding factors that are adjusted in multivariable analyses in the included studies. Opiates are most likely to be used in those with the poorest long-term outcomes. While the studies attempted to adjust for this, the GDG considered that there was considerable risk of residual confounding.
Other considerations	<p>In current practice, opiates are used to ease distress. The patient members noted that patient's anxiety and stress are high during an episode of acute heart failure and this may justify the use of morphine for its anxiolytic effect. However, the limited evidence suggested that the patient may subjectively feel worse following opiate administration. The GDG regarded the key to managing patient's distress lies in rapid diagnosis, initiation of treatment and a safe, reassuring environment. Although routine use is not appropriate a clinician may choose to offer opiates on an individual patient basis where appropriate, after consideration of the potential risks and benefits.</p> <p>The GDG noted that no randomised controlled trial comparing morphine with placebo has been carried out, but acknowledged that ethical concerns would make such a trial unfeasible.</p>

Recommendations	11. Do not routinely offer opiates to people with acute heart failure.

7.3 Diuretic administration strategies

The majority of people presenting with acute heart failure have varying degrees of either peripheral and/or pulmonary oedema. The treatment of peripheral and pulmonary oedema may be life saving as well as significantly improving symptoms and diuretic medications are commonly administered to increase urine output and resolve oedema. In the acute setting, the need for rapid onset of action and uniformity of dose response means that diuretics are most commonly administered by an intravenous route. However, there is significant variation in practice as to the doses used and whether continuous infusions or intermittent boluses are given. This review evaluates the administration strategy of diuretic medications in acute heart failure.

Review question: In patients with acute heart failure which diuretic administration strategy is the most clinically/cost-effective to improve outcome?

For full details see review protocol in Appendix C.

Table 32: PICO characteristics of review question

Population	Adults with acute heart failure
Intervention/s	<p>One of - using one mode of administration:</p> <ul style="list-style-type: none"> ○ Furosemide (Oral, IV Bolus or IV infusion) ○ Bumetanide (Oral, IV Bolus or IV infusion) ○ Torasemide (Oral, IV Bolus or IV infusion) ○ Amiloride (Oral only) ○ Bendroflumethiazide (Oral only) ○ Metolazone (Oral only) ○ Hydrochlorothiazide (Oral only) ○ Indapamide (Oral only) <p>Plus any IV strategy using diuretic plus adjunctive hypertonic saline solution (HSS)</p>
Comparison/s	Any of the interventions listed above using a different mode of administration
Outcomes	<ul style="list-style-type: none"> • Mortality • Dyspnoea • Urine Output • Weight Loss • Length of hospital stay and re-admission rates • Quality of life • Serum creatinine level (or other measure of renal function for example eGFR) • Adverse events (particularly renal adverse events and ototoxicity)
Study design	<p>Systematic reviews</p> <p>Randomised control trials</p>

7.3.1 Clinical evidence

Ten studies were included in the review^{2,14,59,73,117,122,170,172,177,213}. One Cochrane review was identified (Salvador et al., 2005¹⁹⁴) that partially answered this review question. This was cross-checked and all appropriate studies were included. Another systematic review was identified in the search update (Wu et al., 2014^{229,230}). This systematic review was also cross-checked for references. All references satisfying the protocol criteria were included in our original meta-analysis.

No relevant clinical studies comparing oral with intravenous administration strategies were identified.

Four crossover trials are included in the review^{2,59,117,177}. Owing to the paucity of evidence in this area the GDG opted to include these studies and highlight the limitations of this trial design in the quality assessment. Studies in which the population was solely chronic heart failure patients were excluded from the review at the outset.

There was evidence for three diuretic administration strategy comparisons (a summary of study characteristics indicating the relevant comparison is provided in Table 33 below):

- Bolus IV furosemide vs. continuous infusion – 7 RCTs (3 parallel and 4 crossover)^{2,14,59,73,117,177,213}
- Bolus IV furosemide vs. IV furosemide infusion in combination with hypertonic saline solution (HSS) – 2 parallel RCTs^{122,173}
- Continuous IV furosemide vs. IV furosemide infusion in combination with hypertonic saline solution (HSS) – 1 parallel RCT¹⁷⁰

The evidence review is divided accordingly in the GRADE evidence profiles in Table 34, Table 35 and Table 36. See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and lists of excluded studies in Appendix K.

Table 33: Summary of studies included in the review

Study	Intervention / Comparison	Population	Outcomes	Comments
Aaser 1997 ² n=8	Twice daily IV bolus furosemide vs. continuous infusion	Severe chronic heart failure (CHF) who were eligible for heart transplantation	Urine output Withdrawals due to Adverse Events (AEs)	48 hours; Crossover
Allen 2010 ^{12,14} n=41	Twice daily IV bolus furosemide vs. continuous infusion	Admitted to hospital with a diagnosis of acute decompensated heart failure (ADHF)	Mortality Length of hospital stay Weight loss Urine output Renal function Withdrawals due to AEs	72 hours; Parallel
Dormans 1996 ⁵⁹ n=11	Single IV bolus furosemide vs. continuous infusion	Severe HF of differing aetiologies with New York Heart Association (NYHA) Class III/IV. 9 patients were compensated; 11 decompensated	Urine output Renal Function Number of AEs Withdrawals due to AEs	72 hours; Crossover
Felker 2011 ^{73,75} n=307	Low dose bolus IV furosemide (LDB)	ADHF	Mortality Rehospitalisation	72 hours; Parallel

Study	Intervention / Comparison	Population	Outcomes	Comments
	vs. low dose IV continuous (LDC) vs. high dose IV bolus (HDB) vs. high dose IV continuous (HDC)		Number of patients visiting the Emergency Department (ED) Length of hospital stay Weight loss Dyspnoea Urine output Renal function Number of Serious AEs Ototoxicity	
Lahav 1992 ¹¹⁷ n=9	Intermittent IV bolus furosemide vs. continuous infusion	Patients admitted to hospital with NYHA Class III/IV congestive heart failure	Urine output Renal function Number of AEs Withdrawals due to AEs	96 hours; Crossover
Licata 2003 ¹²² n=107	Twice daily IV bolus furosemide vs. twice daily 30 min furosemide infusion combined with HSS	NYHA Class IV refractory CHF	All-cause mortality Cardiac cause mortality Dyspnoea Length of hospital stay Readmission rates Weight loss Urine Output Renal function Number of AEs Ototoxicity	Up to 55 week follow-up; Parallel
Parrinello 2011 ¹⁷⁰ n=133	Twice daily IV infusion of furosemide vs. twice daily 30 min furosemide infusion combined with HSS	NYHA Class IV refractory CHF	Length of hospital stay Readmission rates Weight loss Urine output Renal function Number of AEs	6 days; Parallel
Paterna 2005 ^{171,172} n=94	Twice daily IV bolus furosemide vs. twice daily 30 min furosemide infusion combined with HSS	NYHA Class IV refractory CHF	All-cause mortality Length of hospital stay Readmission rates Weight loss Dyspnoea Urine output Renal function Number of AEs	4-6 days; 30 day follow-up; Parallel
Pivac 1998 ¹⁷⁷ N=20	Single IV bolus furosemide vs. continuous infusion	Congestive cardiac failure (NYHA Class III/IV) with third degree of oedema	Urine output Renal function	48 hours; Crossover
Thomson 2010 ²¹³ n=56	Intermittent IV furosemide vs. continuous IV	Patients aged over 18 with ADHF	Mortality Length of hospital stay Weight loss Urine output Renal function	Unspecified time frame; Parallel

Study	Intervention / Comparison	Population	Outcomes	Comments
			Withdrawals due to AEs	

Table 34: Clinical evidence profile: IV furosemide bolus versus IV furosemide infusion

Quality assessment							Summary of findings				Quality	Importance
									Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Bolus Injection (n) Mean (SD) or Event rate	Continuous Infusion (n) Mean (SD) or Event rate	Relative Risk (95% CI)	Absolute effect Mean Difference (MD) (95% CI)		
All-cause mortality (follow up mean 60 days) ^{73,75}												
1	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious	none	13/156 (8.3%)	16/152 (10.5%)	RR 0.79 (0.39 to 1.59)	22 fewer per 1000 (from 64 fewer to 62 more)	VERY LOW	CRITICAL
Rehospitalisation (follow up mean 60 days) ^{73,75}												
1	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(a)	none	50/156 (32.1%)	42/152 (27.6%)	RR 1.16 (0.82 to 1.64)	44 more per 1000 (from 50 fewer to 177 more)	LOW	IMPORTANT
Number of patients visiting ED (follow up mean 60 days) ^{73,75}												
1	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious	none	16/156 (10.3%)	21/152 (13.8%)	RR 0.74 (0.4 to 1.37)	36 fewer per 1000 (from 83 fewer to 51 more)	VERY LOW	IMPORTANT
Dyspnoea - AUC at 72 hours (follow-up mean 72 Hours; measured with: AUC of serial VAS measurements (higher scores indicate improved symptoms)) ^{73,75}												
1	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(a)	none	n=156 4456 (1468)	n=152 4669 (1573)	-	MD 243 lower (583 lower to 97 higher)	LOW	IMPORTANT
Weight Loss (kg) (follow-up 72 hours - discharge; better indicated by lower values) ^{14,73,213} – means and (sd) provided by study												
3	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	n=207 -1.64 (2.34) -3.08 (5.38) -5.1 (4.6)	n=198 -2.66 (2.44) -3.67 (4.67) -6.8 (6.1)	-	MD 0.78 higher (0.03 to 1.54 higher)	MODERATE	IMPORTANT
Length of Hospital Stay (days) (better indicated by lower values) ^{14,213}												

2	RCT	serious ^(a)	serious ^(b)	no serious indirectness	serious ^(a)	none	n=51 8.86 (3.82) 10.9 (8.3)	n=46 9.85 (11.72) 6.9 (3.7)	-	MD 1.98 higher (2.82 lower to 6.78 higher)	VERY LOW	IMPORTANT
UO - Total Urine Output (ml) (follow-up 24-72 Hours/discharge; assessed with: Urine Output; better indicated by higher values) ^{2,14,117,213}												
4	RCT	very serious ^(a)	very serious ^(b)	no serious indirectness	serious ^(a)	none	n= 114 Crossover study and parallel data combined using generic inverse variance ^(d)	-	Generic Inverse Variance MD 346.19 lower (1231.75 lower to 731.13 higher)	VERY LOW	IMPORTANT	
Sub analysis: UO - Total Urine Output (ml) (follow-up mean 24 hours; measured with: Urine Output; better indicated by higher values) ^{2,213}												
2	RCT	very serious ^(a)	very serious ^(b)	no serious indirectness	very serious ^(a)	none	n= 64 Crossover study and parallel data combined using generic inverse variance ^(d)	-	Generic Inverse Variance MD 250.31 lower (1231.75 lower to 731.13 higher)	VERY LOW	IMPORTANT	
Sub analysis: UO - Total Urine Output (ml) (follow-up mean 48 hours; better indicated by higher values) ¹¹⁷												
1	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	n= 9 Crossover study and parallel data combined using generic inverse variance ^(d)	-	Generic Inverse Variance MD 690.56 lower (834.56 to 546.55 lower)	LOW	IMPORTANT	
Sub analysis: UO - Total Urine Output (ml) (follow-up mean 72 hours; measured with Urine Output; better indicated by higher values) ^{12,14}												
1	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(a)	none	n= 41 Crossover study and parallel data combined using generic inverse variance ^(d)	-	Generic Inverse Variance MD 219 higher (1147.18 lower to 1585.18 higher)	VERY LOW	IMPORTANT	
UO - Net Urine Output (ml) (follow-up 24-72 Hours; measured with Urine Output; better indicated by higher values) ^{73,213}												
2	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	Serious ^(c)	none	n=186 1575 (1100) 4237 (3208)	n=178 2098 (1132) 4249 (3104)	-	MD 313.87 lower (764.83 lower to 137.1 higher)	LOW	IMPORTANT
Sub analysis UO - Net Urine Output (ml) (follow-up mean 24 Hours; better indicated by higher values) ²¹³												

1	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	Serious ^(c)	none	n=30 1575 (1100)	n=26 2098 (1132)	-	MD 523 lower (1109.74 lower to 63.74 higher)	LOW	IMPORTANT
Sub analysis: UO - Net Urine Output (ml) (follow-up mean 72 hours; measured with Urine Output; better indicated by higher values)^{73,75}												
1	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	n=156 4237 (3208)	n=152 4249 (3104)	-	MD 12 lower (716.92 lower to 692.92 higher)	MODERATE	IMPORTANT
Renal Function - Change in creatinine (mg/dl) from baseline (follow-up 72-discharge hours; measured with serum creatinine (md/dl); better indicated by lower values)^{14,73,213}												
3	RCT	very serious ^(a)	very serious ^(b)	no serious indirectness	serious	none	n=207 -0.02 (0.39) 0.05 (0.3) 0.36 (0.29)	n=198 0.13 (0.34) 0.07 (0.3) 0.19 (0.29)	-	MD 0.01 higher (0.14 lower to 0.16 higher)	VERY LOW	IMPORTANT
Sub analysis: Renal Function - Change in creatinine (mg/dl) from baseline (follow-up 72-discharge hours; measured with serum creatinine (mg/dl); better indicated by lower values)^{14,73}												
2	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	n=177 -0.02 (0.39) 0.05 (0.3)	n=172 0.13 (0.34) 0.07 (0.3)	-	MD 0.04 lower (0.13 lower to 0.05 higher)	LOW	IMPORTANT
Sub analysis: Renal Function - Change in creatinine (mg/dl) from baseline (Unspecified endpoint; measured with serum creatinine (mg/dl); better indicated by lower values)²¹³												
1	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	serious	none	n=30 0.36 (0.29)	n=26 0.19 (0.29)	-	MD 0.17 higher (0.02 to 0.32 higher)	VERY LOW	IMPORTANT
Renal Function - Total number of patients with an increase in serum creatinine of > 0.3mg/dL (follow-up mean 72 hours)^{73,75}												
1	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious	none	27/155 (17.4%)	28/146 (19.2%)	RR 0.89 (0.5 to 1.6)	17 fewer per 1000 (from 86 fewer to 83 more)	VERY LOW	IMPORTANT
Renal Function - Total number of patients with an increase in serum creatinine of >0.5mg/dL (at unspecified endpoint)²¹³												
1	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	very serious	none	5/30 (16.7%)	5/26 (19.2%)	RR 0.87 (0.28 to 2.66)	25 fewer per 1000 (from 138 fewer to 319 more)	VERY LOW	IMPORTANT
Toxicity - Total number of patients with any SAE (follow-up mean 60 days)^{73,75}												
1	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious	none	69/156 (44.2%)	67/152 (44.1%)	RR 1 (0.78 to	0 fewer per 1000 (from 97 fewer to	LOW	IMPORTANT

									1.29)	128 more)		
Toxicity - Number of patients with ventricular tachycardia (VT) (follow-up mean 60 days)^{73,75}												
1	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious	none	7/155 (4.5%)	4/152 (2.6%)	RR 1.72 (0.51 to 5.74)	19 more per 1000 (from 13 fewer to 125 more)	VERY LOW	IMPORTANT
Toxicity - Number of patients with new Myocardial Infarction (MI) (follow-up mean 60 days)^{73,75}												
1	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious	none	4/155 (2.6%)	1/152 (0.7%)	RR 3.92 (0.44 to 34.7)	19 more per 1000 (from 4 fewer to 222 more)	VERY LOW	IMPORTANT

- (a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitations across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.
- (b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious ($I^2 = 50$ to 74% , or 2 p -value < 0.1). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious ($I^2 > 75\%$). Urine output and renal function were sub-grouped by time at measurement Length of stay was sub-grouped by country of study. These sub-grouping strategies failed to remove heterogeneity in all cases. Inconsistent outcomes were therefore re-analysed using a random effects model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis.
- (c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID. Default MIDs were set at RRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the control group median standard deviation either side of the null line for continuous variables
- (d) Crossover trial data presented as paired data thus analysed by generic inverse variance (GIV), this is a statistical technique available in Review Manager, which allows imputation of ratio measure of intervention effects. Where outcomes were reported as continuous data in combinable studies this was converted to generic inverse variance to allow meta-analysis.

Table 35: Clinical evidence profile: IV furosemide bolus versus IV furosemide infusion with HSS

Quality assessment							Summary of findings				Quality	Importance
							Bolus Injection (n) Mean (SD) or event rate	Infusion with HSS (n) Mean (SD) or event rate	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			Relative Risk (95% CI)	Absolute effect Mean Difference (MD) (95% CI)		
All-Cause Mortality (follow-up 1-48 months) ^{122,173}												

2	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	50/100 (50%)	24/101 (23.8%)	RR 2.03 (1.48 to 2.79)	245 more per 1000 (from 114 more to 425 more)	LOW	CRITICAL
									22.6%			
Cardiac mortality (follow-up mean 48 months) ¹²²												
1	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	43/54 (79.6%)	20/53 (37.7%)	RR 2.11 (1.46 to 3.06)	419 more per 1000 (from 174 more to 777 more)	LOW	CRITICAL
									37.7%			
Weight loss at discharge (kg) (Better indicated by higher values) ^{122,173}												
2	RCT	very serious ^(a)	Serious ^(b)	no serious indirectness	Serious ^(c)	none	n=100	n=101	-	MD 2.09 lower (3.47 to 0.72 lower)	VERY LOW	IMPORTANT
							8.5 (2.6)	9.9 (4.15)				
							8.1 (2.4)	10.9 (4.1)				
Length of hospital stay (Better indicated by lower values) ^{122,173}												
2	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	n=100	n=101	-	MD 3.5 higher (2.82 to 4.18 higher)	LOW	IMPORTANT
							11.7 (2.6)	8.57 (2.3)				
							10.5 (2.6)	6.57 (2.3)				
Readmission: Number of patients readmitted during follow up due to acute heart failure (follow-up 1-48 months) ^{122,173}												
2	RCT	very serious ^(a)	very serious ^(b)	no serious indirectness	very serious ^(c)	none	55/100 (55%)	25/101 (24.8%)	RR 5.16 (0.21 to 126.19)	1000 more per 1000 (from 196 fewer to 1000 more)	VERY LOW	IMPORTANT
								23.6%				
UO: Total Urine Output/24 hours (ml) (follow-up mean 24 hours; Better indicated by higher values) ^{122,173}												
2	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	n=100	n=101	-	MD 555.35 lower (697.03 to 413.67 lower)	LOW	IMPORTANT
							1650 (535)	2100 (622)				
							1550 (355)	2180 (545)				
Ototoxicity: Total number of patients reporting hearing loss or deafness (follow-up mean 48 months) ¹²²												
1	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	11/54 (20.4%)	0/53 (0%)	Peto OR 8.92 (2.58 to 30.57)	Risk Difference 0.20 [0.09, 0.31]	LOW	IMPORTANT

(a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitations across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitations across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

(b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious ($I^2 = 50$ to 74% , or X^2 p-value ≤ 0.05). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious ($I^2 \geq 75\%$). Weight loss and readmissions were sub-grouped by country of study, however this failed to remove heterogeneity in all cases. Inconsistent outcomes were therefore re-analysed using a random effects model, rather than the default fixed effect model used initially for all outcomes. The point estimate and 95% CIs given in the grade table and forest plots are those derived from the new random effects analysis.

(c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID. Default MIDs were set at RRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the control group weighted mean standard deviation either side of the null line for continuous variables

Table 36: Clinical evidence profile: IV furosemide infusion versus IV furosemide infusion with HSS

Quality assessment							Summary of findings				Quality	Importance
									Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	30 min Infusion (n) Mean (SD) or event rate	Infusion with HSS (n) Mean (SD) or event rate	Relative Risk (95% CI)	Absolute effect Mean Difference (MD) 95% CI		
Weight (kg) at 6 days or discharge (follow-up 0-6 days; better indicated by lower values) ¹⁷⁰												
1	RCT	Serious ^(a)	no serious inconsistency	no serious indirectness	Serious ^(b)	none	n=66 12 (4)	n=67 6.3 (3)	-	MD 3.2 higher (0.29 to 6.11 higher)	LOW	IMPORTANT
Length of hospital stay (days) (better indicated by lower values) ¹⁷⁰												
1	RCT	Serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	n=66 68 (11)	n=67 64.8 (5)	-	MD 5.7 higher (4.5 to 6.9 higher)	MODERATE	IMPORTANT
UO: Total Urine Output / 24 hours (ml) (follow-up mean 24 hours; better indicated by higher values) ¹⁷⁰												
1	RCT	Serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	n=66 1550 (355)	n=67 2180 (545)	-	MD 630 lower (786.09 to	MODERATE	IMPORTANT

[illegible]

(a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitations across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitations across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID. Default MIDs were set at RRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the control group weighted mean standard deviation either side of the null line for continuous variables

7.3.2 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

New cost-effectiveness analysis

New analysis was not prioritised for this area.

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs were provided to aid consideration of cost-effectiveness. These are given in Appendix N.

7.3.3 Evidence statements

7.3.3.1 Clinical

IV furosemide bolus versus IV furosemide infusion

Mortality

Very low quality evidence from 1 study with 308 participants showed there was a relative increase in all-cause mortality at 60 days associated with an IV furosemide infusion compared to IV bolus furosemide. However, the uncertainty around this effect was too large to draw clear conclusions about clear clinical benefit or harm

Rehospitalisation

Low quality evidence from 1 study with 308 participants showed there was a relative increase in rehospitalisation at 60 days associated with IV bolus furosemide compared to an IV furosemide infusion. However, the uncertainty around this effect was too large to draw clear conclusions about clear clinical benefit or harm

Number of patients visiting Emergency Department

Very low quality evidence from 1 study with 308 participants showed there was a relative increase in numbers of patients visiting ED at 60 days associated with an IV furosemide infusion compared to IV bolus furosemide. However, the uncertainty around this effect was too large to draw clear conclusions about clear clinical benefit or harm

Dyspnoea

Low quality evidence from 1 study with 308 participants showed there was a relative improvement in dyspnoea symptoms (as assessed by AUC) at 72 hours associated with an IV furosemide infusion compared to IV bolus furosemide. However, the uncertainty around this effect was too large to draw clear conclusions about clear clinical benefit or harm.

Weight loss (kg)

Moderate quality evidence from 3 studies with 405 participants showed that an IV furosemide infusion was more effective in increasing weight loss at 72 hours (or discharge) compared to IV bolus furosemide, but this effect was not large enough to indicate appreciable clinical benefit.

Length of hospital stay (days)

Very low quality evidence from 2 studies with 97 participants showed there was a relative decrease in length of hospital stay associated with an IV furosemide infusion compared to IV bolus furosemide. However, the uncertainty around this effect was too large to draw clear conclusions about clear clinical benefit or harm.

Urine output (ml)

Total

Very low quality evidence from 4 studies with 114 participants showed there was a relative increase in overall total urine output from 24-72 hours (or discharge) associated with an IV furosemide infusion compared to IV bolus furosemide. However, the uncertainty around this effect was too large to draw clear conclusions about clear clinical benefit or harm. The 2 studies reporting at 24 hours showed a relative increase in urine output associated with an IV furosemide infusion compared to IV bolus furosemide. However, the uncertainty around this effect was far too large to draw clear conclusions about clear clinical benefit or harm. The study reporting at 48 hours showed a clear clinically appreciable increase in urine output at 48 hours with an IV furosemide infusion compared to IV bolus furosemide. The study reporting at 72 hours showed a relative increase in urine output associated with IV bolus furosemide compared to an IV bolus furosemide. However, the uncertainty around this effect was far too large to draw clear conclusions about clear clinical benefit or harm.

Net

Low quality evidence from 2 studies with 364 participants showed there was a relative increase in net total urine output from 24-72 hours associated with an IV furosemide infusion compared to IV bolus furosemide. However, the uncertainty around this effect was too large to draw clear conclusions about clear clinical benefit or harm. The study reporting at 24 hours showed a relative increase in urine output associated with an IV furosemide infusion compared to IV bolus furosemide. However, the uncertainty around this effect was too large to draw clear conclusions about clear clinical benefit or harm. The study reporting at 48 hours showed no clear advantage of an IV furosemide infusion compared to IV bolus furosemide.

Renal Function

Change in serum creatinine (mg/dL) from baseline

Very low quality evidence from 3 studies with 405 participants showed there was no clear advantage of an IV furosemide infusion compared to IV bolus furosemide with regard to change in serum creatinine (mg/dL) from baseline at 72 hours or discharge. When separated by time of measurement, the 2 studies reporting at 24 hours showed no clear advantage of an IV furosemide infusion compared to IV bolus furosemide, the third study reporting at an unspecified follow-up endpoint showed use of an IV furosemide infusion has a smaller change in serum creatinine from baseline compared to IV bolus furosemide, however there was uncertainty as to the clinical benefit.

Numbers of patients with an increase in serum creatinine of > 0.3 mg/dL

Very low quality evidence from 1 study with 301 participants showed there were relatively fewer numbers of patients with an increase in serum creatinine of > 0.3 mg/dL at 72 hours when treated IV bolus furosemide compared to an IV furosemide infusion. However, the uncertainty around this effect was far too large to draw clear conclusions about clear clinical benefit or harm.

Numbers of patients with an increase in serum creatinine of > 0.5 mg/dL

Very low quality evidence from 1 study with 56 participants showed there were relatively fewer numbers of patients with an increase in serum creatinine of > 0.5 mg/dL at 72 hours when treated IV bolus furosemide compared to an IV furosemide infusion. However, the uncertainty around this effect was far too large to draw clear conclusions about clear clinical benefit or harm.

Number of patients with serious adverse events (SAEs)

(AMI; atrial fibrillation; cardiac arrest; ventricular tachycardia; gout, hyperkalaemia, hypokalaemia, hyponatremia, renal failure and renal failure requiring dialysis.)

Low quality evidence from 1 study with 308 participants showed there was no clear advantage of an IV bolus furosemide compared to an IV furosemide infusion with regard number of patients with any SAEs at 60 days.

Total number of patients with ventricular tachycardia (VT)

Very low quality evidence from 1 study with 308 participants showed there was no clear advantage of an IV bolus furosemide compared to an IV furosemide infusion with regard number of patients with ventricular tachycardia (VT) at 60 days.

Total number of patients with new myocardial infarction (MI)

Very low quality evidence from 1 study with 308 participants showed there relatively fewer numbers of patients with new myocardial infarction (MI) at 60 days with an IV furosemide infusion compared to IV bolus furosemide. However, the uncertainty around this effect was far too large to draw clear conclusions about clear clinical benefit or harm.

IV furosemide bolus versus IV furosemide infusion with HSS

All-cause mortality

Low quality evidence from 2 studies with 201 participants showed there was clear clinical benefit of an IV furosemide infusion with HSS compared to IV furosemide bolus in reducing all-cause mortality at 1-48 months.

Weight loss (kg)

Very low quality evidence from 2 studies with 201 participants showed that an IV furosemide infusion with HSS increased weight loss (kg) at discharge compared to IV furosemide bolus. However there was uncertainty as to the clinical benefit.

Length of hospital stay (days)

Low quality evidence from 2 studies with 201 participants showed that there was clear clinical benefit of an IV furosemide infusion with HSS compared to IV furosemide bolus in terms of reducing length of hospital stay.

Number of patients readmitted due to AHF

Very low quality evidence from 2 studies with 201 participants showed that there were fewer numbers of patients admitted during follow up (1-48 months) due to acute heart failure when treated with an IV furosemide infusion with HSS compared to IV furosemide bolus. However, the uncertainty around this effect was too large to draw clear conclusions about clear clinical benefit or harm.

Number of patients reporting hearing loss or deafness

Low quality evidence from 1 study with 107 participants showed that there was clear clinical benefit of an IV furosemide infusion with HSS over IV furosemide bolus with reduced number of patients reporting hearing loss or deafness at 48 weeks.

Urine output (ml)

Low quality evidence from 2 studies with 201 participants showed that there was clear clinical benefit of an IV furosemide infusion with HSS over IV furosemide bolus with increased urine output (ml) at 24 hours.

IV furosemide infusion versus IV furosemide infusion with HSS

Weight (kg)

Low quality evidence from 1 study with 133 participants showed that there was clear clinical benefit of an IV furosemide infusion with HSS compared to IV furosemide infusion with decreased weight (kg) at 6 days (or discharge).

Length of hospital (days)

Moderate quality evidence from 1 study with 133 participants showed that there was a relative reduction in length of hospital stay with an IV furosemide infusion with HSS compared to IV furosemide infusion. However, the uncertainty around this effect was too large to draw clear conclusions about clear clinical benefit.

Urine output (ml)

Moderate quality evidence from 1 study with 133 participants showed that there was clear clinical benefit of an IV furosemide infusion with HSS compared to IV furosemide infusion with increased urine output ml at 24 hours.

7.3.3.2 Economic

No relevant economic evaluations were identified.

7.3.4 Recommendations and link to evidence

Recommendations	<p>12.Offer intravenous diuretic therapy to people with acute heart failure. Start treatment using either a bolus or infusion strategy.</p> <p>13.For people already taking a diuretic, consider a higher dose of diuretic than that on which the person was admitted unless there are serious concerns with patient adherence to diuretic therapy before admission.</p> <p>14.Closely monitor the person's renal function, weight and urine output during diuretic therapy.</p> <p>15.Discuss with the person the best strategies of coping with an increased urine output.</p>
Relative values of different outcomes	The GDG gave most weight to mortality, but also considered dyspnoea, urine output, weight loss, length of hospital stay, re-admission rates, quality of life, serum creatinine level (or other measures of renal function such as eGFR) and adverse events (particularly renal adverse events and ototoxicity) to be relevant. Urine output and weight loss contributed most to the recommendations, though the GDG commented that accurate measurement of urine output is challenging. The GDG considered length of hospital stay to be subject to confounding, so this outcome did not make a major contribution to the recommendations.
Trade-off between clinical benefits and harms	The use of high dose loop diuretics raised concerns relating to ototoxicity, but the group were reassured that no ototoxicity was demonstrated in the DOSE trial (Felker 2011 ^{73,75}). In both bolus and continuous infusion strategies there were benefits in terms of weight loss and urine output but neither strategy was superior and their use may depend on the clinical circumstances. The GDG noted that there is likely to be a

Recommendations	<p>12.Offer intravenous diuretic therapy to people with acute heart failure. Start treatment using either a bolus or infusion strategy.</p> <p>13.For people already taking a diuretic, consider a higher dose of diuretic than that on which the person was admitted unless there are serious concerns with patient adherence to diuretic therapy before admission.</p> <p>14.Closely monitor the person’s renal function, weight and urine output during diuretic therapy.</p> <p>15.Discuss with the person the best strategies of coping with an increased urine output.</p>
	rise in serum creatinine during diuretic therapy so close monitoring is required.
Economic considerations	In the absence of evaluations of cost-effectiveness, the GDG made a qualitative judgement on cost effectiveness. There were no clear differences in clinical outcomes comparing infusion to bolus strategies, and differences in resource use were unlikely to be large, given the time required to administer repeated boluses.
Quality of evidence	The quality of evidence was rated as moderate to very low according to GRADE. The GDG considered the DOSE trial (Felker 2011 ^{73,75}) contributed the most robust evidence, and therefore its results were given more weight in the discussion. Studies involving a crossover design are subject to particular limitations in acute heart failure and so were given less weight.
Other considerations	<p>No evidence was found for comparison of oral and intravenous strategies. The GDG recommended that diuretics are used intravenously to ensure rapid onset of delivery, action and efficacy.</p> <p>The GDG agreed that weight ought to be monitored even though weight loss was the clinical outcome reported in the included studies because weight gain could indicate that the diuretic treatment strategy is not effective.</p> <p>In terms of starting dose, the GDG drew upon the dosing used in the DOSE study (Felker 2011^{73,75}). The GDG noted that a high dose strategy, in which the starting dose is 2.5 times the patient’s usual oral dose, is more effective in terms of weight loss but there is a trend to worsening renal function. This did not translate to longer term harm. However, careful monitoring of renal function is important. In patients who are diuretic-naïve or in whom medication compliance is uncertain a low dose strategy should be commenced initially. In those patients who are treatment-resistant higher doses of diuretic should be used early.</p> <p>There was little difference in effectiveness in the DOSE trial between bolus and continuous infusion. However, the GDG noted that the practical implications of an infusion strategy, such as need for specialised equipment (such as pumps) and trained staff, were not captured by the studies.</p> <p>The GDG considered that in the acute setting and with lower total doses of diuretics, bolus infusion dosing was likely to be more practical. High doses may be given via a divided bolus strategy ensuring that the maximum rate of administration is not exceeded. BNF guidance indicates that administration of intravenous furosemide should not usually exceed 4mg/minute, however, a single dose of up to 80mg may be administered more rapidly. The relative advantage of an infusion strategy increases as the diuretic dose rises, due to the slow rate at which boluses of diuretic need to be administered.</p> <p>The GDG discussed the possibility of addition of a thiazide diuretic to loop diuretic therapy. This is suggested in the European ESC 2012 guidance, but the evidence base</p>

Recommendations	<p>12.Offer intravenous diuretic therapy to people with acute heart failure. Start treatment using either a bolus or infusion strategy.</p> <p>13.For people already taking a diuretic, consider a higher dose of diuretic than that on which the person was admitted unless there are serious concerns with patient adherence to diuretic therapy before admission.</p> <p>14.Closely monitor the person’s renal function, weight and urine output during diuretic therapy.</p> <p>15.Discuss with the person the best strategies of coping with an increased urine output.</p>
	<p>is limited. A randomised controlled trial of this strategy is required to allow definitive guidance. In the interim, such a strategy can be considered.</p> <p>The GDG noted that the use of hypertonic saline solution (HSS) administered along with diuretics and dietary sodium modulation is not currently used in the UK in patients with AHF. The trials in this area largely originate from one Italian study group and there was concern about the relevance of these studies to a UK practice. The GDG were uncertain as to how a short term diuretic strategy could have led to a large reduction in mortality at 48 months of follow-up. The reproduction of these findings in an international multi-centre study would be welcome.</p> <p>The patient representatives raised the importance of ensuring that health care professionals are aware that patients will pass higher urine volumes during intense diuresis. Therefore medical and nursing staff must ensure patient comfort and hygiene at this time.</p>

7.4 Vasodilators

People presenting with acute heart failure and pulmonary oedema in the United Kingdom are commonly offered intravenous vasodilator medications (most frequently nitrates). Pulmonary oedema is associated with left heart failure and high left atrial pressure. Vasodilator medications are generally believed to be beneficial through a number of mechanisms but effectively reducing the work of the left ventricle and reducing left atrial pressure which may contribute to the relief of pulmonary oedema. However, vasodilators may also have adverse effects including headaches and systemic hypotension. Their impact on clinical outcomes in patients with acute heart failure is unclear and therefore this review aims to evaluate their role.

Review question: In patients with acute heart failure are vasodilators more clinically or cost effective than placebo to improve clinical outcomes?

For full details see review protocol in Appendix C.

Table 37: PICO characteristics of review question

Population	Adults with acute heart failure
Intervention/s	<ul style="list-style-type: none"> • Glyceryl trinitrate(GTN)/Nitroglycerin/e • Isosorbide dinitrate • Sodium nitroprusside

Comparison/s	Placebo (medical care)
Outcomes	<ul style="list-style-type: none"> • Mortality • Major cardiovascular events • Length of hospital stay and readmission rates • Quality of life • Dyspnoea • Haemodynamic outcomes: e.g. Pulmonary Capillary Wedge Pressure (PCWP) / Pulmonary Artery Wedge Pressure (PAWP), Right Atrial Pressure (RAP) PCWP, Cardiac Index • Discontinuation of therapy • Adverse events (particularly headache and hypotension)
Study design	Systematic reviews and randomised controlled trials

7.4.1 Clinical evidence

Five studies were included in the review^{44,60,63,142,231}. There were three vasodilators that were compared to placebo:

- Two studies^{63,231} for intravenous (IV) nitroglycerin (see Table 38 Table 37)
- Two studies^{60,142} for oral isosorbide dinitrate (see Table 39)
- One study⁴⁴ for intravenous sodium nitroprusside (see Table 40)

Evidence from these are summarised in the clinical GRADE evidence profiles below (Table 42 and Table 43).

See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and lists of excluded studies in Appendix K.

Studies are briefly summarised in the tables below which are divided by comparisons.

Table 38: Summary of studies included in the review — IV nitroglycerin vs. placebo

Study	Population	Intervention/comparison	Outcomes	Comments
Elkayam et al., 1987 ⁶³ N=40	CAD with concomitant heart failure. Author described population as ADHF in subsequent review ^{63,65} Only patients 'responsive' to open label IV nitroglycerin	IV nitroglycerin versus placebo	Haemodynamic: PAWP	24 hours; Parallel
Young et al., VMAC 2002 ²³¹ N=489	Acute decompensated heart failure	IV nitroglycerin versus placebo	Haemodynamics: PCWP, RAP, cardiac Index Dyspnoea Discontinuation Adverse events	3 hours; Parallel

For oral isosorbide dinitrate the main characteristics are described below.

Table 39: Summary of studies included in the review — oral isosorbide dinitrate vs. placebo

Study	Population	Intervention/comparison	Outcomes	Comments
Dubourg et al., 1984 ⁶⁰	Acute myocardial infarction complicated by moderate	Oral isosorbide dinitrate versus placebo	Haemodynamics: PAWP, PCWP,	8 hours; Parallel

Study	Population	Intervention/comparison	Outcomes	Comments
N=20	left ventricular failure		Cardiac Index	
Mikulic et al., 1975 ¹⁴² N=8	Evidence of congestive heart failure (dyspnoea, S3 gallop, jugular venous distension, oedema and pulmonary congestion on clinical or radiological examination) secondary to ischaemic or primary myocardial disease or acute myocardial infarction	Chewable isosorbide dinitrate versus sublingual nitroglycerin versus placebo	Haemodynamics: LVFP	3 hours; Parallel

The study comparing IV sodium nitroprusside to placebo is briefly summarised below.

Table 40: Summary of studies included in the review — IV sodium nitroprusside vs. placebo

Study	Population	Intervention/comparison	Outcomes	Comments
Cohn et al., 1982 ⁴⁴ N=812	Men with presumed acute myocardial infarction complicated by left ventricular failure	IV sodium nitroprusside versus placebo	Mortality: 21 day and 13 week Haemodynamic: LVFP Toxicity: Hypotension and headache	48 Hours; Parallel

Table 41: Clinical evidence profile: Intravenous nitroglycerin versus placebo

Quality assessment							Summary of findings				Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intravenous nitroglycerin (n) Mean (SD) or Event rate	Placebo (n) Mean (SD) or Event rate	Effect Size			
									Relative Risk (95% CI)	Absolute effect Mean Difference (MD) (95% CI)		
Haemodynamic: Mean change in PCWP (mmHg) from baseline (follow-up 3-24 hours; Better indicated by lower values) ^{64,231}												
1 3 hours	randomised trials	Serious ^(a)	no serious inconsistency	no serious indirectness	Serious ^(c)	none	n=60 -3.8 (5.3)	n=62 -2 (4.2)	-	MD 1.8 lower (3.5 to 0.1 lower)	LOW	IMPORTANT
1 24 hours	randomised trials	very serious ^(a)	no serious inconsistency	no serious indirectness	Serious ^(c)	none	n=15 -7 (6)	n=16 -2 (4)	-	MD 5 lower (8.61 to 1.39 lower)	VERY LOW	IMPORTANT
2 Total	randomised trials	very serious ^(a)	Serious ^(c)	no serious indirectness	Serious ^(c)	none	n=75	n=78	-	MD 2.38 lower (3.92 to 0.84 lower)	VERY LOW	IMPORTANT
Haemodynamic: Mean change in right atrial pressure (mmHg) from baseline at 3 hours (follow-up mean 3 hours; Better indicated by lower values) ²³¹												
1	randomised trials	Serious ^(a)	no serious inconsistency	no serious indirectness	Serious ^(c)	none	n=60 -2.6 (3.5)	n=62 0 (4.4)	-	MD 2.6 lower (4.01 to 1.19 lower)	LOW	IMPORTANT
Haemodynamic: Mean change in cardiac index (L/min per m2) at 3 hours (Better indicated by higher values) ²³¹												
1	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	serious ^c	none	n=60 0.2 (0.5)	n=62 0 (0.6)	-	MD 0.2 higher (0 to 0.4 higher)	LOW	IMPORTANT

Dyspnoea: Number of patients reporting markedly, moderately or minimally better at 3 hours ²³¹												
1	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	serious ^c	none	103/143 (72%)	92/142 (64.8%)	RR 1.11 (0.95 to 1.3)	71 more per 1000 (from 32 fewer to 194 more)	LOW	IMPORTANT
Global clinical status: Number of patients reporting markedly, moderately or minimally better at 3 hours ²³¹												
1	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	no serious imprecision	none	93/143 (65%)	93/142 (65.5%)	RR 0.99 (0.84 to 1.18)	7 fewer per 1000 (from 105 fewer to 118 more)	MODERATE	IMPORTANT
Toxicity: Number of patients with any Adverse Event (follow-up 3 hours) ²³¹												
1	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	serious ^c	none	39/143 (27.3%)	20/142 (14.1%)	RR 1.94 (1.19 to 3.15)	132 more per 1000 (from 27 more to 303 more)	LOW	IMPORTANT
Toxicity: Number of patients discontinuing therapy due to drug (follow-up 3 hours) ²³¹												
1	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	very serious ^c	none	0/143 (0%)	1/142 (0.7%)	OR 0.13 (0 to 6.77)	6 fewer per 1000 (from 7 fewer to 39 more)	VERY LOW	IMPORTANT
Toxicity: Number of patients with headache (follow-up 3 hours) ²³¹												
1	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	no serious imprecision	none	17/143 (11.9%)	3/142 (2.1%)	RR 5.63 (1.69 to 18.78)	98 more per 1000 (from 15 more to 376 more)	MODERATE	IMPORTANT

(a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitations across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitations across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

(b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I^2 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I^2 75% or more). Mean change in PCWP was sub grouped by time at measurement. This sub-grouping strategy removed the heterogeneity for this outcome

(c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID. Default MIDs were set at RRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the control group median standard deviation either side of the null line for continuous variables

Table 42: Clinical evidence profile: Oral isosorbide dinitrate versus placebo

Quality assessment							Summary of Findings				Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Oral ISDN (n)	Placebo (n)	Effect Size			
							Mean (SD) or Event rate	Mean (SD) or Event rate	Relative Risk (95% CI)	Absolute effect Mean Difference (MD) (95% CI)		
Haemodynamic: Mean PAWP (mmHg) (follow-up 8 hours; Better indicated by lower values) ⁶⁰												
1	randomised trials	very serious ^a	no serious inconsistency	no serious indirectness	serious ^b	none	n=10 20.5 (3.8)	n=10 27.2 (6.6)	-	MD 6.7 lower (11.44 to 1.96 lower)	+ VERY LOW	IMPORTANT
Haemodynamic: Mean PCWP (mmHg) (follow-up 8 hours; Better indicated by lower values) ⁶⁰												
1	Randomised trial	very serious ^a	no serious inconsistency	no serious indirectness	serious ^b	none	n=10 13.1 (3.5)	n=10 18.8 (6.2)	-	MD 5.7 lower (10.13 to 1.27 lower)	+ VERY LOW	IMPORTANT
Haemodynamic: Cardiac Index (l/min/m2) (follow-up 8 hours; Better indicated by higher values) ⁶⁰												
1	randomised trials	very serious ^a	no serious inconsistency	no serious indirectness	very serious ^b	none	n=10 2.95 (0.6)	n=10 3.15 (0.7)	-	MD 0.2 lower (0.74 lower to 0.34 higher)	+ VERY LOW	IMPORTANT

(a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitation *s* across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

(b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID. Default MIDs were set at RRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the control group median standard deviation either side of the null line for continuous variables

Narrative results

Left ventricular filling pressure (Mikulic et al., 1975¹⁴²)

“Chewable isosorbide dinitrate produced significantly lower LVFP compared to placebo from 5 through 180 minutes ($p < 0.05$) while Sublingual nitroglycerin differed significantly from placebo ($p < 0.05$) only at 10 and 15 mins” – Very low quality

Table 43: Clinical evidence profile: Intravenous sodium nitroprusside versus placebo

Quality assessment							Summary of findings				Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intravenous nitroprusside (n) Mean (SD) or Event rate	Placebo (n) Mean (SD) or Event rate	Effect Size			
									Relative Risk (95% CI)	Absolute effect Mean Difference (MD) (95% CI)		
Mortality: All cause ⁴⁴												
48 hours	randomised trial	very serious ^a	no serious inconsistency	no serious indirectness	very serious ^b	none	11/407 (2.7%)	9/405 (2.2%)	RR 1.22 (0.51 to 2.9)	5 more per 1000 (from 11 fewer to 42 more)	+ VERY LOW	CRITICAL
21 days	randomised trial	very serious ^a	no serious inconsistency	no serious indirectness	very serious ^b	none	47/407 (11.5%)	42/405 (10.4%)	RR 1.11 (0.75 to 1.65)	11 more per 1000 (from 26 fewer to 67 more)	+ VERY LOW	CRITICAL
13 weeks	randomised trial	very serious ^a	no serious inconsistency	no serious indirectness	serious ^b	none	69/407 (17%)	77/405 (19%)	RR 0.89 (0.66 to 1.2)	21 fewer per 1000 (from 65 fewer to 38 more)	+ VERY LOW	CRITICAL
Mortality: 13 week mortality according to time of onset of infarction to start of infusion ⁴⁴												
< 9 hours	randomised trial	very serious ^a	no serious inconsistency	no serious indirectness	no serious imprecision	none	250/370 (67.6%)	256/366 (69.9%)	RR 0.97 (0.88 to	21 fewer per 1000 (from 84 fewer to 42 more)	++ LOW	CRITICAL

									1.06)			
> 9 hours	randomised trial	very serious ^a	no serious inconsistency	no serious indirectness	no serious imprecision	none	370/740 (50%)	366/732 (50%)	RR 1 (0.91 to 1.1)	0 fewer per 1000 (from 45 fewer to 50 more)	++ LOW	CRITICAL
Haemodynamic: Number of patients achieving LVFP reduction by 60%⁴⁴												
	randomised trial	very serious ^a	no serious inconsistency	no serious indirectness	no serious imprecision	none	283/407 (69.5%)	16/405 (4%)	RR 17.6 (10.84 to 28.57)	656 more per 1000 (from 389 more to 1000 more)	++ LOW	IMPORTANT
Toxicity: Number of patients reaching hypotensive limit⁴⁴												
	randomised trial	very serious ^a	no serious inconsistency	no serious indirectness	no serious imprecision	none	54/407 (13.3%)	2/405 (0.5%)	RR 26.87 (6.59 to 109.46)	128 more per 1000 (from 28 more to 536 more)	++ LOW	IMPORTANT
Toxicity: Number of patients reporting headache⁴⁴												
	randomised trial	very serious ^a	no serious inconsistency	no serious indirectness	serious ^b	none	142/407 (34.9%)	105/405 (25.9%)	RR 1.35 (1.09 to 1.66)	91 more per 1000 (from 23 more to 171 more)	+ VERY LOW	IMPORTANT
Toxicity: Number of patients with severe headache⁴⁴												
	randomised trial	very serious ^a	no serious inconsistency	no serious indirectness	no serious imprecision	none	18/407 (4.4%)	5/405 (1.2%)	RR 3.58 (1.34 to 9.56)	32 more per 1000 (from 4 more to 106 more)	++ LOW	IMPORTANT

(a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitations across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitations across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.

(b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID. Default MIDs were set at RRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the control group median standard deviation either side of the null line for continuous variables

7.4.2 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs were provided to aid consideration of cost-effectiveness. These are given in Appendix N.

7.4.3 Evidence statements

7.4.3.1 Clinical

Intravenous nitroglycerin versus placebo

In this comparison none of the studies reported mortality data.

Haemodynamic outcomes

- Very low quality evidence from 2 studies with 153 participants showed there was a decrease in mean pulmonary capillary wedge pressure (PCWP) (mmHg) at 3-24 hours associated with IV nitroglycerin compared to placebo. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit. When these studies were sub grouped based on time of PCWP measurement the same effect was found at 3 and 24 hours. The quality of the evidence was low and very low respectively.
- Low quality evidence from 1 study with 153 participants showed there was a decrease in mean right atrial pressure (mmHg) at 3 hours associated with IV nitroglycerin compared to placebo. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit.
- Low quality evidence from 1 study with 153 participants showed there was a relative increase in cardiac index (l/min/m²) at 3 hours associated with IV nitroglycerin compared to placebo. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit.

Dyspnoea

- Low quality evidence from 1 study with 153 participants showed there was no clear advantage of IV nitroglycerin compared to placebo in the number of patients reporting marked, moderate or minimal improvement in dyspnoea symptoms at 3 hours.

Global clinical status

- Moderate quality evidence from 1 study with 153 participants showed there was no clear advantage of IV nitroglycerin compared to placebo in the number of patients reporting marked, moderate or minimal improvement in global symptoms at 3 hours.

Toxicity

- Low quality evidence from 1 study with 153 participants showed there were a relatively greater number of patients with any AE at 3 hours associated with IV nitroglycerin compared

to placebo. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit and harm.

- Very low quality evidence from 1 study with 153 participants showed there was no clear advantage of IV nitroglycerin compared to placebo in the number of patients discontinuing the study drug at 3 hours.
- Moderate quality evidence from 1 study with 153 participants showed there were a greater number of patients reporting headache at 3 hours associated with IV nitroglycerin compared to placebo.

Oral isosorbide dinitrate versus placebo

In this comparison the studies only reported haemodynamic outcomes

Haemodynamic outcomes

- Very low quality evidence from 1 study with 20 participants showed there was a decrease in mean pulmonary capillary wedge pressure (PCWP) and pulmonary arterial wedge pressure (PAWP) (mmHg), and no clear advantage in change in cardiac index (l/min/m²) at 8 hours associated with PO isosorbide dinitrate compared to placebo. However, the uncertainty around these effects was too large to draw conclusions about clear clinical benefit.

Intravenous sodium nitroprusside versus placebo

All-cause mortality

- Very low quality evidence from 1 study with 812 participants showed there was no clear advantage of sodium nitroprusside over placebo in all-cause mortality at 48 hours, 21 days or 13 weeks. When only patients with confirmed acute MI were analysed (n=736) according to if they had received the study medication before or after 9 hours of the onset of the MI no clear advantage was found between intravenous sodium nitroprusside and placebo. The evidence ranged from low to very low quality.

Haemodynamic

- Low quality evidence from 1 study with 812 participants showed there was increased number of patients achieving a reduction in LVFP (mmHg) by greater than 60% with sodium nitroprusside compared to placebo

Toxicity

- Low quality evidence from 1 study with 812 participants showed there was increased number of patients reaching the prespecified hypotensive limit (mmHg) with sodium nitroprusside compared to placebo
- Very low quality evidence from 1 study with 812 participants showed there was increased number of reporting headache and severe headache with sodium nitroprusside compared to placebo. The evidence was deemed to be of very low and low quality respectively. With regard to headache the uncertainty surrounding the effect was too large to draw clear conclusions about clinical benefit and harm.

7.4.3.2 Economic

No relevant economic evaluations were identified.

7.4.4 Recommendations and link to evidence

	<p>16. Do not routinely offer nitrates to people with acute heart failure.</p> <p>17. If intravenous nitrates are used in specific circumstances, such as for people with concomitant myocardial ischaemia, severe hypertension or regurgitant aortic or mitral valve disease, monitor blood pressure closely in a setting where at least level 2 care^e can be provided.</p>
Recommendations	18. Do not offer sodium nitroprusside to people with acute heart failure.
Relative values of different outcomes	Limited clinical endpoints were available, so the GDG focussed on patient-reported outcomes, including dyspnoea, patient global status and patient-reported adverse events, in particular, headache. Haemodynamic outcomes were also available, although it was recognised that these are not necessarily linked to longer term clinical benefit.
Trade-off between clinical benefits and harms	<p>Intravenous nitroglycerin was not associated with any clear global symptomatic improvement or patient reported breathlessness over placebo. Headache was more common in those receiving nitrates as well as other adverse events such as pain and nausea. Hypotension was common with both nitrates and nitroprusside.</p> <p>Both nitrates and nitroprusside appear to have favourable effects on haemodynamic measures, but these did not appear to translate into clinical benefit in the studies reviewed. Nitroprusside was associated with a trend towards increased survival, but this did not persist after 13 weeks.</p> <p>Both nitrates and nitroprusside may increase the risk of harm in patients with hypotension, in particular in those with aortic stenosis.</p>
Economic considerations	No economic evidence on nitrates or sodium nitroprusside was identified. Unit costs were presented to the GDG, and other costs such as drug administration were considered. No evidence of clinical benefit was identified to justify the increased cost associated with routine use of vasodilators.
Quality of evidence	<p>The evidence was mostly of low to very low quality using GRADE criteria. The VMAC study (Young et al., 2002²³¹) was considered to be the most relevant study due to the inclusion of patient reported outcomes (e.g. improvement in global clinical status). However, the study had limitations, as relatively low mean doses of intravenous nitroglycerin were used and there was a potential patient selection bias, as the study was restricted to patients whom the investigators considered appropriate to receive placebo for three hours. Therefore, interpretation of the study results is controversial within the research community. The other study for the Intravenous nitroglycerin versus placebo included only patients with a positive effect in open label run-in which is likely to over-estimate the effect.</p> <p>The GDG acknowledged a further general limitation of the evidence in that patient-reported outcomes may be captured at a time when the researcher is aware of the patient's haemodynamic changes, so introducing ascertainment bias.</p>
Other considerations	Nitrates are used not infrequently (although variably) throughout the UK to treat

^e Level 2 care is for people needing more detailed observation or intervention, including support for a single failing organ system or postoperative care and for those stepping down from higher levels of care. From Intensive Care Society, Levels of Critical Care for Adult Patients (2009).

Recommendations	<p>16. Do not routinely offer nitrates to people with acute heart failure.</p> <p>17. If intravenous nitrates are used in specific circumstances, such as for people with concomitant myocardial ischaemia, severe hypertension or regurgitant aortic or mitral valve disease, monitor blood pressure closely in a setting where at least level 2 care^e can be provided.</p> <p>18. Do not offer sodium nitroprusside to people with acute heart failure.</p>
	<p>acute heart failure, and there is considerable variation in practice across Europe. There is limited evidence of benefit. Adverse events may be increased, in particular, hypotension. The GDG did note, however, that in specific circumstances, such as patients with concomitant myocardial ischaemia, severe hypertension or severe regurgitant aortic or mitral valve disease, nitrates may have a role (although the specific evidence in these areas has not been reviewed). When intravenous nitrates are used care should be taken to ensure that appropriate monitoring is undertaken and that staff are adequately trained. Comparison between different modes of administration of nitrates has not been specifically reviewed, however the GDG consensus was that when deemed necessary nitrates should be given intravenously in order to allow accurate dose titration and cessation in case of adverse effects.</p> <p>The GDG discussed the Cotter study (1998)^{47,48} but expressed concern about the generalisability of the methodology to a UK setting. The lack of a placebo group, the presence of confounders and a high risk of patient selection bias led to exclusion of this study from the meta-analysis.</p> <p>Nitroprusside is of limited availability and is only likely to be used in specialist circumstances and settings, particularly within a critical care setting. The definition of myocardial infarction has changed over the years and this should be taken into consideration when interpreting older data, such as those from Cohn 1982, the nitroprusside study included in this review.</p>

7.5 Inotropes and vasopressors

In some patients with severe acute heart failure there may be an imminent danger to life due to poor cardiac output, systemic hypotension and hypoperfusion. Inotropic medications which increase the force of myocardial contraction and vasopressors which increase systemic vascular resistance and raise blood pressure, may be key to preventing further clinical deterioration and allowing time to investigate the underlying cause of the heart failure. However their use may be associated with side effects and there is variation in what medications are used and in whom. Therefore this review investigates the use of these medications and their effect on clinical outcomes.

Review question: In patients with acute heart failure are inotropes or vasopressors safe and clinically / cost effective compared to standard medical treatment or each other to improve outcome?

For full details see review protocol in Appendix C. Please see Table 44 for a breakdown of the population, intervention/s, the comparator/s, outcomes, and study design.

Table 44: PICO characteristics of review question

Population	Adults with acute heart failure
-------------------	---------------------------------

Intervention/s	<ul style="list-style-type: none"> • Inotropes (milrinone, enoximone, dobutamine, dopamine) • vasopressors (adrenaline, noradrenaline/norepinephrine, vasopressin)
Comparison/s	Standard medical care (any form of standard medical care provided for the management of acute heart failure) plus placebo, or each other.
Outcomes	<ul style="list-style-type: none"> • Mortality (hospital mortality) • Major cardiovascular events • Length of hospital stay (or intensive care unit stay) and re-admission rates • Renal functions • Quality of life • Adverse events
Study design	Systematic reviews and randomised controlled trials

7.5.1 Clinical evidence

We searched for systematic reviews and randomised controlled trials (RCTs) investigating the effectiveness of inotropes and/or vasopressors when compared with each other, or with standard medical care (any form of standard medical care provided for the management of acute heart failure) generally coupled with placebo. Studies comparing one inotrope or one vasopressor to another within the same class of drugs were excluded. The trials were divided into the following comparison groups:

Inotrope evidence

- Two studies compared milrinone to placebo (Cuffe et al, 2002 OPTIME-HF trial⁵⁰, Seino et al, 1996¹⁹⁷) (summarised in Table 45).
- Two studies compared dobutamine to placebo (Adamopoulos et al, 2006⁵ and Zairis et al, 2004 CASINO trial as described in Cleland 2004^{42,43}) (summarised in Table 46).
- Three studies compared a treatment combining a low dose furosemide plus a low dose dopamine to a high dose furosemide treatment (Chen et al, 2013 – ROSE trial^{35,36}, Giamouzis et al, 2010 – DAD-HF trial^{88,89} and Triposkiadis et al, 2014 - DAD-HF II^{214,215}) (summarised in Table 47).

Vasopressor evidence

- No evidence was identified that investigated the clinical efficacy of vasopressors versus usual care.

Inotrope versus vasopressor evidence

- One study compared dopamine and norepinephrine in patients hospitalised for shock (De Backer et al, 2010⁵⁴). This RCT had a planned subgroup of patients with cardiogenic shock and was therefore included as the only inotrope versus vasopressor comparison.

Evidence is summarised in the GRADE evidence profiles below (Table 49, and Table 51). See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and lists of excluded studies in Appendix K.

Summary of included studies

Some of the characteristics of the included studies are summarised in the tables below which are divided by treatment comparisons (for details see study evidence tables in Appendix G).

Table 45: Summary of RCTs (milrinone vs. placebo)

Study	Population	Comparison	Outcomes	Comments
Cuffe et al, 2002 OPTIME-HF trial ^{50,51}	949 patients with acute exacerbation of chronic heart failure	Milrinone Intravenous infusion of milrinone administered, without an initial loading dose, at an initial infusion of 0.5 µg/kg/min and investigators were encouraged to continue this rate for 48 hours (dose could be adjusted up- or downwards. vs. placebo	<ul style="list-style-type: none"> • Mortality (during hospital stay and 60 day follow-up) • Length of hospital stay • Adverse events: arrhythmia (during hospital stay and 60 day follow-up) • Adverse events: myocardial infarction (during hospital stay and 60 day follow-up) 	<ul style="list-style-type: none"> • Population consisted of patients, for whom inotropic therapy was indicated but not, in the opinion of the investigators, essential.
Seino et al, 1996 ¹⁹⁷	52 patients with acute decompensated heart failure	Milrinone 50 µg/kg intravenous loading followed by 0.5 µg/kg/min continuous infusion of milrinone.. Duration 6 hrs. vs. placebo	<ul style="list-style-type: none"> • Number of patients improved on a 4 point scale using a combination of subjective ratings of symptoms (e.g. dyspnoea, palpitation) and ratings of physical findings (e.g. moist rales in the lung, gallop) • Adverse event: arrhythmia 	<ul style="list-style-type: none"> • This trial is a combination of two studies. One with the aim of finding the most effective dose of milrinone and the other part is the RCT comparison with placebo. Only the RCT part is analysed here. • The scale for number of patients improved is not a validated measure. • Main outcomes reported were not specified in our protocol: cardiovascular hemodynamic measurements – not reported below.
Triposkiadis et al, 2014 – DAD-HF II trial ^{88,89}	111 patients with acute decompensated heart failure	Low dose dopamine combined with low dose furosemide - 40 mg intravenous furosemide bolus followed by continuous IV administration of 5 mg/h furosemide combined with 5 µg/kg-1 min-1 dopamines. vs. Low dose furosemide - 40 mg intravenous furosemide bolus followed by high dose furosemide 10 mg/h	<ul style="list-style-type: none"> • Mortality (in hospital and at 60 day follow-up) • Rehospitalisation within 60 days • Dyspnoea score (Borg index) • Renal function (total urine volume in ml, incidence of patients with an above 0.3 mg/dL rise in serum creatinine, incidence of patients with an above 20% decrease in eGFR) 	<ul style="list-style-type: none"> • In this study dopamine was used for renal effects. • This study also included a high dose furosemide comparator, but since the low dose corresponds to UK this comparator was not prioritised.

Study	Population	Comparison	Outcomes	Comments
		continuous IV administration. Duration 8 hours.		

Table 46: Summary of RCTs (dobutamine vs. placebo)

Study	Population	Comparison	Outcomes	Comments
Adamopoulos et al, 2006 ⁵	46 patients with acutely decompensated chronic heart failure	Dobutamine - Continuous 24-hour infusion initially a 5 µg/kg/min without a loading dose; if a symptomatic reduction was not achieved after 2 hours the rate of dobutamine infusion was gradually doubled vs. placebo	<ul style="list-style-type: none"> • Mortality 	<ul style="list-style-type: none"> • The main focus of the study was a comparison between levosimendan with dobutamine. • The primary outcomes were not those in our protocol (such as echocardiographic and hemodynamic parameters as well as proinflammatory and proapoptotic markers) – not reported below
Zairis et al, 2004 CASINO trial as described in Cleland 2004 ^{42,43}	299 patients with decompensated heart failure of which 193 were in the dobutamine vs. placebo comparison	Dobutamine - Infused intravenously - placebo bolus then 10 mcg/kg per min vs. placebo	<ul style="list-style-type: none"> • Mortality (at 1 and 6 months follow-up) 	<ul style="list-style-type: none"> • The CASINO trial has never reported in full • The main focus of the study was a comparison between levosimendan with dobutamine. • The study was stopped early due to a clear mortality benefit in favour of levosimendan.

Table 47: Summary of RCTs (dopamine with furosemide vs. medical care including furosemide)

Study	Population	Comparison	Outcomes	Comments
Chen et al, 2013, ROSE trial ^{35,36}	241 patients hospitalised for acute heart failure who had renal dysfunction on admission	Dopamine (2 µg/kg/min for 72 hrs) was added to furosemide (intravenous loop diuretic treatment with a recommended total daily dose equal to 2.5 times the total daily oral outpatient furosemide [or equivalent] dose up to a maximum of 600mg/d) compared to medical care plus loop diuretic as described above	<ul style="list-style-type: none"> • Mortality • Self-rated clinical status (dyspnoea and global well-being scores) • Total urine volume • Worsening renal function • Serious adverse events 	<ul style="list-style-type: none"> • The trial included a nesiritide arm which is not analysed for this review.
Giamouzis et al, 2010 – DAD-HF	60 patients with acute decompensated	Low dose dopamine combined with low dose furosemide (40 mg	<ul style="list-style-type: none"> • Mortality (in hospital and at 60 day follow-up) 	<ul style="list-style-type: none"> • High dose furosemide was not a

Study	Population	Comparison	Outcomes	Comments
trial ^{88,89}	heart failure	intravenous furosemide bolus followed by continuous IV administration of 5 mg/h furosemide combined with 5 µg/kg/min dopamine) versus. high dose furosemide – (40 mg intravenous furosemide bolus followed by high dose furosemide 20 mg/h continuous IV administration). Duration 8 hours.	<ul style="list-style-type: none"> Rehospitalisation within 60 days Dyspnoea score (Borg index) Renal function (total urine volume in ml, incidence of patients with an above 0.3 mg/dL rise in serum creatinine, incidence of patients with an above 20% decrease in eGFR) 	comparator specified in the protocol.
Tripodskiadis et al, 2014 – DAD-HF II trial ^{88,89}	111 patients with acute decompensated heart failure	Low dose dopamine combined with low dose furosemide (40 mg intravenous furosemide bolus followed by continuous IV administration of 5 mg/h furosemide combined with 5 µg/kg-1 min-1 dopamines) versus. Low dose furosemide (40 mg intravenous furosemide bolus followed by high dose furosemide 10 mg/h continuous IV administration). Duration 8 hours.	<ul style="list-style-type: none"> Mortality (in hospital and at 60 day follow-up) Rehospitalisation within 60 days Dyspnoea score (Borg index) Renal function (total urine volume in ml, incidence of patients with an above 0.3 mg/dL rise in serum creatinine, incidence of patients with an above 20% decrease in eGFR) 	<ul style="list-style-type: none"> This study also included a high dose furosemide comparator, but since the low dose corresponds to UK this comparator was not prioritised.

Table 48: Summary of randomised controlled trials (Dopamine vs. norepinephrine in a subgroup of patients with cardiogenic shock)

Study	Population	Comparison	Outcomes	Comments
De Backer et al, 2010 ⁵⁴	1679 patients with shock of which 280 patients presented with cardiogenic shock	Dopamine - Dopamine. 2 mcg/kg min with the same increments up to a maximal dose of 20 mcg/kg min. vs. Norepinephrine - 0.02 mcg/kg min with the same increments up to a maximal dose of 0.19 mcg/kg min.	<ul style="list-style-type: none"> Mortality (28 day follow-up) 	<ul style="list-style-type: none"> The source of cardiogenic shock was not acute heart failure for all patients in the subgroup. Of the 280 patients 39 had cardiogenic shock after cardiopulmonary bypass and 9 due to tamponade.

Table 49: GRADE profile - milrinone vs. placebo

Quality assessment							Summary of Findings (most recent meta-analysis)				Quality	Importance
							Milrinone Number of event / Total N (%) or Mean (SD)	Placebo Number of event / Total N (%) or Mean (SD)	Effect			
									Relative Risk (95% CI)	Absolute effect / Mean difference (MD) (95% CI)		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other						
In hospital mortality ^{50,51}												
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	18/477 (3.8%)	11/472 (2.3%)	RR 1.62 (0.77 to 3.39)	14 more per 1000 (from 5 fewer to 56 more)	VERY LOW	CRITICAL
Mortality – 60 day follow-up ^{50,51}												
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	49/474 (10.3%)	41/463 (8.9%)	RR 1.17 (0.79 to 1.73)	15 more per 1000 (from 19 fewer to 65 more)	VERY LOW	CRITICAL
Number of people improved (subjective rating of symptoms: e.g. dyspnoea, palpitation; and physical findings: e.g. moist rales in the lung, gallop) expressed on a 4-point scale ¹⁹⁷												
1	RCT	Very serious ^(a, c))	No serious inconsistency	No serious indirectness	No serious imprecision	None	11/27 (40.7%)	0/25 (0%)	OR 10.98 (2.93 to 41.05)	41 more per 1000 (from 22 to 60 more)	LOW	IMPORTANT
Length of hospital stay ^{50,51}												
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	No serious imprecision	None	7 (6.2)	7 (6.6)		MD 0 higher (0.81 lower to 0.81 higher)	MODERATE	IMPORTANT
Adverse event: arrhythmia – during hospitalisation ^{51,197}												
2	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	No serious imprecision	None	26/504 (5.2%)	7/497 (1.4%)	RR 3.47 (1.56 to 7.74)	17 more per 1000 (from 4 more to 47 more)	MODERATE	CRITICAL

Quality assessment							Summary of Findings (most recent meta-analysis)				Quality	Importance
							Milrinone Number of event / Total N (%) or Mean (SD)	Placebo Number of event / Total N (%) or Mean (SD)	Effect			
									Relative Risk (95% CI)	Absolute effect / Mean difference (MD) (95% CI)		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other						
								0.7%				
Adverse event: arrhythmia – during follow-up ^{50,51}												
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Serious imprecision ^(b)	None	26/462 (5.6%)	16/446 (3.6%)	RR 1.57 (0.85 to 2.88)	20 more per 1000 (from 5 fewer to 67 more)	LOW	CRITICAL
Adverse event: myocardial infarction – during hospitalisation ^{50,51}												
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	7/477 (1.5%)	2/472 (0.42%)	RR 3.46 (0.72 to 16.59)	10 more per 1000 (from 1 fewer to 66 more)	VERY LOW	CRITICAL
Adverse event: myocardial infarction – during follow-up ^{50,51}												
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	10/462 (2.2%)	5/448 (1.1%)	RR 1.94 (0.67 to 5.63)	10 more per 1000 (from 4 fewer to 52 more)	VERY LOW	CRITICAL

- (a) The RCTs varied in methodological quality. Each study had at least 1 serious limitations. None of the studies described clear allocation concealment. Both studies were said to be double blinded. Outcomes were covered mostly by one of the two studies, and the overall risk of bias for each outcome was assessed according to the risk of bias for the majority of the evidence (according to the weight of the study in the meta-analysis).
- (b) If the CI for the meta-analysis was consistent with an appreciable clinically beneficial effect (or harmful as the case may be) of the intervention and an effect that indicates no clear clinical advantage, imprecision was graded as serious; if the CI was consistent with both appreciable clinical benefit and an appreciable clinical harm, then imprecision was graded as very serious.
- (c) The scale used to measure this outcome was not independently validated.

Table 50: GRADE profile - dobutamine vs. placebo

Quality assessment	Summary of Findings	Quality	Importance
--------------------	---------------------	---------	------------

							Dobutamine Number of event / Total N (%) or Mean (SD)	Placebo Number of event / Total N (%) or Mean (SD)	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			Relative Risk (95% CI)	Absolute effect or Mean Difference (MD) (95% CI)		
In-hospital mortality ⁵												
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	5/23 (21.7%)	4/23 (17.4%)	RR 1.25 (0.38 to 4.07)	43 more per 1000 (from 108 fewer to 534 more)	VERY LOW	CRITICAL
Mortality – 1 month follow-up ^{42,43}												
1	RCT	Very serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	12/96 (12.5%)	8/97 (8.2%)	RR 1.52 (0.65 to 3.54)	43 more per 1000 (from 29 fewer to 209 more)	VERY LOW	CRITICAL
Mortality - 6 months follow-up ^{42,43}												
1	RCT	Very serious ^(a)	No serious inconsistency	No serious indirectness	Serious imprecision ^(b)	None	30/76 (39.5%)	19/77 (24.7%)	RR 1.6 (0.99 to 2.58)	148 more per 1000 (from 2 fewer to 390 more)	VERY LOW	CRITICAL

- (a) The RCTs varied in methodological quality. One of them did not describe clear allocation concealment whereas the other study was only described as a summary and it is therefore methodologically unclear. Each outcome was covered by one of the two studies, and the overall risk of bias for each outcome was assessed according to the risk of bias for this study.
- (b) If the CI for the meta-analysis was consistent with an appreciable clinically beneficial (or harmful as the case may be) effect of the intervention and an effect that indicates no clear clinical advantage, imprecision was graded as serious; if the CI was consistent with both appreciable clinically benefit and an appreciable clinically important harm, then imprecision was graded as very serious.

Table 51: GRADE clinical evidence profile – dopamine/furosemide vs. HD furosemide

Quality assessment	Summary of Findings			Quality	Importance
	Dopamine/F	Furosem	Effect		

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	urosemide Number of event / Total N (%) or Mean (SD)	ide Number of event / Total N (%) or Mean (SD)	Relative Risk (95% CI)	Absolute effect or Mean Difference (MD) (95% CI)		
All-cause mortality at 180 days (time to event) ^{35,36}												
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	24/122 (19.7%)	25/119 (21%)	HR 0.95 (0.54 to 1.67)	9 fewer per 1000 (from 91 fewer to 115 more)	VERY LOW	CRITICAL
All-cause mortality - in-hospital ^{89,214}												
2	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	1/86 (1.2%)	3/85 (3.5%)	RR 0.33 (0.04 to 3.05)	18 fewer per 1000 (from 26 fewer to 55 more)	VERY LOW	CRITICAL
								2.7%				
All-cause mortality – 60 day follow-up ^{89,214}												
2	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	7/86 (8.1%)	7/85 (8.2%)	RR 0.99 (0.36 to 2.7)	1 fewer per 1000 (from 55 fewer to 146 more)	VERY LOW	CRITICAL
All-cause mortality – 1 year follow-up ^{214,215}												
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	7/86 (8.1%)	7/85 (8.2%)	RR 0.99 (0.36 to 2.7)	1 fewer per 1000 (from 55 fewer to 146 more)	VERY LOW	CRITICAL
CV mortality - in-hospital ^{89,214}												

Quality assessment							Summary of Findings				Quality	Importance
							Dopamine/Furosemide	Furosemide	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Relative Risk (95% CI)	Absolute effect or Mean Difference (MD) (95% CI)		
2	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	0/86 (0%)	3/85 (3.5%)	RR 0.14 (0.01 to 2.66)	1 fewer per 1000 (from 55 fewer to 146 more)	VERY LOW	CRITICAL
CV Mortality – 60 day follow-up ^{89,214}												
2	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	4/86 (4.7%)	7/85 (8.2%)	RR 0.57 (0.17 to 1.86)	37 fewer per 1000 (from 71 fewer to 74 more)	VERY LOW	CRITICAL
								8.6%				
CV Mortality – 1 year follow-up ^{214,215}												
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	17/56 (30.4%)	14/55 (25.5%)	RR 1.19 (0.65 to 2.18)	48 more per 1000 (from 89 fewer to 300 more)	VERY LOW	CRITICAL
Heart failure hospitalisation within 60 days ^{89,214}												
2	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Serious imprecision ^(b)	None	16/86 (18.6%)	10/85 (11.8%)	RR 1.58 (0.76 to 3.26)	61 more per 1000 (from 25 fewer to 240 more)	LOW	IMPORTANT
								10.6%				
Heart failure hospitalisation within 1 year ^{214,215}												

Quality assessment							Summary of Findings				Quality	Importance
							Dopamine/Furosemide	Furosemide	Effect			
									Relative Risk (95% CI)	Absolute effect or Mean Difference (MD) (95% CI)		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)				
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	28/56 (50%)	26/55 (47.3%)	RR 1.06 (0.72 to 1.55)	28 more per 1000 (from 132 fewer to 260 more)	VERY LOW	IMPORTANT
Length of hospital stay - days ^{89,214}												
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Serious imprecision ^(b)	None	Mean (sd) 6.1 (3) N=30	Mean (sd) 5.3 (2.4) N=30		MD 0.8 higher (0.57 lower to 2.17 higher)	VERY LOW	IMPORTANT
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Unclear / n/a ^(c)	None	Median 4.5 (IQR 3-5.75) N=56	Median 4 (IQR 3-5) N=55		P=0.342 ^(c)	VERY LOW	IMPORTANT
Total urine volume (ml) – (measured at 7-8 hours and at 72 hours) ^{36,89}												
2	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	No serious imprecision	None	-2,230 (1,485) N=30; -8,524 (3,386) N=122	-2,176 (1,193) N=30; -8,296 (2,941) N=119		MD 127.16 lower (646.07 lower to 391.76 higher)	MODERATE	IMPORTANT
Global well-being score (AUC ^(d)) at 72 hours – Better indicated by lower values ^{35,36}												

Quality assessment							Summary of Findings				Quality	Importance
							Dopamine/Furosemide	Furosemide	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Relative Risk (95% CI)	Absolute effect or Mean Difference (MD) (95% CI)		
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	No serious imprecision	None	-4,553 (1,384) N=122	-4,704 (1,438) N=119		MD 151 higher (205.39 lower to 507.39 higher)	MODERATE	IMPORTANT
Dyspnoea score (Borg index and AUC ^(d)) up to 72 hours – Better indicated by lower values ^{88,89}												
1	RCTs	Serious ^(a)	No serious inconsistency	No serious indirectness	Serious imprecision ^(b)	None	2.5 (1.3) N=30; -4,936 (1,534) N=122	2.8 (1.8) N=30; -4,998 (1,509) N=119		SMD 0.0 lower (0.23 lower to 0.22 higher)	VERY LOW	IMPORTANT
Worsening renal function – incidence of patients >0.3 mg/dL rise in serum creatinine level ^{36,89}												
3	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Serious imprecision ^(b)	None	31/208 (14.9%)	37/204 (18.1%) 18.6%	RR 0.82 (0.53 to 1.27)	33 fewer per 1000 (from 87 fewer to 50 more)	VERY LOW	IMPORTANT
Worsening renal function – incidence of patients with > 20% decrease in eGFR ^{89,214}												
2	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	8/86 (9.3%)	14/85 (16.5%)	RR 0.57 (0.25 to 1.28)	87 fewer per 1000 (from 152 fewer to 57 more)	VERY LOW	IMPORTANT

Quality assessment							Summary of Findings				Quality	Importance
							Dopamine/Furosemide	Furosemide	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Relative Risk (95% CI)	Absolute effect or Mean Difference (MD) (95% CI)		
								20.3%				
Rate of serious adverse events - total ^{89,214}												
2	RCT	Serious ^(a)	no serious inconsistency	no serious indirectness	serious ¹	none	38/178 (21.3%)	30/174 (17.2%)	RR 1.24 (0.81 to 1.90)	41 more per 1000 (from 33 fewer to 155 more)	LOW	
								12.5%				
Rate of serious adverse events in hospital ^{214,215}												
1	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	8/56 (14.3%)	6/55 (10.9%)	RR 1.31 (0.49 to 3.53)	34 more per 1000 (from 56 fewer to 276 more)	VERY LOW	CRITICAL
Rate of serious adverse events at 60 days ^{35,36}												
1:	RCT	Serious ^(a)	No serious inconsistency	No serious indirectness	Serious imprecision ^(b)	None	30/122 (24.6%)	24/119 (20.2%)	RR 1.22 (0.76 to 1.96)	44 more per 1000 (from 48 fewer to 194 more)	LOW	CRITICAL

(a) None of the studies describe allocation concealment.

(b) If the CI for the meta-analysis was consistent with an appreciable clinically beneficial (or harmful as the case may be) effect of the intervention and an effect that indicates no clear clinical advantage, imprecision was graded as serious; if the CI was consistent with both appreciable clinically benefit and an appreciable clinically important harm, then imprecision was graded as very serious

(c) Medians and interquartile ranges are reported. It is therefore difficult to say what the level of imprecision and effect size is.

For the global well-being and dyspnoea scale the ROSE trial used a visual analogue scale ranging from 0-100. To account for measurements taken at different time points the authors used the area under the curve (AUC) with time on the x-axis and scale on the y-axis. Here converted into a negative number to bring it in line with other outcomes. Not to be confused with the area under the curve in receiver operating characteristics curves where values can only range from 0 to 1.

Dopamine versus norepinephrine

De Backer et al, 2010⁵⁴

Mortality (28-day follow-up):

In this RCT, dopamine was compared with norepinephrine for the treatment of patients with shock. They carried out a planned subgroup analysis in patients with cardiogenic shock. They describe that among those with cardiogenic shock (N = 135 in the dopamine group and N = 145 in the norepinephrine group) the rate of mortality was significantly lower in the norepinephrine group ($p = 0.03$). A forest plot of the hazard ratio and a Kaplan-Meier curve was provided, but no clear data could be extracted from those figures. VERY LOW QUALITY evidence.

7.5.2 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs were provided to aid consideration of cost-effectiveness. These are given in Appendix N.

7.5.3 Evidence statements

7.5.3.1 Clinical

Inotropes versus placebo

Milrinone

Mortality

In very low quality evidence from one study comprising 949 patients with acute exacerbation of chronic heart failure, higher rates of mortality in patients treated with milrinone compared to placebo were reported. This was the case during hospital stay and at 60-day follow-up. However, there was considerable uncertainty which makes it impossible to draw clear conclusions about clinical harm associated with milrinone.

Number of people improved

(Subjective rating of symptoms: e.g. dyspnoea, palpitation; and physical findings: e.g. moist rales in the lung, a gallop rhythm; expressed in a 4-point scale)

Low quality evidence from one study with 52 patients with acute heart failure indicated a higher proportion of patients rated as improved in the milrinone group compared to the placebo group.

Length of hospital stay

Moderate quality evidence from one study with 949 participants showed no clear advantage in the length of hospital stay associated with milrinone.

Adverse events: arrhythmia

Moderate quality evidence from two studies with 1001 participants showed a higher proportion of arrhythmias associated with milrinone during hospital stay. However, at 60-day follow-up, no clear pattern was observed with regard to arrhythmias associated with milrinone (low quality evidence).

Adverse events: myocardial infarction (MI)

Evidence from one study comprising 949 participants with acute exacerbations of chronic heart failure indicated a higher proportion of MI in patients receiving milrinone (very low quality evidence). This was the case during hospital stay and at 60-day follow-up. However, there was much uncertainty around this effect that it was unclear whether this constituted a clinically harmful effect of milrinone.

Dobutamine

Mortality

Very low quality evidence from two studies with 239 participants indicated higher rates of mortality associated with dobutamine. However, there was too much uncertainty to draw any clear conclusions about clinical harm of dobutamine during hospital stay, after one month and at 6 month follow-up.

Low dose dopamine with low dose furosemide vs. medical care including furosemide

Mortality

Very low quality evidence from three randomised controlled trials (RCTs) comprising 412 patients with acute decompensated heart failure (ADHF) reported similar rates of mortality in the dopamine / low dose furosemide and the medical care (including furosemide) groups. This was the case for in-hospital, 60-day, 180-day and 1-year as well as cardiovascular mortality at these follow-up times.

Heart failure hospitalisations

Two RCTs comprising 171 patients reported higher rates of rehospitalisation associated with the group receiving low dose dopamine / low dose furosemide compared to high dose furosemide at 60-day follow-up (low quality evidence). However, rates of heart failure hospitalisations were similar in the two groups at 1-year follow-up (very low quality evidence).

Global well-being (visual analogue scale)

In one RCT with 241 participants there was no clear difference in self-rated global well-being between people who received dopamine in combination with furosemide and those who received medical care including furosemide (moderate quality evidence) at 72-hour follow-up.

Dyspnoea rating (Borg index and visual analogue scale)

In two RCTs with 241 participants, those on low dose dopamine / low dose furosemide rated themselves no more improved in dyspnoea (according to the Borg index and a visual analogue scale) than those on medical care including furosemide (very low quality evidence).

Renal functions

Total urine volume (ml):

Very low quality evidence from two RCTs comprising 301 people with ADHF showed similar levels of urine volume in the low dose dopamine / low dose furosemide and the medical care including furosemide groups at 8- and 72-hour follow-up (moderate quality evidence).

Worsening renal function:

- Incidence of patients with a higher than 0.3 mg/dL rise in serum creatinine level: In three RCTs with 412 participants, similar proportions of patients in both groups showed a higher than 0.3 mg/dL increase in serum creatinine.
- Incidence of patients with a higher than 20% decrease in estimated glomerular filtration rate (eGFR): In two RCTs with 171 participants a lower proportion of patients showed a higher than 20% decrease in eGFR in the low dose dopamine / low dose furosemide group compared to those receiving medical care including furosemide. However, considerable uncertainty makes it unclear whether this is an appreciable benefit (very low quality evidence).

Serious adverse events

Two RCTs with a total of 352 participants indicated similar rates of serious adverse events in the dopamine with furosemide group and the medical care including furosemide group in hospital and at 60-day follow-up (very low / low quality evidence).

Length of hospital stay

Two RCTs comprising 171 patients showed that the patients in the low dose dopamine / low dose furosemide group and the patients in the medical care (including furosemide) group had similar length of hospital stay (very low quality evidence).

Inotrope versus vasopressor

Dopamine vs. norepinephrine

Mortality

Very low evidence from one study with 280 patients with cardiogenic shock suggested lower mortality rates in those receiving norepinephrine compared to those receiving dopamine. However,

due to the way in which it was reported, there is considerable uncertainty around this effect and does not allow clear conclusions about appreciable benefits of norepinephrine over dopamine.

7.5.3.2 Economic

No relevant economic evaluations were identified.

7.5.4 Recommendations and link to evidence

Recommendations	<p>19. Do not routinely offer inotropes or vasopressors to people with acute heart failure.</p> <p>20. Consider inotropes or vasopressors in people with acute heart failure with potentially reversible cardiogenic shock. Administer these treatments in a cardiac care unit or high dependency unit or an alternative setting where at least level 2 care^f can be provided.</p>
Relative values of different outcomes	The GDG was most interested in mortality, but also considered adverse events such as myocardial infarction and arrhythmia to be important, and renal function in the context of one study that had used dopamine in conjunction with a diuretic. The GDG also looked at length of hospital stay, rehospitalisation and the Borg dyspnoea index.
Trade-off between clinical benefits and harms	The evidence reviewed did not demonstrate any sustained benefit from the use of vasopressors or inotropes in acute heart failure. Conducting trials in this area is very difficult due to the severely unwell patient group and the 'rescue' nature of therapy. There is a trend to harm in terms of increased mortality, myocardial infarction and arrhythmia associated with their use. However, the data are relatively weak in the acute heart failure setting. The GDG noted that use of vasopressors and/or inotropes may be most appropriate in rescuing patients from life-threatening systemic hypoperfusion, in order to allow other therapies to act and address potentially reversible causes.
Economic considerations	No economic evidence on nitrates or sodium nitroprusside was identified. Unit costs were presented to the GDG, and other costs such as drug administration were considered. No evidence of clinical benefit was identified to justify the increased cost associated with routine use of inotropes or vasopressors.
Quality of evidence	Most evidence was very low quality according to GRADE criteria. One of the included studies has never been published (Zairis et al, 2004 CASINO trial) as described in Cleland 2004 ^{42,43}). Following discussion, it was decided this evidence should remain included due to the importance of the mortality outcome that was reported. The Seino et al., (1996) study ¹⁹⁷ used an unvalidated scale to measure subjective and objective improvement, so this study was given less weight in the discussion. The De Backer et al., (2010) ⁵⁴ study reported significantly reduced mortality associated with norepinephrine compared to dopamine in a subgroup of people with cardiogenic shock. However, this evidence was downgraded using the GRADE criteria, since only limited data on this sub-group were reported. The largest study included in the

^f Patients requiring more detailed observation or intervention, including support for a single failing organ system or post-operative care and those stepping down from higher levels of care (Intensive Care Society, Levels of Critical Care for Adult Patients, 2009).

Recommendations	<p>19. Do not routinely offer inotropes or vasopressors to people with acute heart failure.</p> <p>20. Consider inotropes or vasopressors in people with acute heart failure with potentially reversible cardiogenic shock. Administer these treatments in a cardiac care unit or high dependency unit or an alternative setting where at least level 2 care^f can be provided.</p>
	<p>review (OPTIME-HF trial – Cuffe et al, 2002) recruited patients for whom inotropic therapy was indicated but not, in the opinion of the investigators, essential. As this is not the population that would usually receive such medication in the UK, the applicability of this evidence was questioned. The GDG noted that one of the outcomes used in this study – number of patients improved – was an unusual composite outcome.</p>
Other considerations	<p>In order to administer inotropes and vasopressors, it is necessary to ensure that the level of care is appropriate. This should include an adequate nurse-patient ratio, the facility for invasive monitoring and rapid access to higher levels of care if required. The GDG agreed this may be achieved most easily in a level 2 care setting (see glossary).</p> <p>There is only limited evidence available assessing dobutamine in acute heart failure, and this suggests it is associated with higher longer term mortality, but the results from the CASINO trial^{42,43} were never fully published. The GDG noted that in chronic heart failure there is more robust evidence of harm of dobutamine, and this is relevant to considering its role in acute heart failure.</p> <p>The GDG noted that in some trials (ROSE^{35,36}, DAD-HF^{88,89} and the DAD-HF II^{214,215}), dopamine was used mainly for its diuretic effect rather than as a vasopressor or an inotrope. Combined with low dose furosemide, dopamine allowed equivalent diuresis to higher doses of furosemide alone; but the effects on renal function were similar. There was no difference in mortality, but an increase in rehospitalisation was reported in the dopamine group. The GDG considered that low dose dopamine might be appropriate to assist diuresis in certain patients, but recognised that a stronger evidence base was required, so recommended a further trial is carried out to address this question.</p> <p>The GDG discussed the use of dopamine as a vasopressor compared to norepinephrine (de Backer et al. 2010). The relevant part of the study was in a small group of patients with cardiogenic shock which is not representative of acute heart failure. Norepinephrine appeared to be associated with improved survival compared to dopamine, but there is significant uncertainty around this effect. The GDG therefore was not able to recommend one agent over the other. The GDG noted that when assessing a patient with suspected cardiogenic shock, the patient's usual blood pressure should be taken into consideration.</p> <p>The GDG did not review the evidence for levosimendan as it does not currently have a UK licence.</p> <p>The GDG noted that the majority of patients in the UK presenting with acute heart failure will have an acute decompensation of chronic heart failure. As such, the patient members emphasised that it is important, in common with all intensive and potentially harmful treatments, that inotropes and vasopressors should only be commenced if this is in accordance with any advance treatment directive the patient may have.</p> <p>The GDG concluded that the evidence base for the use of inotropes and vasopressors</p>

Recommendations	<p>19. Do not routinely offer inotropes or vasopressors to people with acute heart failure.</p> <p>20. Consider inotropes or vasopressors in people with acute heart failure with potentially reversible cardiogenic shock. Administer these treatments in a cardiac care unit or high dependency unit or an alternative setting where at least level 2 care^f can be provided.</p>
	<p>in acute heart failure is poor and that further research is welcomed. Current use of these agents should be limited to stabilising a patient in cardiogenic shock so as to allow time for reversible causes (e.g. myocardial ischaemia, arrhythmia, structural valve disease) to be identified and corrected or treated.</p>

8 Initial non-pharmacological treatment

8.1 Ventilatory support

Non-invasive ventilation is a potential alternative to invasive ventilation to assist some patients with respiratory distress. Patients with acute heart failure resulting in acute pulmonary oedema are frequently offered non-invasive ventilation. However, whether or not non-invasive ventilation impacts on patients' mortality, the need for invasive ventilation and the length of hospital stay is unclear and is investigated in this review.

Review question: In people with confirmed acute heart failure and cardiogenic pulmonary oedema is non-invasive positive pressure ventilation (CPAP and/or bilevel NIPPV) more clinical and cost effective than standard medical care alone to improve outcome?

For full details see review protocol in Appendix C. Please see Table 52 for a breakdown of the population, intervention/s, the comparator/s, outcomes, and study design.

Table 52: PICO characteristics of review question

Population	Adults with acute heart failure and cardiogenic pulmonary oedema.
Intervention/s	Continuous positive airway pressure (CPAP) or bilevel positive airway pressure (BiPAP)
Comparison/s	Standard medical care (any form of standard medical care provided for the management of cardiogenic pulmonary oedema excluding non-invasive positive pressure ventilation (NIPPV) or alternative methods of ventilatory support e.g. oxygen by face mask, diuretics, nitrates, etc).
Outcomes	<ul style="list-style-type: none"> • Mortality (in-hospital and at the end of follow-up) • Myocardial infarction • Intubation rate • Length of hospital stay • Quality of Life
Study design	Systematic reviews, and randomised controlled trials

8.1.1 Clinical evidence

We searched for systematic reviews and randomised controlled trials (RCT) investigating the effectiveness of non-invasive positive pressure ventilation (NIPPV), which can be continuous positive airway pressure (CPAP) or bilevel positive airway pressure (BiPAP), compared against standard medical care (any standard medical care for the management of cardiogenic pulmonary oedema, excluding NIPPV or alternative methods of ventilatory support, such as diuretics and nitrates) usually accompanied by oxygen mask. Although a number of systematic reviews are available we focused on the most recent reviews of which one was a Cochrane review (Vital et al., 2013^{220,221} assessed as up-to-date April 2011) and the other systematic review (Weng et al., 2010²²⁸) included studies published before December 2009. A further systematic review was identified (Mariani et al., 2011), but was deemed not to be methodologically rigorous enough to be included as evidence (see exclusion list in Appendix K). Other older systematic reviews were double checked for completeness of included trials.

There were inconsistencies between the two systematic reviews (included / excluded studies and extracted numbers), it was therefore decided not to update the Cochrane review, but order all the individual trials of both meta-analyses and assess them against the inclusion criteria of the protocol. The characteristics of the published meta-analyses are briefly highlighted in Table 53. From the

searches conducted as well as from the two systematic reviews, 22 trials were included^{8,28,49,56,61,76,83,95,111,116,121,124,134,157,168,169,178,183,203,208,209,226}. The reference list contains three publications of one trial (3CPO trial^{91,93,95}) that was included in this review. The main characteristics of the RCTs are summarised in Table 54. The main aim of these studies was to assess whether non-invasive ventilation lowered mortality rate and the need for endotracheal intubation. Evidence is summarised in the clinical GRADE evidence profile below (Table 57).

Two studies (Plaisance et al., 2007¹⁷⁸ and Sharon et al., 2000²⁰³) had comparisons which did not directly satisfy the protocol (early vs. later CPAP, and BiPAP vs. high-dose intravenous isosorbide dinitrate, a comparator not deemed to be 'medical care', respectively). These trials were not entered into the overall meta-analysis, but were analysed separately (see Table 58 and Table 59) and were downgraded in GRADE due to the indirectness of the evidence.

See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and lists of excluded studies in Appendix K.

Summary of included studies

The characteristics of the included studies are summarised below.

Table 53: Summary of studies included in published meta-analyses

Study	No. of studies	Population	Sensitivity analyses / confounders investigated	Outcomes	Comments
Vital et al., 2013 ^{220,221}	23	Adults with acute cardiogenic pulmonary oedema – according to criteria of the American Heart Association (AHA) and/ or the European Society of Cardiology (ESC) or based on symptoms and clinical signs indicative of acute heart failure	<ul style="list-style-type: none"> Setting, Hypercapnic patients, Patients with SP ≥ 14 cmH₂O with difference between SP and PEEP ≥ 8 cm H₂O and PEEP ≥ 10 cmH₂O 	<p><i>Primary outcomes:</i></p> <ul style="list-style-type: none"> Hospital mortality <p><i>Secondary outcomes:</i></p> <ul style="list-style-type: none"> Intubation rate Hospital length of stay Incidence of MI Intolerance to allocated treatment Arterial blood gases Vital signs Treatment failure Adverse events 	A trial was included that duplicates another population (Lin et al. 1991); two studies were included that would not really match our protocol characteristics (Sharon 2000 and Thys 2002). Heterogeneity analysis is not clearly described, i.e. many factors described in the introduction but later unclear how the 'heterogeneity' was analysed or how stratified subgroups were selected.
Weng et al., 2010 ²²⁸	20		<p>Additional meta-analyses were carried out to explore the influence of the 3CPO trial:</p> <ul style="list-style-type: none"> Omitting each trial in turn and recalculating 	<ul style="list-style-type: none"> Mortality Intubation rate Incidence of new MI 	It could be argued that the depth of analysis meant the number of analysed outcomes was small. However, those would be

Study	No. of studies	Population	Sensitivity analyses / confounders investigated	Outcomes	Comments
			<p>the pooled effect size estimate to assess which trial influences the estimate the most</p> <ul style="list-style-type: none"> • Contribution of the 3CPO trial to the overall estimate of effect across the entire range of weightings • Cumulative meta-analyses on basis of trial quality and date of publication • Sensitivity analysis using allocation concealment as subgroup • Bayesian hierarchical meta-analysis to account for between trial heterogeneity 		considered to be the most critical outcomes.

In the following three tables the trials included in the current meta-analysis are briefly described. For additional details see evidence extraction tables in Appendix G.

Table 54: Summary of RCTs included in the current meta-analysis

Study	Setting	Comparison	Outcomes	% switched treatment	Comments
Agmy et al., 2009 ⁸	Intensive care unit or cardiac care unit	CPAP vs. BiPAP vs. medical care	<ul style="list-style-type: none"> • Intubation rate 	N/A	<ul style="list-style-type: none"> • Results only presented on the trial registry website • Only one outcome reported even though several outcomes were investigated • Only very limited information provided (baseline and methodological information)
Bersten et al., 1991 ²⁸	Emergency department and intensive care unit	CPAP vs. medical care	<ul style="list-style-type: none"> • Intubation rate • Stay in ICU • Hospital stay • In-hospital mortality 	None	<ul style="list-style-type: none"> • Method of randomisation poorly described
Crane et al., 2004 ⁴⁹	Emergency department	CPAP vs. BiPAP vs. medical care	<ul style="list-style-type: none"> • Mortality • Intubation rate 	1/20 (5%) patient from CPAP to BiPAP	<ul style="list-style-type: none"> • Only one patient received non-invasive ventilation beyond 2 hrs

Study	Setting	Comparison	Outcomes	% switched treatment	Comments
Delclaux et al., 2000 ⁵⁶	Intensive care unit	CPAP vs. medical care	<ul style="list-style-type: none"> • Mortality • Intubation rate 	4/16 (25%) patients from medical care switched to CPAP	<ul style="list-style-type: none"> • Patients were stratified by whether or not they had an underlying cardiac condition. • Only 6 (14%) of the 42 patients with a cardiac condition had an acute cardiac disease.
Ducros et al., 2011 ⁶¹	Pre-hospital setting	CPAP vs. medical care	<ul style="list-style-type: none"> • Mortality • Intubation rate • ICU length of stay 	Unclear (on admission to hospital treatment decisions were made by the cardiologist which was not determined by the study protocol)	<ul style="list-style-type: none"> • Mobile physician-staffed intensive care units (ICUs) not comparable to paramedic-staffed ambulances • Study was prematurely stopped due to low recruitment rate. • Relatively low average severity of condition.
Ferrer et al., 2003 ⁷⁶	Intensive care unit	BiPAP vs. medical care	<ul style="list-style-type: none"> • Mortality • Intubation rate 	Overall 2/54 (4%) in the medical care group received non-invasive ventilation (unclear whether they were patients with cardiogenic pulmonary oedema)	<ul style="list-style-type: none"> • Patients included if they had severe acute hypoxemic respiratory failure (only a subgroup of those were patients with cardiogenic pulmonary oedema) • Trial not included in Cochrane analysis
Frontin et al., 2011 ⁸³	Pre-hospital setting	CPAP vs. medical care	<ul style="list-style-type: none"> • Intubation rate • Mortality • Length of ICU stay • Length of hospital stay 	2/62 (3%) of those receiving CPAP switched to oxygen mask – it is unclear which treatment was administered upon admission to hospital	<ul style="list-style-type: none"> • Mobile physician-staffed ICUs not comparable to paramedic-staffed ambulances • Trial protocol completed on arrival to ICU and all patients were treated at the discretion of the attending admitting physician.
3CPO trial: Gray et al., 2009; Gray et al., 2008 and Goodacre et	Emergency department	CPAP vs. BiPAP vs. medical care	<ul style="list-style-type: none"> • Mortality • Intubation rate • Critical care admission 	Interventions were defined only for the first 2 hrs after which treatment was	<ul style="list-style-type: none"> • Largest trial of non-invasive ventilation to date (Health Technology Assessment) – good trial methodology

Study	Setting	Comparison	Outcomes	% switched treatment	Comments
al., 2011 ^{91,93,95}			<ul style="list-style-type: none"> Length of hospital stay 	at the discretion of clinicians. 19.4 % of all patients later received a different treatment to the one they were assigned to.	<ul style="list-style-type: none"> Early mortality not reported (only 5 and 7 day) Lower rates than in other studies of patients who had myocardial ischemia or infarction as the cause of pulmonary oedema.
Kelly et al., 2002 ^{110,111}	Emergency department	CPAP vs. medical care	<ul style="list-style-type: none"> Mortality rate Intubation rate Length of stay 	2/31 (6%) in medical care group received CPAP	<ul style="list-style-type: none"> None of the patients needed intubation, suggesting differences in the population group Baseline differences with patients assigned to CPAP having more severe disease with a slightly greater acidosis and hypercapnia on admission.
L'Her et al., 2004 ¹¹⁶	Emergency department	CPAP vs. medical care	<ul style="list-style-type: none"> Mortality Intubation rate Length of hospital stay 	10/46 (22%) in the medical care group later received non-invasive ventilation and 2/43 (5%) in the CPAP group received a different type of non-invasive pressure ventilation	<ul style="list-style-type: none"> Study population restricted to an elderly group with a mean age of 84 yrs
Levitt et al., 2001 ¹²¹	Emergency department	BiPAP vs. medical care	<ul style="list-style-type: none"> Mortality Intubation rate Length of hospital stay 	4/21 (19%) did not tolerate BiPAP treatment and received medical care	<ul style="list-style-type: none"> Data published as preliminary findings – study terminated early due to findings of another study which reported a higher myocardial infarction rate in the BiPAP group
Lin et al., 1995 ^{123,124}	Emergency department, during hospitalisation and long term follow-	CPAP vs. medical care	<ul style="list-style-type: none"> Mortality Intubation rate Length of hospital stay 	Unclear	<ul style="list-style-type: none"> One-year follow-up results reported

Study	Setting	Comparison	Outcomes	% switched treatment	Comments
	up				
Masip et al., 2000 ¹³⁴	Intensive care unit	BiPAP vs. medical care	<ul style="list-style-type: none"> • Mortality • Intubation rate • Length of hospital stay 	Unclear	<ul style="list-style-type: none"> • Rates of patient with previous myocardial infarction, chronic obstructive pulmonary disease and diabetes were higher in the control group
Nava et al., 2003 ¹⁵⁷	Emergency department	BiPAP vs. medical care	<ul style="list-style-type: none"> • Mortality • Intubation rate • Length of hospital stay • New Myocardial infarction 	Unclear	<ul style="list-style-type: none"> • Randomisation was balanced according to whether the patient had an admission PaCO₂ below or above 45mmHg • Comparatively high overall intubation rate > 20%
Park et al., 2001 ¹⁶⁸	Emergency department	CPAP vs. BiPAP vs. medical care	<ul style="list-style-type: none"> • Mortality • Intubation rate 	Unclear	<ul style="list-style-type: none"> • Relatively low pressure levels were used
Park et al., 2004 ^{168,169}	Emergency department	CPAP vs. BiPAP vs. medical care	<ul style="list-style-type: none"> • Mortality • Intubation rate • Length of hospital stay 	None	<ul style="list-style-type: none"> • There were more patients with previous myocardial infarction at baseline in the medical care group • Comparatively high intubation rate in medical care group (11/26 - 42%)
Räsänen et al., 1985 ¹⁸³	Intensive care unit	CPAP vs. medical care	<ul style="list-style-type: none"> • Mortality • Intubation rate 	Unclear (treatment on admission at discretion of in-hospital physician). In 7/10 BiPAP treatment was discontinued due to considerable clinical improvement	<ul style="list-style-type: none"> • Study was interrupted after 10 minutes of the assigned treatment and patients classified as treatment failures were removed from the study and subsequent respiratory therapy in these patients was determined by the physician in charge
Takeda et al., 1997 ²⁰⁹	Intensive care unit	CPAP vs. medical care	<ul style="list-style-type: none"> • Mortality • Intubation rate • Length of stay 	Unclear	<ul style="list-style-type: none"> • Comparatively high intubation rate in medical care group (6/15 - 40%)
Takeda et al., 1998 ^{208,209}	Cardiac care unit	CPAP vs. medical care	<ul style="list-style-type: none"> • Mortality • Intubation rate 	Unclear	<ul style="list-style-type: none"> • All patients had cardiogenic pulmonary oedema due to acute myocardial infarction or ischemia

Study	Setting	Comparison	Outcomes	% switched treatment	Comments
Weitz et al., 2007 ²²⁶	Pre-hospital	BiPAP vs. medical care	<ul style="list-style-type: none"> • Mortality • Length of stay 	Unclear	<ul style="list-style-type: none"> • Physician-staffed emergency team not comparable to paramedic-staffed pre-hospital teams

Table 55: Summary of RCTs comparing early against later CPAP included in the current meta-analysis

Study	Setting	Comparison	Outcomes	% switched treatment	Comments
Plaisance et al, 2007 ¹⁷⁸	Pre-hospital	Early CPAP vs. late CPAP	<ul style="list-style-type: none"> • Mortality • Intubation rate 	Unclear	<ul style="list-style-type: none"> • All participants received standard medical care after 15 minutes – reported in a separate GRADE table

Table 56: Summary of RCTs comparing BiPAP against high dose isosorbide dinitrate included in the current meta-analysis

Study	Setting	Comparison	Outcomes	% switched treatment	Comments
Sharon et al, 2000 ²⁰³	Pre-hospital (Israel)	BiPAP vs. High-dose intravenous isosorbide dinitrate	<ul style="list-style-type: none"> • mortality • Intubation rate 	Unclear	<ul style="list-style-type: none"> • Comparison not standard medical practice therefore included in another GRADE table • Study was prematurely terminated due to significant deterioration of patients in the BiPAP group

Table 57: GRADE clinical evidence profile – Non-invasive positive pressure ventilation (NIPPV) vs. medical care (evidence from RCTs – new meta-analysis)

Quality assessment							Summary of Findings						
							NIPPV Number of event / Total N (%) or Mean (SD)	Medical care Number of event / Total N (%) or Mean (SD)	Effect				
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations					Relative Risk (95% CI)	Absolute effect or Mean Difference (MD) (95% CI)	Quality
In-hospital mortality (including up to 7 day mortality) ^{28,49,56,61,76,83,95,111,116,121,124,134,157,168,169,183,208,209,226}													
19	RCTs	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	Publication bias ^(a)	129/1323 (9.8%)	131/930 (14.1%) 17.1%	RR 0.72 (0.57 to 0.90)	40 fewer per 1000 (from 14 fewer to 61 fewer)	MODERATE	CRITICAL	
30 day mortality ^{83,95}													
2	RCTs	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(b)	None	113/762 (14.8%)	67/429 (15.6%) 13.8%	RR 0.93 (0.7 to 1.23)	10 fewer per 1000 (from 41 fewer to 32 more)	LOW	CRITICAL	
Mortality by setting – Pre-hospital ^{61,83,226}													
3	RCTs	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	7/180 (3.9%)	9/172 (5.2%) 5%	RR 0.73 (0.28 to 1.92)	13 fewer per 1000 (from 36 fewer to 46 more)	LOW	CRITICAL	
Mortality by setting – Intensive care unit ^{56,76,134,183,209}													
5	RCTs	Serious ^(a)	No serious inconsistency	No serious indirectness	Serious ^(b)	None	12/92 (13%)	20/90 (22.2%) 20%	RR 0.58 (0.31 to 1.09)	84 fewer per 1000 (from 138 fewer to 18 more)	LOW	CRITICAL	
Mortality by setting – cardiac care unit ^{208,209}													

Quality assessment							Summary of Findings					
							NIPPV Number of event / Total N (%) or Mean (SD)	Medical care Number of event / Total N (%) or Mean (SD)	Effect			
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			Relative Risk (95% CI)	Absolute effect or Mean Difference (MD) (95% CI)	Quality	Importance
1	RCTs	Serious ^(a)	No serious inconsistency	No serious indirectness	Very serious ^(b)	None	1/11 (9.1%)	7/11 (63.6%)	RR 0.14 (0.02 to 0.98)	547 fewer per 1000 (from 13 fewer to 623 fewer)	VERY LOW	CRITICAL
Mortality by setting – Emergency department ^{49,56,95,111,116,121,124,157,168,169}												
9	RCTs	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	92/1020 (9%)	80/637 (12.6%)	RR 0.72 (0.54 to 0.97)	40 fewer per 1000 (from 4 fewer to 66 fewer)	HIGH	CRITICAL
								14.3%				
								24.5%				
Intubation rate by setting – overall ^{8,28,49,56,61,76,83,95,111,116,121,124,134,157,168,169,183,208,209}												
19	RCTs	Serious ^(a)	No serious inconsistency	No serious indirectness	No serious imprecision	None	87/1398 (6.2%)	132/961 (13.7%)	RR 0.52 (0.41 to 0.67)	118 fewer per 1000 (from 81 fewer to 145) fewer)	MODERATE	CRITICAL
								24.6%				
Intubation rate by setting – pre-hospital ^{61,83}												
2	RCTs	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious imprecision ^(b)	None	5/167 (3%)	9/162 (5.6%)	RR 0.54 (0.18 to 1.57)	25 fewer per 1000 (from 44 fewer to 31 more)	LOW	CRITICAL
								5.4%				
Intubation rate by setting – intensive care unit ^{8,56,76,134,183,209}												

Quality assessment							Summary of Findings					
							NIPPV Number of event / Total N (%) or Mean (SD)	Medical care Number of event / Total N (%) or Mean (SD)	Effect			
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations				Relative Risk (95% CI)	Absolute effect or Mean Difference (MD) (95% CI)	Quality
6	RCTs	Serious ^(a)	No serious inconsistency	No serious indirectness	No serious imprecision	None	24/180 (13.3%)	42/131 (32.1%)	RR 0.44 (0.29 to 0.69)	168 fewer per 1000 (from 93 fewer to 213 fewer)	MODERATE	CRITICAL
								30%				
Intubation rate by setting – cardiac care unit ^{208,209}												
1	RCT	Very serious risk of bias ^(a)	No serious inconsistency	No serious indirectness	No serious imprecision	None	2/11 (18.2%)	8/11 (72.7%)	RR 0.25 (0.07 to 0.92)	545 fewer per 1000 (from 58 fewer to 676 fewer)	VERY LOW	CRITICAL
Intubation rate by setting – emergency department ^{28,49,95,111,116,121,124,157,168,169}												
10	RCTs	Serious ^(a)	No serious inconsistency	No serious indirectness	No serious imprecision	None	56/1040 (5.4%)	73/657 (11.1%)	RR 0.59 (0.43 to 0.81)	80 fewer per 1000 (from 37 fewer to 111 fewer)	MODERATE	CRITICAL
								19.5%				
Incidence of new MI - CPAP ^{49,95,168,169}												
4	RCTs	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(b)	None	98/420 (23.3%)	98/424 (23.1%)	RR 1.01 (0.79 to 1.28)	1 more per 1000 (from 30 fewer to 40 more)	LOW	CRITICAL
								14.3%				
Incidence of new MI - BiPAP ^{49,95,121,157,168,169}												
6	RCTs	No serious risk of	No serious inconsistency	No serious indirectness	Serious ^(b)	None	115/498 (23.1%)	108/510 (21.2%)	RR 1.09 (0.87 to 1.37)	14 more per 1000 (from 21 fewer to	MODERATE	CRITICAL
								15.8%				

Quality assessment							Summary of Findings				Quality	Importance	
							NIPPV Number of event / Total N (%) or Mean (SD)	Medical care Number of event / Total N (%) or Mean (SD)	Effect				
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations					Relative Risk (95% CI)	Absolute effect or Mean Difference (MD) (95% CI)	
		bias									58 more)		
Length of hospital stay by setting - CCU (Better indicated by lower values) ^{123,124}													
1	RCTs	Serious ^(a))	No serious inconsistency	No serious indirectness	Very serious ^(b)	none	4 (3) N=50	4.5 (3.5) N=50	-		MD 0.5 lower (1.78 lower to 0.78 higher)	VERY LOW	IMPORTANT
Length of hospital stay by setting - ICU (Better indicated by lower values) ^{28,111,209}													
3	RCTs	Serious ^(a))	No serious inconsistency	No serious indirectness	No serious imprecision	none	1.2 (0.4) N=19; 0 N=27; 2.9 (1.1) N=15	2.7 (2) N=20; 0 N=31; 3.9 (1.1) N=15	-		MD 1.22 lower (1.81 to 0.63 lower)	MODERATE	IMPORTANT
Length of hospital stay by setting - HDU (Better indicated by lower values) ^{110,111}													
1:	RCTs	Serious ^(a))	No serious inconsistency	No serious indirectness	No serious imprecision	None	1.1 (0.2) N=27	0.4 (0.2) N=31	-		MD 0.7 higher (0.6 to 0.8 higher)	MODERATE	IMPORTANT
Length of hospital stay by setting - Total length of stay (Better indicated by lower values) ^{28,95,111,116,121,124,134,157,169}													
9 ^(d)	RCTs	Serious ^(a))	No serious inconsistency	No serious indirectness	No serious imprecision	None	8.7 (8.3) N=19; 11.4 (8.71) N=702; 13.7 (2) N=27; 12 (11) N=43; 7.3 (8) N=21; 8.5 (4.5) N=50; 14.2 (5)	7.9 (4.1) N=20; 10.5 (8.71) N=367; 15 (2.7) N=31; 9 (7) N=46; 8.1 (6.4) N=17; 9 (4.5) N=50; 14.3 (4)	-		MD 0.04 higher (0.51 lower to 0.58 higher)	MODERATE	

Quality assessment							Summary of Findings						
							NIPPV Number of event / Total N (%) or Mean (SD)	Medical care Number of event / Total N (%) or Mean (SD)	Effect				
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations					Relative Risk (95% CI)	Absolute effect or Mean Difference (MD) (95% CI)	Quality
							N=19; 5.4 (3) N=65; 10 (7) N=27	N=18; 5.1 (2.3) N=65; 12 (8) N=26					

- (a) The funnel plot was asymmetrical and this indicates the presence of publication bias.
- (b) If the CI for the meta-analysis was consistent with an appreciable clinically beneficial effect (or harmful as the case may be) of the intervention and an effect that indicates no clear clinical advantage, imprecision was graded as serious; if the CI was consistent with both appreciable clinically benefit and an appreciable clinical harm, then imprecision was graded as very serious.
- (c) The RCTs varied in methodological quality. The majority of studies had at least two serious limitations. All the studies were unblinded and only a few studies had adequate allocation concealment. Lack of blinding is not considered important in all reported outcomes because knowledge of the interventions is unlikely to affect the outcomes. Each outcome was covered by a different combination of studies, and the overall risk of bias for each outcome was assessed according to the risk of bias for the majority of the evidence (according to the weight of the study in the meta-analysis).
- (d) For length of stay in Gray 200,9 only means and a mean difference with 95% CI was reported. Standard deviations are therefore calculated according to the mean difference and its CI, and represent an approximation of the actual standard deviations of the study population.

Narrative report on quality of life

Quality of life (EQ-5D) data were reported in the 3-CPO trial^{91,95} for patients at 1, 3 and 6 months follow-up for each of the following: medical care, CPAP and BiPAP. Only overall response rates and percentages per group were provided so it was unclear what the total number of responders was in each group. No group differences in EQ-5D were reported (LOW QUALITY EVIDENCE).

Table 58: GRADE clinical evidence profile – Early vs. late continuous positive airway pressure (CPAP)

Quality assessment							Summary of Findings						
							Non-invasive positive pressure ventilation Number of event / Total N (%) or Mean (SD)	Medical care Number of event / Total N (%) or Mean (SD)	Effect				
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations					Relative Risk (95% CI)	Absolute effect or Mean Difference (MD) (95% CI)	Quality
In-hospital mortality ¹⁷⁸													
1	RCT	No serious risk of bias	No serious inconsistency	Serious indirectness ^(a)	Serious imprecision ^(b)	None	2/63 (3.2%)	8/61 (13.1%)	RR 0.24 (0.05 to 1.09)	100 fewer per 1000 (from 125 fewer to 12 more)		LOW	CRITICAL
Intubation rate ¹⁷⁸													
1	RCT	No serious risk of bias	No serious inconsistency	Serious indirectness ^(a)	Serious imprecision ^(b)	None	6/63 (9.5%)	16/61 (26.2%)	RR 0.36 (0.15 to 0.87)	168 fewer per 1000 (from 34 fewer to 223 fewer)		LOW	CRITICAL

(a) The comparison was early vs. later CPAP (after 15 minutes), and as this did not fully match the comparator in the protocol it was thus classified as indirect evidence and downgraded.

(b) If the CI for the meta-analysis was consistent with an appreciable clinically beneficial effect of the intervention and an effect that does not warrant this shows no clear clinical advantage, imprecision was graded as serious. If the CI was consistent with both appreciable clinically benefit and an appreciable clinically important harm, then imprecision was graded as very serious.

Table 59: GRADE clinical evidence profile – Bilevel positive airway pressure (BiPAP) vs. high dose intravenous isosorbide dinitrate

Quality assessment						Summary of Findings			Quality	Importance
						Non-invasive	Medical care	Effect		

No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	positive pressure ventilation Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Relative Risk (95% CI)	Absolute effect or Mean Difference (MD) (95% CI)		
In-hospital mortality²⁰³												
1	RCTs	Very serious ^(a)	No serious inconsistency	Serious indirectness ^(b)	Very serious imprecision ^(c)	None	2/20 (10%)	0/20 (0%)	Peto OR 7.79 (0.47 to 129.11)	100 more per 1000 (from 50 fewer to 250 more)	VERY LOW	CRITICAL
Intubation rate²⁰³												
1	RCTs	Very serious ^(a)	No serious inconsistency	Serious indirectness ^(b)	No serious imprecision	None	16/20 (80%)	4/20 (20%)	RR 4 (1.62 to 9.87)	600 more per 1000 (from 124 more to 1000 more)	VERY LOW	CRITICAL
Incidence of new MI²⁰³												
1	RCTs	Very serious ^(a)	No serious inconsistency	Serious indirectness ^(b)	No serious imprecision	None	11/20 (55%)	2/20 (10%)	RR 5.5 (1.39 to 21.71)	450 more per 1000 (from 39 more to 1000 more)	VERY LOW	CRITICAL

(a) Allocation was inadequately concealed.

(b) The comparison was BiPAP vs. high dose intravenous isosorbide dinitrate, and this does not fully match the comparator in the protocol and was thus classified as indirect evidence and downgraded.

(c) If the CI for the meta-analysis was consistent with an appreciable clinically beneficial effect of the intervention and an effect that does not warrant this shows no clear clinical advantage, imprecision was graded as serious. If the CI was consistent with both appreciable clinically benefit and an appreciable clinically important harm, then imprecision was graded as very serious.

8.1.2 Economic evidence

Published literature

One economic evaluation was identified with the relevant comparison and has been included in this review (Gray et al., 2009^{93,95}). This is summarised in the economic evidence profile below (Table 60). See also the full study evidence tables in Appendix H. See also the economic article selection flow chart in Appendix E.

Table 60: Economic evidence profile: non-invasive positive pressure ventilation (CPAP or BiPAP) vs. medical care alone

Study	Applicability	Limitations	Other comments	Incremental cost (£)	Incremental effects (QALYs)	Cost effectiveness (£/QALY)	Uncertainty
Gray et al., 2009 ^{93,95} (UK)	Partially applicable	Minor limitations	<ul style="list-style-type: none"> Population: patients presenting with severe acute cardiogenic pulmonary oedema (ACPO) in 26 emergency departments in the UK. Comparators: <ol style="list-style-type: none"> Standard oxygen therapy (minimum of 2 days of treatment) CPAP (minimum of 2 days of treatment) BiPAP (minimum of 2 days of treatment) Within trial cost-utility analysis (cost per QALYs gained) of 6 months and lifetime cost-utility model both with and without data imputation for missing values. Follow-up of 6 months with a lifetime extrapolation. 	<p>Lifetime without data imputation: 2 vs 1: 456 3 vs 1: 691</p> <p>Lifetime with data imputation: 2 vs 1: 1,761 3 vs 1: 1,257</p>	<p>Lifetime without data imputation: 2 vs 1: 0.174 3 vs 1: 0.008</p> <p>Lifetime with data imputation: 2 vs 1: 0.244 3 vs 1: 0.11</p>	<p>Lifetime without data imputation: 2 vs 1: 2,621 3 vs 1: 86,375</p> <p>Lifetime with data imputation: 2 vs 1: 7,217 3 vs 1: 11,427</p>	<p>At a willingness to pay of £20,000 the probability of CPAP being cost-effective compared to medical care were 71% and 74% without and with data imputation respectively. Under this approach the probability that CPAP (lifetime treatment effect without data imputation) is cost-effective was reduced from 71% to 63%.</p> <p>Using the RCT follow-up time (6 months) for the analysis had a large impact on the ICERs:</p> <p>6 months without data imputation:</p> <ul style="list-style-type: none"> 2 vs 1: £92,000/QALY 3 vs 1: £10,923/QALY <p>6 months with data imputation:</p> <ul style="list-style-type: none"> 2 vs 1: £18,273/QALY 3 vs 1: £23,125/QALY

(a) Other trials were included in our clinical review while this analysis was based only on one of them.

(b) This analysis was published in two papers and discrepancies were noted between papers. Inconsistent results when a 6-month time horizon was considered were not explained.

Abbreviations: Abbreviations: BiPAP = Bilevel positive airway pressure; CPAP = continuous positive airway pressure; EQ-5D = Euroqol five dimensions (scale: 0.0 [death] to 1.0 [full health]; <0.0 = worse than death); PSA = probabilistic sensitivity analysis; QALYs = quality-adjusted life years.

New cost-effectiveness analysis

New analysis was not prioritised for this area.

Economic considerations

The included evaluation (Gray et al., 2009^{93,95}, a within trial analysis using a 6-month time horizon) finds a difference in the base case results and those from a modelled lifetime extrapolation. The intervention with the highest number of QALYs at 6 months is BiPAP, while after lifetime extrapolation CPAP has the highest number of QALYs. However, over both time horizons the optimal strategy is non-invasive ventilation.

8.1.3 Evidence statements

8.1.3.1 Clinical

Non-invasive positive pressure ventilation vs. medical care

Mortality

Moderate quality evidence from 19 RCTs (N=2,253) showed that non-invasive ventilation was effective in reducing the rate of in-hospital mortality (including up to 7 day). However, two of those RCTs (N=1,129) found no clear difference for mortality measured at 30-day follow-up (low quality evidence). Non-invasive ventilation showed no advantage in reducing mortality over medical care in 3 RCTs (N=352) that administered ventilation in the pre-hospital setting (low quality evidence). A relative reduction in mortality was seen in the intensive care setting (5 RCTs, N=182). However, there was uncertainty around this effect and it is unclear whether this constituted clear appreciable benefit (low quality evidence). One study with 22 participants provided very low quality evidence for cardiac care unit mortality which was lower in those receiving non-invasive ventilation. High quality evidence showed that the rate of mortality in the emergency department was lower in those receiving non-invasive ventilation compared to medical care (9 RCTs N=1,657).

Need for intubation by setting

Moderate quality evidence from 19 RCTs (N=2,359) showed that non-invasive ventilation was effective in reducing the rate of overall rate of intubation in people with cardiogenic pulmonary oedema. Non-invasive ventilation showed a relative advantage in reducing intubation rates over medical care in 2 RCTs (N=329) that administered ventilation in the pre-hospital setting (low quality evidence). However, there was a large amount of uncertainty around this effect and it is unclear whether this constituted any benefit (low quality evidence). A clinically effective reduction in intubation rate was seen in the intensive care setting (5 RCTs, N=182) which was moderate quality evidence. One study with 22 participants provided very low quality evidence for lower cardiac care intubation rates which were lower in those receiving non-invasive ventilation with some uncertainty about the clear clinical effectiveness. High quality evidence showed that

mortality in the emergency department was reduced in those receiving non-invasive ventilation compared to medical care with uncertainty about the clinical effectiveness (10 RCTs, N=1,697).

Incidence of new myocardial infarction

- CPAP vs. medical care: Low quality evidence from 4 RCTs (N=844) showed no clear advantage associated with CPAP in prevention of new myocardial infarction.
- BiPAP vs. medical care: Moderate quality evidence from 6 RCTs (N=1,008) showed no clear advantage associated with BiPAP in prevention of new myocardial infarction.

Length of hospital stay

Overall the evidence of length of stay was very variable and inconsistent between length of stay in different hospital settings. Very low quality evidence from 1 RCT (N=100) a reduction in the average length of Cardiac Care Unit stay. However, the uncertainty around this effect indicates no clear advantage associated with non-invasive ventilation. Moderate quality evidence from three RCTs (N=127) indicate that non-invasive ventilation was effective in reducing the time spent in intensive care units. However, moderate quality evidence from one RCT (N=58) on length of stay in High Dependency Unit showed that non-invasive ventilation was associated with longer stay. When considering the total length of hospital stay a 8 RCTs (N=544) moderate quality evidence indicated a slightly shorter average length of stay in the group receiving non-invasive ventilation. However, the uncertainty around this effect does not allow clear conclusions to be drawn about any advantage of non-invasive ventilation over standard medical care on reducing total length of hospital stay.

Quality of life

Quality of life (EQ-5D) data were reported in the 3-CPO trial^{91,95} for patients after 1, 3 and 6 months follow-up for each the medical care, CPAP and BiPAP group. Only overall response rates and percentages per group were provided so it was unclear what the total number of responders was in each group. No group differences in EQ-5D were reported (MODERATE QUALITY EVIDENCE).

Early CPAP vs. late CPAP

Mortality and need for intubation

One RCT comprising 124 participants provided low quality evidence for a reduction in the rates of mortality and need for intubation associated with early administration of CPAP compared to CPAP administered after a 15-minute delay. However, there was uncertainty around this effect which made it unclear whether this constituted a clear appreciable benefit with early administration of CPAP.

BiPAP vs. high-dose intravenous isosorbide dinitrate

Mortality

Very low quality evidence from one study comprising 40 participants indicated an increased risk in mortality associated with BiPAP ventilation when compared to high dose intravenous isosorbide dinitrate. However, there was vast uncertainty around this result and it is therefore unclear whether this constitutes benefit or harm associated with either BiPAP or high-dose intravenous isosorbide dinitrate. This study was terminated early due to increased risk associated with BiPAP.

Need for intubation and new myocardial infarction

One RCT (N=40) provided very low evidence showing an increase in the rate of patients needing intubation and more patients experiencing a new myocardial infarction in the BiPAP group when

compared to patients who received high-dose intravenous isosorbide dinitrate. This study was terminated early due to increased risk associated with BiPAP.

8.1.3.2 Economic

- One cost-utility analysis showed that CPAP was cost-effective compared to standard oxygen therapy for treating patients with severe acute cardiogenic pulmonary oedema (ICER: £2,621 per QALY gained). This analysis was assessed as partially applicable with minor limitations.
- The same cost-utility analysis showed that BiPAP was NOT cost-effective compared to standard oxygen therapy for treating patients with severe acute cardiogenic pulmonary oedema (ICER: £86,375 per QALY gained).

8.1.4 Recommendations and link to evidence

Recommendations	<p>21. Do not routinely use non-invasive ventilation (continuous positive airways pressure [CPAP] or non-invasive positive pressure ventilation [NIPPV]) in people with acute heart failure and cardiogenic pulmonary oedema.</p> <p>22. If a person has cardiogenic pulmonary oedema with severe dyspnoea and acidaemia consider starting non-invasive ventilation without delay:</p> <ul style="list-style-type: none"> • at acute presentation or • as an adjunct to medical therapy if the person's condition has failed to respond.
Relative values of different outcomes	The GDG was most interested in mortality, intubation rates, incidence of major adverse cardiovascular events and quality of life. The group considered length of hospital stay to be subject to considerable variation depending on the healthcare setting and study population differences.
Trade-off between clinical benefits and harms	The meta-analysis showed non-invasive ventilation was associated with lower mortality and intubation rates, without any increase in major cardiovascular adverse events, including myocardial infarction. There was no improvement in quality of life associated with use of non-invasive ventilation, or change in length of hospital stay, but one study (Kelly 2002) did report a shorter intensive care stay and longer cardiac care unit stay associated with non-invasive ventilation.
Economic considerations	The cost-utility analysis conducted alongside the 3CPO trial suggested that non-invasive ventilation is likely to be cost-effective. However the GDG expressed caution over these results because they are driven by small differences in health-related quality of life. On balance, the GDG concluded that the cost effectiveness of non-invasive ventilation is most certain in patients with cardiogenic pulmonary oedema with severe dyspnoea and acidaemia.
Quality of evidence	<p>The evidence varied from high to very low according to GRADE criteria. The 3CPO trial provided the highest quality evidence (Gray et al. 2009), and reached different conclusions from the meta-analysis, which included small studies with a high risk of bias (including publication bias).</p> <p>The 3CPO study was considered highly applicable to the UK setting and was well conducted. The inclusion criteria of the 3CPO included, amongst others, an arterial</p>

	<p>blood pH less than 7.35 and a respiratory rate greater than 20. However the GDG noted that the most common reason for non-randomisation was being too unwell and that this may have excluded the sickest patients. The 3CPO study was qualitatively different in terms of both design and results to the other studies but this did not show up in terms of heterogeneity in the meta-analysis.</p>
Other considerations	<p>Although the results of the meta-analysis conducted showed a small mortality benefit for patients receiving non-invasive ventilation, the GDG considered that the effect of publication bias and heterogeneity of study population and healthcare setting had overshadowed the significantly larger and well conducted study in this area (3CPO). As such, the GDG chose to focus mainly on the results from this study, particularly due to its applicability to the UK healthcare setting.</p> <p>The GDG recognised that non-invasive ventilation is beneficial to some patients and is currently frequently used. The evidence suggests that early use is associated with greater benefit, but the GDG recognised that this may only be an effect on short term mortality.</p> <p>The GDG noted no difference was observed in the trials in which a direct comparison between CPAP and NIPPV was made. Thus in patients in whom NIV is considered, the evidence does not support the use of one strategy over the other.</p> <p>Some studies, but not 3CPO, found NIV was associated with reduced need for invasive ventilation and an improvement in blood oxygenation measures.</p> <p>Taking into account all patients with acute heart failure and pulmonary oedema, the GDG recommends that non-invasive ventilation should not be offered as a routine strategy. The GDG considered that in patients with severe dyspnoea, a case-by-case decision should be made, since non-invasive ventilation may reduce subsequent need for invasive ventilation and improve oxygenation, allowing time for other therapies to act. The decision to utilise NIV should be made in a timely fashion, and once the decision has been made to use NIV, the strategy should be implemented promptly. Patients with acute heart failure and pulmonary oedema who fail to respond to initial medical therapy can be considered for non-invasive ventilation given that other treatment options may be limited.</p> <p>The GDG agreed that it was not necessary to make a recommendation on the use of supplementary oxygen as an alternative method of ventilatory support as its usage is standard practice.</p> <p>The GDG noted that the majority of patients in the UK presenting with acute heart failure will have an acute decompensation of chronic heart failure. As such, the patient members emphasised that it is important, in common with all intensive and potentially harmful treatments, that NIV should only be commenced if this is in accordance with any advance treatment directive the patient may have.</p>

8.2 Invasive ventilation

Some acute heart failure patients present with respiratory distress due to pulmonary oedema. In addition to pharmacological interventions and non-invasive ventilation, some patients require invasive intervention. Invasive ventilation is not only invasive but is associated with significant morbidity and could in itself result in mortality. The dilemma that faces healthcare professionals in

these cases is whether one could predict who amongst the acute heart failure patients is either most likely to benefit from or most likely not to be harmed by invasive ventilation.

Review question: What are the predictors of outcome in invasively ventilated acute heart failure patients?

For full details see review protocol in Appendix C.

Table 61: PICO characteristics of review question

Population	Adults with acute heart failure who are invasively ventilated
Prognostic Factors	Any
Outcomes	Mortality Length of invasive ventilation Major cardiovascular events Length of hospital stay Re-admission rates Admission to critical care units Quality of life Adverse events (organ failure)
Study design	Studies using only univariate analysis will be excluded

8.2.1 Clinical evidence

This prognostic review aimed to identify independent risk factors for adverse outcomes in patients with acute heart failure who are invasively ventilated. The objective was to find out whether any particular characteristics of a person indicate who may be at risk from, or may benefit from invasive ventilation.

Three studies were included in the review^{33,70,166}. Evidence from these are summarised in the results section below. Evidence was not pooled because the multivariable analyses used different methods and different confounders to adjust results. The review is divided into the adverse outcomes for which prognostic factors were reported. These were:

- Mortality
- Prolonged weaning (> 7 days)

Prognostic factors for these outcomes are then described by study with forest plots in Appendix I.

See also the study selection flow chart in Appendix D, and exclusion list in Appendix K.

Summary of included studies

The main characteristics of the included studies are briefly described in Table 62 for full study details see study evidence tables in Appendix G.

Table 62: Summary of studies included in the review

Study	Population	Key confounders reported	Outcomes
Fedullo 1991 ⁷⁰	Patients with acute cardiogenic pulmonary oedema treated with invasive ventilation N=79 Setting: Single ICU in France	Model 1: Age; Previous history of hospitalisation for pulmonary oedema; Congestive heart failure; Diabetes; Cardiovascular disease; Taking a Ca ²⁺ channel blocker; Taking a diuretic; MI at onset; Anterior MI; Peripheral oedema;	In hospital mortality

Study	Population	Key confounders reported	Outcomes
		Respiratory rate; Systolic blood pressure; Arterial pH; Arterial pCO ₂ HCO ₃ Model 2: All above plus; Use of vasopressors; Use of lidocaine; Improvement in CXR; Awake and responsive; On ventilator; Peak CPK > 1000U/ml; High HR; Low Systolic BP; High Systolic BP; Arterial pCO ₂ ; a/A for O ₂ ; HCO ₃	
Brezins 1993 ³³	Patients with acute myocardial infarction admitted to the cardiac care unit who were treated with invasive ventilation because of pulmonary oedema that was not responding to classic treatment. N=69 Setting: Single ICU in Israel	Age (yr); Gender; Shock; Anterior MI; TranSPMral MI; Previous MI; Past angina; Past hypertension; Diabetes; Past smoking; Thrombolytic therapy; CPR before ventilation; Non-severe LV dysfunction; Severe LV dysfunction; VT/VF; Atrial fibrillation; Pacing; Severe VSD, Mitral regurgitation or tamponade; LV score; CK (LN units); BP (mm Hg); HR (beats/min)	In hospital mortality 1 year mortality
Papaioannou 2010 ^{165,166}	Primary diagnosis of severe acute respiratory failure due to acute pulmonary oedema. All patients were invasively ventilated. N=32 Setting: Single ICU in Greece	Lack of clear reporting on confounders used in regression analyses	Length of invasive ventilation Prolonged weaning (>7 days)

Results of the prognostic factors for each outcome are described by study. The following table gives an overview of risk of bias of the included studies (Table 63) The quality assessment for each prognostic factor is then assessed across studies:

Table 63: Risk of bias of studies included in the review

Study	Quality assessment – methodological flaws of studies						
	Representative population sample	Minimal attrition bias	Prognostic factor measured appropriately	Outcomes adequately measured	Important confounders accounted for	Appropriate statistical analysis	Risk of bias
Fedullo 1991 ⁷⁰	✓	✓	✓	✓	✓	✗ (a)	Serious
Brezins 1993 ³³	✗ (b)	✓	✓	✓	✓	✗ (c)	Very serious
Papaioannou 2010 ^{165,166}	✗ (d)	✓	✓	✓	✗ (e)	✗ (f)	Very serious

(a) Small sample size: n=88 patient episodes of cardiogenic pulmonary oedema and n=79 individual patients leads to < 10 deaths per variable thus too small for multivariable analysis.

(b) Included only patients with acute MI as underlying pulmonary oedema aetiology

(c) Small sample size: n= 69 patients leads to < 10 deaths per variable too small for multivariable analysis.

(d) Excluded patients with inappropriate acoustic windows, significant valvular pathologies and ventricular arrhythmia or atrial fibrillation

(e) Lack of clear reporting of confounders used in regression analyses

(f) Small sample size: n= 32 patients leads to < 10 episodes of prolonged (>7 days) weaning per variable thus too small for multivariable analysis

Prognostic factors by adverse event outcome

Two studies provided evidence for risks of mortality associated with invasive ventilation.

Mortality:

Fedullo 1991⁷⁰:

[SERIOUS LIMITATIONS]

In hospital survivors: n=56

In hospital non-survivors: n=26

Table 64: Independent risk factors for in-hospital mortality: Fedullo 1991⁷⁰

Risk factor	Coefficient ^(a)	SE of coefficient	Adjusted Odds Ratio with 95% CI
Model 1^(b)			
SBP <130 mm Hg	3.94	0.95	51.41 [8-330.9]
Previous hospitalisation with cardiogenic pulmonary oedema	-2.62	1.00	0.07 [0.01-0.51]
Use of calcium channel blockers	2.3	0.88	9.97 [1.78-55.7]
Anterior MI (diagnosed within 48 hours of intubation)	1.81	0.76	6.11 [1.39-26.84]
Age, year (10 year odds ratio)	0.124	0.048	3.45 [1.34-8.84]
Model 2^(c):			
SBP <130 mm Hg	2.82	NR ^(d)	16.78 [NR ^(d)]
Use of vasopressors at 24 hours	2.52	NR ^(d)	12.43 [NR ^(d)]

(a) Beta-coefficient from stepwise logistic regression multivariable analysis

(b) Model 1: Uses age + all variables which by the univariate analysis were statistically associated ($p < 0.05$) with mortality at the time of intubation: Age (10 year OR); Previous history of hospitalisation for pulmonary oedema; Congestive heart failure; Diabetes; Cardiovascular disease; Taking a Ca²⁺ channel blocker; Taking a diuretic; MI at onset; Anterior MI; Peripheral oedema; Respiratory rate; Systolic blood pressure; Arterial pH; Arterial pCO₂; HCO₃

(c) Model 2: Uses age + all variables which by the univariate analysis were statistically associated ($p < 0.05$) with mortality at the time of intubation, and those available at 24 hours. Those variables listed above plus: Use of vasopressors; Use of lidocaine; Improvement in CXR; Awake and responsive; On ventilator; Peak CPK > 1000U/ml; High HR (during 24-48 hours after intubation); Low Systolic BP; High Systolic BP; Arterial pCO₂; a/A for O₂; HCO₃

(d) NR - Standard error or confidence intervals were not reported

The results of model 2 were then combined as predictors for mortality to indicate the probability of death according to a combination of systolic blood pressure above or below 130 mmHg and use or absence of vasopressors at 24 hours:

Table 65: Model 2 results for the 2 prognostic factors systolic blood pressure and use of vasopressors at 24 hours (when divided up categorically)

Admission SBP (mm Hg)	Vasopressors at 24h	N ^(a)	Probability of death (%)
>130	No	40	3

Admission SBP (mm Hg)	Vasopressors at 24h	N ^(a)	Probability of death (%)
>130	Yes	10	28
<130	No	11	34
<130	Yes	21	87

(a) n = number in each group (total n is 82 rather than 88 as six patients died <24 hours after hospital admission)

Brezins 1993³³:

[VERY SERIOUS LIMITATIONS]

Overall n=69

In hospital deaths: n=46

Deaths within one year: n=12

Survivors at one year: n=11

“In univariate analysis shock was the only variable associated with in hospital mortality rates. However multivariate analysis did not show any significant relationship between these variables and in hospital survival and survival at one year.”

Confounders adjusted for: Age (yr); Gender; Shock; Anterior MI; TranSPMral MI; Previous MI; Past angina; Past hypertension; Diabetes; Past smoking; Thrombolytic therapy; CPR before ventilation; Non-severe LV dysfunction; Severe LV dysfunction; VT/VF; Atrial fibrillation; Pacing; Severe VSD, Mitral regurgitation or tamponade; LV score; CK (LN units); BP (mm Hg); HR (beats/min)

Length of invasive ventilation and prolonged weaning > 7 days:

One study provided evidence for factors that may predict prolonged weaning from invasive ventilation.

Papaïannou 2010^{165,166}:

[VERY SERIOUS LIMITATIONS]

Duration of weaning < 7 days: n=20

Duration of weaning >7 days: n=12

All data reported verbatim from Papaïannou 2010^{165,166}. A major study limitation is the lack of clear reporting on confounders used in the regression analyses. This study focused on echocardiographic prognostic factors that may predict prolonged weaning. Particular parameters found to be prognostic were tricuspid annular plane systolic excursion (TAPSE), systolic right ventricular tissue Doppler imaging velocity (Sm), the ratio of early vs. late diastolic right ventricular tissue Doppler imaging velocity (Em/Am), left ventricular ejection fraction (LVEF), and right ventricular fractional area change (RVFAC) – see below.

Linear univariate regression analysis revealed significant associations between duration of ventilation and:

TAPSE: Beta slope = -0.89, SE = 0.14, p<0.001

Sm: Beta slope = -0.57. SE = 0.09, p<0.001

Em/Am: Beta slope = -0.27. SE = 0.05, p<0.001

Logistic univariate regression analysis revealed that the following can predict length of weaning > 7 days:

TAPSE: beta = 0.76, SE = 0.043, $p < 0.001$

LVEF: beta = 0.87, SE = 0.03, $p < 0.001$

Sm: beta = 0.75, SE = 0.03, $p < 0.001$

Em/Am: beta = 0.32, SE = 0.05, $p < 0.001$

RVFAC: beta = 0.74, SE = 0.03, $p < 0.001$

The authors state that in multivariate analysis after adjustment of predictors found in univariate models for age, systolic blood pressure, heart rate, body surface area and duration of intravenous therapy the prognostic factors described above were independently associated with the outcome of interest ($p < 0.05$).

8.2.2 Economic evidence

Published literature

No relevant economic evaluations were identified.

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

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs were provided to aid consideration of cost-effectiveness. These are given in Appendix N.

8.2.3 Evidence statements

8.2.3.1 Clinical

Mortality

- One study with serious limitations including 79 patients suggested age, an SBP <130 mm Hg, use of calcium channel blockers and anterior myocardial infarction (diagnosed within 48 hours of intubation) were independent risk factors, and previous hospitalisation with cardiogenic pulmonary oedema was an independent protective factor of in hospital mortality in patients with acute cardiogenic pulmonary oedema treated with invasive ventilation. This was the case when considering only variables available at baseline.
- One study with serious limitations including 79 patients suggested an SBP <130 mm Hg, and use of vasopressor medication were independent risk factors of in hospital mortality in patients with acute cardiogenic pulmonary oedema treated with invasive ventilation. This was the case when considering variables available at baseline, and at 24 hours of invasive ventilation.
- One study with very serious limitations including 69 patients did not find any independent risk factors of in hospital or one year mortality in patients with acute myocardial infarction admitted to the cardiac care unit who were treated with invasive ventilation because of pulmonary oedema that was not responding to classic treatment.

Length of invasive ventilation and prolonged weaning > 7 days

- One study with very serious limitations including 32 patients found significant associations between length of invasive ventilation and the echocardiographic parameters TAPSE, Sm and Em/Am in patients with a primary diagnosis of severe acute respiratory failure due to acute pulmonary oedema who were invasively ventilated.
- One study with very serious limitations including 32 patients found that the echocardiographic parameters TAPSE, LVEF, Sm, Em/Am and RVFAC are independent risk factors of prolonged weaning > 7 days in patients with a primary diagnosis of severe acute respiratory failure due to acute pulmonary oedema who were invasively ventilated.

8.2.3.2 Economic

- No relevant economic evaluations were identified

8.2.4 Recommendations and link to evidence

Recommendations	23. Consider invasive ventilation in people with acute heart failure that, despite treatment, is leading to or is complicated by: <ul style="list-style-type: none"> • respiratory failure or • reduced consciousness or physical exhaustion.
Relative values of different outcomes	In-hospital mortality, one-year mortality and length of invasive ventilation were the only outcomes that were available of those prioritised by the GDG.
Trade-off between clinical benefits and harms	Decisions on use of invasive ventilation are made on an individual patient basis, including establishing whether a patient has made a prior decision regarding ventilation. The GDG could not identify any clinical variables from the available evidence that was predictive of outcomes for patients being considered for ventilation in acute heart failure.
Economic considerations	No relevant economic evaluations were identified. The clinical review considered the importance of prognostic risk factors on outcomes when using invasive ventilation; the clinical effectiveness of invasive ventilation versus another intervention was not directly considered. The GDG made a qualitative decision that invasive ventilation for patients with respiratory failure and reduced consciousness/physical exhaustion would be cost-effective versus non-invasive ventilation. The GDG considered the clinical outcomes for these patient types to be particularly poor without invasive ventilation.
Quality of evidence	The GDG noted that the included studies had serious or very serious limitations relating to the fact that they were small studies (n < 100), with no robust multivariable models, i.e. some studies included too many variables in analysis for the number of patients included (i.e. less than 10 patients per variable). Significant changes in practice have occurred since the publication of the earlier studies (Brezins 1993 and Fedullo 1991).
Other considerations	<p>The evidence presented was not sufficient to make recommendations on prognostic factors for mortality or for the length of invasive ventilation in patients with acute heart failure. Due to the paucity and very serious limitations of the included data the GDG made a consensus recommendation, and considered the ESC heart failure guideline 2012.</p> <p>The GDG did not want to specify which treatment should be given before ventilation as individual clinical circumstances will vary.</p> <p>The GDG noted that the majority of patients in the UK presenting with acute heart failure will have an acute decompensation of chronic heart failure. As such, the patient members emphasised that it is important, in common with all intensive and potentially harmful treatments, that invasive ventilation should only be commenced</p>

if this is in accordance with any advance treatment directive the patient may have.

8.3 Ultrafiltration

Patients being admitted to hospital with acute heart failure will commonly have significant excess fluid which may have accumulated throughout their peripheries and abdomen. Whilst diuretic medications are the main treatment used to remove this fluid, the medications can cause side effects, may take many days to be effective, or some patients may be resistant to their actions. Ultrafiltration is a technique where excess water and salt can be filtered from the blood rapidly, and with advancing technology it has been suggested that it should play more of a role in the management of patients with significant oedema. However, it is an invasive therapy, is not free of side effects and has cost implications. This review question investigates the use of ultrafiltration in comparison to diuretic treatment.

Review question: In patients with acute heart failure is ultrafiltration more clinical / cost effective than diuretic therapy alone or in addition to diuretic therapy to improve outcome?

For full details see review protocol in Appendix C.

Table 66: PICO characteristics of review question

Population	Adults (18 years and above) with AHF
Intervention/s	- Ultrafiltration alone - Ultrafiltration in addition to diuretic therapy
Comparison/s	- Standard medical care with diuretics
Outcomes	- Mortality - Major cardiovascular events - Length of hospital stay - Re-admission rates - Dyspnoea - Clinical status - Weight Loss - Renal function - Quality of life - Adverse events (particularly renal and cardiovascular events)
Study design	Systematic reviews and meta-analyses Randomised control trials

8.3.1 Clinical evidence

We looked for systematic reviews and randomised controlled trials comparing the effectiveness of ultrafiltration versus standard medical care with diuretics.

Seven studies were included in the review^{20,22,23,45,90,99,133}. A variety of ultrafiltration systems, fluid removal rates and durations was used. The main characteristics of these studies are summarised in Table 67. The aim of all these trials was to assess whether ultrafiltration was effective in preventing

death, major cardiovascular events and worsening renal function amongst other outcomes as listed in Table 66. Evidence from these is summarised in the clinical GRADE evidence profiles below. The analysis was divided into two comparisons:

- Five studies^{20,23,45,90,99} that compared ultrafiltration without additional diuretics (i.e. where the pharmaceutical diuretic was discontinued) to diuretics (as summarised in the GRADE profile in Table 68).
- Two studies^{22,133} that compared ultrafiltration with additional diuretics to medical care (including diuretics). This is summarised in the GRADE profile in Table 69.

Four systematic reviews^{115,132,227,233} were identified when searches were rerun. These were cross checked for additional studies to the review. Apart from one reference²⁰, all studies fitting the protocol criteria were already included in the original meta-analysis. Our meta-analysis is not replaced by these reviews, due to the more comprehensive list of outcomes analysed. See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and lists of excluded studies in Appendix K.

Table 67: Summary of studies included in the review

Study	Population	Ultrafiltration	Comparison
Badawy et al, 2012 ²⁰	Hospitalised patients with ADHF (n=40)	Continuous veno-venous hemodiafiltration with either a Multifiltrate or a Prismaflex machine. <u>Fluid removal</u> : The rate of ultrafiltration defined as the net fluid lost by the machine per hour was determined by the attending physician but never exceeded 200 mL/h <u>Duration</u> : 72 hours <u>Concurrent medication / care</u> : Heparin. All baseline cardiac medications according to the ICU protocol were continued except diuretics.	<u>Furosemide</u> : A loading bolus of 1 mg/kg and then a continuous furosemide infusion starting with 20 mg/h. The rate of continuous infusion could be increased to maintain the urine output > 1 mL/kg per hour. <u>Concurrent medication</u> : All baseline cardiac medications according to the ICU protocol.
Bart et al, 2005 (RAPID-CHF) ²²	Hospitalised patients with AHF (n=40)	System 100, CHF Solutions Inc., Brooklyn Park, Minnesota <u>Fluid removal</u> : Determined by the attending physician (to a maximum of 500cc/h) <u>Duration</u> : Single 8 hour session. Median time from consent to UF initiation was 3.69 hours. <u>Concurrent medication/care</u> : Diuretics were held during the 8hr of UF thereafter they were administered at the discretion of the treating physician. No further courses of UF were permitted until after 24 hours. Percentage of patients receiving following	Usual care <u>Duration</u> : 24 hours <u>Concurrent medication/care</u> : Percentage of patients receiving following medications: IV Diuretics: 95%; Nesiritide: 50%; IV inotropes 10%

		medications: IV Diuretics: 65%; Nesiritide: 20%; IV inotropes 0	
Bart et al, 2012 (CARRESS-HF) ^{22,23}	ADHF with worsened renal function (an increase of serum creatinine of at least 26.5µmol/l within 12 weeks before or 10 days after index admission) (n=188)	Aqualex system 100 (CHF solutions) <u>Fluid removal</u> : 200ml/hour started a median of 8 hours post randomisation <u>Duration</u> : Median 40 hours (IQR: 28 to 67) <u>Concurrent medication/care</u> : 30% received loop IV diuretics before 96 hours; 3% received IV vasodilators; 3% received inotropes	Diuretic-based stepped pharmacological therapy aimed at maintaining urine output at 3-5litres/day <u>Duration</u> : Median 92 hours (interquartile range (IQR): 56 to 138) <u>Concurrent medication/care</u> : In addition to loop diuretics 46% received metolazone; 5% received IV vasodilators; 12% received inotropes
Costanzo et al, 2007 (UNLOAD) ⁴⁵	Patients hospitalised with AHF (n=200)	Aqualex System 100 (CHF solutions) <u>Fluid removal</u> : Average rate of 241ml/h <u>Duration</u> : 12.3 +/- 12 hours. <u>Concurrent medication/care</u> : Heparin. Treatment with intravenous diuretics prohibited.	Average IV loop diuretic dose 181 +/- 121mg. 68 patients received diuretics as bolus, 32 as a continuous infusion. <u>Duration</u> : 48 hours <u>Concurrent medication/care</u> : Standard care
Giglioli et al, 2011 (ULTRADISCO) ⁹⁰	Patients hospitalised with AHF (n=30)	PRISMA System (HOSPAL-GAMBRO DASCO, Medolla, Italy) <u>Fluid removal</u> : from 100 to 300mL/h adjusted for SBP (mmHg): <100: 100mL/h; >100 ≤110: 200mL/h; >110: 300mL/h <u>Duration</u> : Median length of treatment 46 hours (IQR 39-71) <u>Concurrent medication/care</u> : Unfractionated heparin. No concurrent IV diuretic administration.	Continuous infusion of furosemide at initial dose of 250mg/24 hours. Dose lowered or if plasma creatinine > 44µmol/l when clinical score had decreased by a third or when a reduction in SBP or HR by 15% was observed. Initial dose increased to 500mg/25 hours if achievement of negative fluid balance was not sufficient to reach >2000mL/day. <u>Duration</u> : Median length of treatment was 57 hours (IQR48-85) <u>Concurrent medication/care</u> : Nil progressed to inotrope therapy
Hanna et al, 2012 ⁹⁹	ADHF (n=36)	(NXstage System One; NXStage System inc. Lawrence MA) <u>Fluid removal</u> : 400mL/h for 6 hours and then decreased to 200ml/hour (changes permitted as clinically directed) . <u>Duration</u> : Mean time to achieve primary end point	IV diuretics at doses and frequencies designated by the treating physician. <u>Duration</u> : Mean time to achieve primary end point (PCWP <18mmHg for at least 4 hours) 34.8 hours (6.7) <u>Concurrent medication/care</u> : Received IV vasoactive medication at clinician discretion

		(PCWP <18mmHg for at least 4 hours) 22 hours (4.2). <u>Concurrent medication/care:</u> Heparin.; Diuretics stopped except for spironolactone (<=25mg/day) received IV vasoactive medication at clinician discretion	
Marenzi et al, 2014 (CUORE) ¹³³	Hospitalised patients with congestive HF, NYHA class III or IV, LVEF ≤40% and estimated weight gain due to peripheral fluid overload ≥ 4 kg in the preceding 2 months. (n=56)	A simplified device (Dedyc, Bellco, Mirandola, Italy) <u>Fluid removal:</u> 100-500 mL/h at discretion of the treating physician <u>Duration:</u> Single or double session of ultrafiltration. Mean time was 19 (±9) hours. The session duration was left to the discretion of the treating physician. <u>Concurrent medication/care:</u> Heparin; additional medical therapy was left to the discretion of the cardiologist responsible for the patient. The intravenous dosage of diuretics started before randomisation was left unchanged. Pharmacological withdrawal, including diuretics was not advised during ultrafiltration sessions.	IV diuretics according to guideline recommendations at discretion of the treating physician. <u>Concurrent medication/care:</u> Additional medical therapy was left to the discretion of the cardiologist responsible for the patient.

Table 68: GRADE profile for the comparison of ultrafiltration versus diuretic therapy

Quality assessment							Summary of findings				Quality	Importance	
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ultrafiltration (n) Mean (SD) or Event rate	Pharmacological Therapy (n) Mean (SD) or Event rate*	Effect Size				
									Relative Risk (95% CI)	Absolute effect Mean Difference (MD) (95% CI)			
All-cause mortality at 60 days ^{22,24}													
1	randomised trials	serious	no serious inconsistency	no serious indirectness	serious ^(c)	none	n=94	n=94	Hazard Ratio 1.32 (0.46 to 3.81)		LOW	CRITICAL	
All-cause mortality total ^{20,46,99}													
3	Randomised trials	serious	no serious inconsistency	no serious indirectness	very serious ^(c)	none	16/133 (12%)	20/132 (15.2%)	Risk Ratio 0.78 (0.43 to 1.44)	44 fewer per 1000 (from 114 fewer to 88 more)	VERY LOW	CRITICAL	
								20%					
All-cause mortality at 30 Days ²⁰													
1	randomised trials	very serious	no serious inconsistency	no serious indirectness	very serious ^(c)	none	3/20 (15%)	5/20 (25%)	Risk Ratio 0.60 (0.17 to 2.18)	100 fewer per 1000 (from 207 fewer to 295 more)	VERY LOW	CRITICAL	
All-cause mortality at 90 Days ^{46,99}													
2	randomised trials	serious	no serious inconsistency	no serious indirectness	very serious ^(c)	none	13/113 (11.5%)	15/112 (13.4%)	RR 0.85 (0.42 to 1.69)	26 fewer per 1000 (from 102 fewer to 121 more)	VERY LOW	CRITICAL	
								17.6%					

Quality assessment							Summary of findings				Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ultrafiltration (n) Mean (SD) or Event rate	Pharmacological Therapy (n) Mean (SD) or Event rate*	Effect Size			
									Relative Risk (95% CI)	Absolute effect Mean Difference (MD) (95% CI)		
Length of hospital stay ^{20,24,46,99}												
2	randomised trials	serious	no serious inconsistency	no serious indirectness	no serious imprecision	none	n=120 12 (6) 6.3 (4.9)	n=120 19 (7) 5.8 (3.8)	-	MD 0.12 lower (1.29 lower to 1.04 higher)	MODERATE	IMPORTANT
1	randomised trials	serious	no serious inconsistency	no serious indirectness	N/A ^(d)	none	Median: 4.53 days	Median: 9.61days	p<0.019		MODERATE	IMPORTANT
1	randomised trials	serious	no serious inconsistency	no serious indirectness	N/A ^(d)	none	Median: 7 days	Median: 6 days	p=NS		MODERATE	IMPORTANT
Number of patients readmitted for any cause (follow-up 60-90 days) ^{24,99}												
2	randomised trials	serious	no serious inconsistency	no serious indirectness	serious ^(c)	none	54/109 (49.5%)	43/110 (39.1%) 37.5%	RR 1.27 (0.94 to 1.72)	101 more per 1000 (from 23 fewer to 270 more)	LOW	IMPORTANT
Number of patients readmitted for any cause (follow-up 60 days) ^{22,24}												
1	randomised trials	serious	no serious inconsistency	no serious indirectness	serious ^(c)	none	46/90 (51.1%)	37/93 (39.8%)	RR 1.28 (0.93 to 1.77)	111 more per 1000 (from 28 fewer to 306 more)	LOW	IMPORTANT
Number of patients readmitted for any cause (follow-up 90 days) ⁹⁹												
1	randomised trials	serious	no serious inconsistency	no	very	none	8/19	6/17	RR	67 more per 1000	VERY	IMPORTANT

Quality assessment							Summary of findings				Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ultrafiltration (n) Mean (SD) or Event rate	Pharmacological Therapy (n) Mean (SD) or Event rate*	Effect Size			
									Relative Risk (95% CI)	Absolute effect Mean Difference (MD) (95% CI)		
	sed trials	usa	inconsistency	serious indirectness	serious ^(c)		(42.1%)	(35.3%)	1.19 (0.52 to 2.74)	(from 169 fewer to 614 more)	LOW	NT
Number of patients readmitted due to HF (follow-up 60-90 days) ^{24,46}												
2	randomised trials	seriousa	seriousb	no serious indirectness	serious ^(c)	none	39/179 (21.8%)	52/180 (28.9%) 29%	RR 0.75 (0.53 to 1.08)	72 fewer per 1000 (from 136 fewer to 23 more)	VERY LOW	IMPORTANT
Number of patients readmitted due to HF (follow-up 60 days) ^{22,24}												
1	randomised trials	seriousa	no serious inconsistency	no serious indirectness	very serious ^(c)	none	23/90 (25.6%)	24/93 (25.8%)	RR 0.99 (0.6 to 1.62)	3 fewer per 1000 (from 103 fewer to 160 more)	VERY LOW	IMPORTANT
Number of patients readmitted due to HF (follow-up 90 days) ^{45,46}												
1	randomised trials	seriousa	no serious inconsistency	no serious indirectness	serious ^(c)	none	16/89 (18%)	28/87 (32.2%)	RR 0.56 (0.33 to 0.96)	142 fewer per 1000 (from 13 fewer to 216 fewer)	LOW	IMPORTANT
Change in score on dyspnoea 100mm VAS from baseline (follow-up 96 hours; Better indicated by higher values) ^{22,24}												
1	randomised trials	seriousa	no serious inconsistency	no serious indirectness	no serious imprecision	none	n=94 16.5 (29.2)	n=94 20.5 (27.8)	-	MD 4 lower (12.15 lower to 4.15 higher)	MODERATE	IMPORTANT
Mean dyspnoea score (follow-up 48 hours; Better indicated by higher values) ^{45,46}												

Quality assessment							Summary of findings				Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ultrafiltration (n) Mean (SD) or Event rate	Pharmacological Therapy (n) Mean (SD) or Event rate*	Effect Size			
									Relative Risk (95% CI)	Absolute effect Mean Difference (MD) (95% CI)		
1	randomised trials	very serious	no serious inconsistency	no serious indirectness	serious ^(c)	none	n=80 6.4 (0.502)	n=83 6.1 (0.697)	-	MD 0.3 higher (0.11 to 0.49 higher)	VERY LOW	IMPORTANT
End NYHA class at 36 hours (follow-up 36 hours; Better indicated by lower values) ⁹⁰												
1	randomised trials	serious	no serious inconsistency	no serious indirectness	serious ^(c)	none	n=15 2 (0.5)	n=15 2.4 (0.52)	-	MD 0.4 lower (0.77 to 0.03 lower)	LOW	IMPORTANT
Change in score on global well-being scale from baseline (follow-up 96 hours; Better indicated by higher values) ^{22,24}												
1	randomised trials	very serious3	no serious inconsistency	no serious indirectness	serious ^(c)	none	n=94 13.7 (27.9)	n=94 22.8 (25.8)	-	MD 9.1 lower (16.78 to 1.42 lower)	VERY LOW	IMPORTANT
Number of patients achieving clinical decongestion (follow-up 96 hours) ^{22,24}												
1	randomised trials	very serious	no serious inconsistency	no serious indirectness	very serious ^(c)	none	8/82 (9.8%)	7/80 (8.8%)	RR 1.11 (0.42 to 2.93)	10 more per 1000 (from 51 fewer to 169 more)	VERY LOW	IMPORTANT
Mean change from baseline in body weight (kg) (follow-up up to 48 hours; Better indicated by lower values) ^{46,99}												
2	randomised trials	serious	seriousb	no serious indirectness	serious ^(c)	none	n=100 -5 (3.1) -4.7 (3.5)	n=101 -3.1 (3.5) -1 (2.5)	-	MD 2.25 lower (3.15 to 1.35 lower)	VERY LOW	IMPORTANT
Mean change from baseline in body weight (kg) (follow-up up to 48 hours; Better indicated by lower values) ²⁰												

Quality assessment							Summary of findings				Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ultrafiltration (n) Mean (SD) or Event rate	Pharmacological Therapy (n) Mean (SD) or Event rate*	Effect Size			
									Relative Risk (95% CI)	Absolute effect Mean Difference (MD) (95% CI)		
1	randomised trial	very serious	no serious inconsistency	no serious indirectness	serious ^(c)	none	n=20 -6.3 (3.5)	n=20 -3.7 (3.2)	-	MD 2.6 lower (4.68 lower to 0.52 lower)	VERY LOW	IMPORTANT
Mean change from baseline in body weight (kg) (follow-up 96 hours; Better indicated by lower values) ^{22,24}												
1	randomised trials	serious	no serious inconsistency	no serious indirectness	no serious imprecision	none	n=94 -5.7 (3.9)	n=94 -5.5 (5.1)	-	MD 0.2 lower (1.5 lower to 1.1 higher)	MODERATE	IMPORTANT
Mean change from baseline in body weight (kg) (follow-up 60 days; Better indicated by lower values) ^{22,24}												
1	randomised trials	serious	no serious inconsistency	no serious indirectness	no serious imprecision	none	n=94 -16 (7.96)	n=94 -17 (7.96)	-	MD 1 higher (1.27 lower to 3.27 higher)	MODERATE	IMPORTANT
Weight (kg): % of baseline (follow-up 36 hours; Better indicated by lower values) ⁹⁰												
1	randomised trials	serious	no serious inconsistency	no serious indirectness	serious ^(c)	none	n=15 90.9 (1.7)	n=15 93.1 (1.8)	-	MD 2.2 lower (3.45 to 0.95 lower)	LOW	IMPORTANT
End creatinine score (μmol/l) (follow-up 36 hours; Better indicated by lower values) ⁹⁰												
1	randomised trials	very serious	no serious inconsistency	no serious indirectness	serious ^(c)	none	n=15 147.63 (65.42)	n=15 170.61 (54.81)	-	MD 22.98 lower (66.17 lower to 20.2 higher)	VERY LOW	IMPORTANT
End creatinine score (μmol/l) (follow-up 72 hours; Better indicated by lower values) ²⁰												
1	randomised trial	very serious	no serious inconsistency	no serious indirectness	serious ^(c)	none	n=20 97.24 (53.04)	n=20 141.44 (88.4)	-	MD 44.2 lower (89.38 lower to 0.98 higher)	VERY LOW	IMPORTANT

Quality assessment							Summary of findings				Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ultrafiltration (n) Mean (SD) or Event rate	Pharmacological Therapy (n) Mean (SD) or Event rate*	Effect Size			
									Relative Risk (95% CI)	Absolute effect Mean Difference (MD) (95% CI)		
				ess								
Change in serum creatinine (μmol/l) (follow-up 96 hours; Better indicated by lower values) ^{24,99}												
2	randomised trials	serious	no serious inconsistency	no serious indirectness	serious ^(c)	none	n=113 20.9 (61.9) 194.48 (10.6.08)	n=111 -3.5 (46.9) 167.96 (79.56)e	-	MD 23.97 higher (8.77 to 39.17 higher)	LOW	IMPORTANT
Change in serum creatinine (μmol/l) (follow-up 60 days; Better indicated by lower values) ^{22,24}												
1	randomised trials	serious	no serious inconsistency	no serious indirectness	serious ^(c)	none	n=94 -10.608 (77.37)	n=94 -35.36 (77.37)	-	MD 24.75 higher (2.63 to 46.87 higher)	LOW	IMPORTANT
Number of patients with rise in serum creatinine 26.5 μmol/litre (follow-up 24 hours) ^{45,46}												
1	randomised trials	serious	no serious inconsistency	no serious indirectness	serious ^(c)	none	13/90 (14.4%)	7/91 (7.7%)	RR 1.88 (0.79 to 4.49)	68 more per 1000 (from 16 fewer to 268 more)	LOW	IMPORTANT
Number of patients with rise in serum creatinine 26.5 μmol/litre (follow-up 48 hours) ^{46,99}												
2	randomised trials	serious	no serious inconsistency	no serious indirectness	Serious ^(c)	none	24/87 (27.6%)	19/91 (20.9%) 21.9%	RR 1.31 (0.78 to 2.22)	68 more per 1000 (from 48 fewer to 267 more)	LOW	IMPORTANT
Total number of patients with any SAE (follow-up 60 days) ^{22,24}												
1	randomised	serious	no serious inconsistency	no serious indirectness	serious ^(c)	none	68/94 (72.3%)	54/94 (57.4%)	RR 1.26 (1.02	149 more per 1000 (from 11 more to	LOW	IMPORTANT

Quality assessment							Summary of findings				Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ultrafiltration (n) Mean (SD) or Event rate	Pharmacological Therapy (n) Mean (SD) or Event rate*	Effect Size			
									Relative Risk (95% CI)	Absolute effect Mean Difference (MD) (95% CI)		
	trials		ncy	ess					to 1.56)	322 more)		
Total number of patients with heart failure SAE (follow-up 60 days) ^{22,24}												
1	randomised trials	serious	no serious inconsistency	no serious indirectness	very serious ^(c)	none	31/94 (33%)	28/94 (29.8%)	RR 1.11 (0.72 to 1.69)	33 more per 1000 (from 83 fewer to 206 more)	VERY LOW	IMPORTANT
Total number of patients with other cardiovascular SAEs (follow-up 60 days) ^{22,24}												
1	randomised trials	serious	no serious inconsistency	no serious indirectness	very serious ^(c)	none	6/94 (6.4%)	5/94 (5.3%)	RR 1.2 (0.38 to 3.8)	11 more per 1000 (from 33 fewer to 149 more)	VERY LOW	IMPORTANT
Total number of patients with new dialysis dependence (at hospital discharge) ²⁰												
1	randomised trials	very serious	no serious inconsistency	no serious indirectness	very serious ^(c)	none	1/17 (5.7%)	1/15 (6.7%)	RR 0.88 (0.06 to 12.91)	8 fewer per 1000 (from 63 fewer to 794 more)	VERY LOW	IMPORTANT
Total number of patients with renal failure SAE (follow-up 60 days; Better indicated by lower values) ^{22,24}												
1	randomised trials	serious	no serious inconsistency	no serious indirectness	very serious ^(c)	none	17/94 (18.1%)	14/94 (14.9%)	RR 1.21 (0.64 to 2.32)	31 more per 1000 (from 54 fewer to 197 more)	VERY LOW	IMPORTANT

* In case of divided cells within a row the top cell refers to the overall control event rate whereas the lower cell presents the median control even rate on which the absolute effect is based.

- (a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitations across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitations across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.*
- (b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I^2 50 - 74%, or chi square p value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I^2 75% or more). Number of patients readmitted due to HF, and mean change from baseline in body weight (kg), were sub grouped by time at measurement. This sub-grouping strategy removed the heterogeneity for all outcomes*
- (c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID. Default MIDs were set at RRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the control group median standard deviation either side of the null line for continuous variables*
- (d) Median values are reported for the length of stay outcome. These are reported in the GRADE table and unable to be assessed for imprecision.*
- (e) End scores and change from baseline scores are reported for the change in serum creatinine at 96 hours outcome and are analysed together using mean difference.*

Narratively reported data (not pooled because outcomes were not reported in sufficient detail):

Costanzo 2007:

Renal Function: Changes in serum creatinine were similar in the group receiving ultrafiltration and the control group throughout the study with $p > 0.05$ at all time points measured (48 hours-90 days). Figures too small to extract data.

Quality of Life: At each assessment "Minnesota Living with Heart Failure" scores were similarly improved in the two groups. No effect measures or statistics provided.

Hanna 2012:

Quality of Life: Ninety day follow up for quality of life was not statistically different. No effect measures or statistics provided.

Adverse events: There were no significant differences in the adverse events between both groups. Adverse events not described in detail.

Table 69: GRADE profile for ultrafiltration + diuretic therapy versus usual care

Quality assessment							Summary of findings				Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Ultrafiltration +/- IV Diuretics (n) Mean (SD) or Event rate	Usual Care (n) Mean (SD) or Event rate	Effect size			
									Relative (95% CI)	Absolute effect Mean Difference (MD) (95% CI)		
All-cause mortality total ^{22,133}												
2	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	very serious ^b	none	12/49 (24.5%)	7/47 (14.9%)	RR 1.56 (0.72 to 3.37)	83 more per 1000 (from 42 fewer to 353 more)	VERY LOW	CRITICAL
All-cause mortality (follow-up 30 days) ²²												
1	randomised trials	very serious ^a	no serious inconsistency	no serious indirectness	very serious ^b	none	1/20 (5%)	0/20 (0%)	Peto OR 7.39 (0.15 to 372.38) ^d	50 more per 1000 (from 80 fewer to 180 more)	VERY LOW	CRITICAL
All-cause mortality (follow-up 1 year) ¹³³												
1	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	very serious ^b	none	11/29 (37.9%)	7/27 (25.9%)	RR 1.46 (0.66 to 3.22)	119 more per 1000 (from 88 fewer to 576 more)	VERY LOW	CRITICAL
Rehospitalisations due to congestive heart failure (follow-up 1 year) ¹³³												
1	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	no serious imprecision	none	3/27 (11%)	14/29 (48%)	HR 0.14 (0.04 to 0.49)		MODE RATE	CRITICAL
Change in body weight from baseline at hospital discharge - kg ¹³³												

1	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	no serious imprecision	none	n=19 -7.5 (5.6)	n=18 -7.9 (9)	-	MD 0.4 higher (3.5 lower to 4.3 higher)	MODE RATE	CRITICAL
Renal function: serum creatinine (μmol/l) at hospital discharge¹³³												
1	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	serious ^b	none	n=27 159.12 (61.88)	n=29 167.96 (61.88)	-	MD 8.84 lower (41.27 lower to 23.59 higher)	LOW	IMPORTANT
Renal function: serum creatinine (μmol/l) at 6 months¹³³												
1	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	serious ^b	none	n=24 159.12 (53.04)	n=27 203.32 (97.24)	-	MD 44.20 lower (86.57 lower to 1.83 lower)	LOW	IMPORTANT
Renal function: serum creatinine (μmol/l) at 1 year¹³³												
1	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	very serious ^b	none	n=19 159.12 (53.04)	n=17 159.12 (44.2)	-	MD 0 lower (31.78 lower to 31.78 higher)	VERY LOW	IMPORTANT
Length of hospital stay (days)^{22,133}												
1	randomised trials	serious ^a	no serious inconsistency	no serious indirectness	serious ^b	none	n=19 7.4 (4.6)	n=18 9.1 (1.9)	-	MD 1.7 lower (3.95 lower to 0.55 higher)	LOW	IMPORTANT
1	randomised trials	very serious ^a	no serious inconsistency	no serious indirectness	N/A ^c	none	n=20 Median: 6 days	n=20 Median: 5 days	NR		LOW	IMPORTANT
Number of patients with marked, moderate or mild improvement in dyspnoea (follow-up 24 hours)²²												
1	randomised trials	very serious ^a	no serious inconsistency	no serious indirectness	serious ^b	none	16/19 (84.2%)	15/19 (78.9%)	RR 1.07 (0.79 to 1.44)	55 more per 1000 (from 166 fewer to 347 more)	VERY LOW	IMPORTANT
Number of patients with marked, moderate or mild improvement in global symptoms (follow-up 24 hours)²²												
1	randomised trials	very serious ^a	no serious inconsistency	no serious indirectness	serious ^b	none	17/19 (89.5%)	14/19 (73.7%)	RR 1.21 (0.89 to 1.66)	155 more per 1000 (from 81 fewer to 486 more)	VERY LOW	IMPORTANT

- (a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitations across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitations across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality.*
- (b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed the lower MID or the upper or lower 95% CI crossed the upper MID. Outcomes were downgraded by two increments if the upper CI simultaneously crossed the upper MID and the lower CI crossed the lower MID. Default MIDs were set at RRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the control group median standard deviation either side of the null line for continuous variables*
- (c) Median values are reported for the length of stay outcome. These are reported in the GRADE table and unable to be assessed for imprecision.*
- (e) In cases of zero event rate in one arm Peto OR rather than RR is used*

Narratively reported data (not pooled because outcomes were not reported in sufficient detail):

Bart 2005:

Renal Function: Difference in change in creatinine from baseline ($\mu\text{mol/l}$) at 24 hours was +0.1 between groups. No significant difference.

Weight loss: Weight loss was greater in the UF group but failed to reach statistical significance ($p=0.24$).

8.3.2 Economic evidence

Published literature

One study was included with the relevant comparison.¹⁸² This is summarised in the economic evidence profile below (Table 70) and the economic evidence tables in Appendix H.

One study³¹ that met the inclusion criteria was selectively excluded due to the availability of more applicable evidence. This is summarised in Appendix L, with reasons for exclusion given.

See also the study selection flow chart in Appendix E.

Table 70: Economic evidence profile: ultrafiltration versus intravenous diuretic

Study	Applicability	Limitations	Other comments	Incremental cost per patient (£)	Incremental effects (incremental risk of death at 90 days)	Cost effectiveness (£ per death averted (all-cause))	Uncertainty
DOH 2007 ¹⁸² UK	Partially applicable ^(a)	Potentially serious limitations ^(b)	UK costs based on resource use in UNLOAD trial. Population: Patients hospitalized for acute decompensated heart failure.	Ultrafiltration = £608 increase ^(c)	Ultrafiltration = 0.02 decrease ^(d)	£30,400 ^(d)	Alternative care settings (ICU, day case) were addressed in a sensitivity analysis. A PSA was not conducted.

(a) Costs from the UK setting, however resource utilisation was measured in the USA setting.

(b) Not a cost-effectiveness analysis. Health outcomes based on 90-day follow up results from UNLOAD trial. The current data on ultrafiltration for heart failure are too limited in length of follow-up to extrapolate cost and survival estimates for a reasonable comparison of cost per life-year or cost per quality-adjusted life-year gained.

(c) 2005/2006 UK pounds. Cost components were: preparation, heparin, furosemide, ultrafiltration equipment and consumables, haematocrit testing, hospital care, readmission and emergency care.

(d) Calculated by the NCGC based on the UNLOAD trial data.⁴⁵

(e) Abbreviations: PSA = probabilistic sensitivity analysis

DOH2007 assessed the cost of care only, by attaching local unit costs to resource use data reported in the USA setting from the UNLOAD trial.⁴⁵ No cost effectiveness analysis was performed due to no statistically significant differences in primary outcomes between the two study groups in the UNLOAD trial. There were a small number of deaths in each treatment group (IV diuretics n=11/100, and ultrafiltration n=9/100), and a cost effectiveness analysis was performed here using these mortality data along with costs reported in each study. However we recognise the limitations of this analysis due to the few events. Also of note in this study; the accumulation of resource utilisation was measured in the USA setting which diminishes the study's applicability to the UK NHS setting. DOH2007 performed sensitivity analyses of both strategies in alternative care settings (ICU and day care), but did not assess the impact of a higher level of staff qualification, which may be a requirement specific to ultrafiltration.

New cost-effectiveness analysis

New analysis was not prioritised for this area.

8.3.3 Evidence statements

Clinical

Ultrafiltration versus diuretic therapy

All-cause mortality:

- Low quality evidence from 1 study with 188 participants showed similar rates of mortality in people receiving ultrafiltration and those who had diuretic therapy at 60 days follow up.
- Very low quality evidence from 3 studies with 265 participants showed similar total rates of mortality in people receiving ultrafiltration and those who had diuretic therapy as well as at different follow-up times, i.e. at 30 and at 90 days.

Length of hospital stay:

- Moderate quality evidence from 2 studies with 240 participants suggested there may be no clear overall advantage of ultrafiltration compared to diuretic therapy in length of hospital stay.
- Moderate quality evidence from 1 study with 36 participants showed that the median length of hospital stay was significantly ($p<0.019$) shorter for patients treated with ultrafiltration compared to diuretic therapy; however we are unable to comment on the uncertainty surrounding this outcome.
- Moderate quality evidence from 1 study with 188 participants showed that there was no clinically significant difference in the median length of hospital stay for ultrafiltration compared to diuretic therapy; however we are unable to comment on the uncertainty surrounding this outcome.

Readmission rate:

- Low quality evidence from 2 studies with 219 participants suggested there may be fewer readmissions to hospital due to any cause at a follow up of 60-90 days in the diuretic therapy group compared to ultrafiltration. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit or harm. When sub grouped by length of follow up time low quality evidence from 1 study with 183 participants suggested there may be fewer readmissions to hospital due to any cause at a follow up of 60 days in the diuretic therapy group compared to ultrafiltration. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit or harm whereas at 90 days very low quality evidence from 1 study with 36 participants suggested there may be no clear advantage for either strategy.

- Very low quality evidence from 2 studies with 359 participants suggested there may be fewer readmissions to hospital due to heart failure at a follow up of 60-90 days in the ultrafiltration group compared to diuretic therapy. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit or harm. When sub grouped by length of follow up time very low quality evidence from 1 study with 183 participants suggested there was no clear advantage of either strategy in reducing readmissions to hospital due to heart failure at 60 days follow up. However at 90 days low quality evidence from 1 study with 176 participants suggested there may be fewer readmissions to hospital due to heart failure in the ultrafiltration group compared to diuretic therapy. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit or harm.

Dyspnoea:

- Moderate quality evidence from 1 study with 188 participants suggested there may be no clear advantage of ultrafiltration compared to diuretic therapy in the change in score on dyspnoea 100mm VAS from baseline at 96 hours.
- Very low quality evidence from 1 study with 163 participants suggested that ultrafiltration may improve the mean dyspnoea score at 48 hours compared to diuretic therapy. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit or harm.
- Low quality evidence from 1 study with 30 participants suggested that ultrafiltration may improve the NYHA score at 36 hours compared to diuretic therapy. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit or harm.

Clinical Status:

- Very low quality evidence from 1 study with 188 participants suggested that there may be no clear advantage of ultrafiltration compared to diuretic therapy in the change in score on global well-being scale from baseline at 96 hours.
- Very low quality evidence from 1 study with 162 participants suggested that there may be no clear advantage of ultrafiltration compared to diuretic therapy in the number of patients achieving clinical decongestion at 96 hours.

Weight Loss:

- Very low quality evidence from 2 studies with 201 participants suggested that there may be increased weight loss from baseline (kg) within 48 hours with ultrafiltration compared to diuretic therapy. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit or harm. At 72 hours there was very low evidence from 1 study with 40 participants indicating a clinical benefit in favour of ultrafiltration, but there was uncertainty around this effect. At 96 hours and 60 days moderate quality evidence from 1 study with 188 participants showed no clear advantage of either strategy.
- Low quality evidence from 1 study with 30 participants suggested that there may be increased weight loss based on a percentage of baseline weight (kg) at 36 hours with ultrafiltration compared to diuretic therapy. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit or harm.

Renal Function:

Very low quality evidence from 1 study with 30 participants suggested that there may be no clear advantage of ultrafiltration compared to diuretic therapy in the change in serum creatinine.

- ($\mu\text{mol/l}$) at 36 hours. Very low quality evidence from 1 study at 72 hours suggested a possible effectiveness of ultrafiltration compared to diuretic therapy in decreasing serum creatinine level. Low quality evidence from 2 studies with 224 participants suggested there may be smaller increases in serum creatinine ($\mu\text{mol/l}$) at 96 hours with diuretic therapy compared to ultrafiltration. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit or harm. Low quality evidence from 1 study with 188 participants suggested that there may be no clear advantage of ultrafiltration compared to diuretic therapy in the change in serum creatinine ($\mu\text{mol/l}$) at 60 days.
- Low quality evidence from 1 study with 181 participants suggested that there may be fewer patients with a rise in serum creatinine $>26.5\mu\text{mol/l}$ in the diuretic therapy group compared to ultrafiltration at 24 hours. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit or harm. Low quality evidence from 1 study with 178 participants suggested that there may be fewer patients with a rise in serum creatinine $>26.5\mu\text{mol/l}$ in the diuretic therapy group compared to ultrafiltration at 48 hours. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit or harm.

Serious Adverse Events:

- Very low quality evidence from 2 studies with 220 participants suggested that there was no clear advantage of ultrafiltration compared to diuretic therapy in numbers of patients with heart failure SAEs, other cardiovascular SAEs at 60 days follow up or renal failure / dialysis dependence SAEs (at hospital discharge and 60 days). Low quality evidence from the same study suggested there may be fewer patients with any SAE in the diuretic therapy group compared to ultrafiltration. However, the uncertainty around this effect was too large to draw conclusions about clear clinical benefit or harm.

Ultrafiltration +/- diuretic therapy versus medical care

All-cause mortality:

- Very low quality evidence from 2 studies with 96 participants showed similar rates of all-cause mortality in ultrafiltration +/- diuretic therapy compared to usual care at 30 day and 1 year follow up.

Rehospitalisation rate due to congestive heart failure:

- Moderate quality evidence from 1 study with 56 participants showed that ultrafiltration was effective in lowering the rate of rehospitalisations at 1 year.

Length of hospital stay:

Low quality evidence from 2 studies with 96 participants showed no clear difference in length of hospital stay.

Weight Loss:

- Very low quality evidence from 1 study with 37 participants showed that ultrafiltration and medical care were associated with similar decreases in body weight at hospital discharge.

Renal Function:

- Low quality evidence from 1 study with 56 participants suggested that there may be no clear advantage of ultrafiltration compared to diuretic therapy in the change in serum creatinine

($\mu\text{mol/l}$) at hospital discharge and 1 year. However, the same study provided low quality for a decrease in serum creatinine level associated with ultrafiltration at 6 months with some uncertainty about this effect.

Dyspnoea:

- Very low quality evidence from 1 study with 38 participants suggested that there was no clear advantage of ultrafiltration +/- diuretic therapy compared to usual care in the number of patients with marked, moderate or mild improvement in dyspnoea at 24 hours.

Clinical Status:

- Very low quality evidence from 1 study with 38 participants suggested that there was no clear advantage of ultrafiltration +/- diuretic therapy compared to usual care in the number of patients with marked, moderate or mild improvement in global symptoms at 24 hours.

Economic

- One study found that ultrafiltration was more costly and more effective than intravenous diuretics (£30,400 per death averted) for patients with acute heart failure. This study was assessed as partially applicable with potentially serious limitations.

8.3.4 Recommendations and link to evidence

24. Do not routinely offer ultrafiltration to people with acute heart failure.	
Recommendations	25. Consider ultrafiltration for people with confirmed diuretic resistance ^g .
Relative values of different outcomes	The GDG regarded mortality and rehospitalisation for heart failure as the key outcomes. Other important outcomes were weight loss, changes in dyspnoea, clinical status, measures of renal function and length of hospital stay.
Trade-off between clinical benefits and harms	There was no difference in the mortality rate between people treated with ultrafiltration or diuretic therapy. Ultrafiltration was associated with fewer rehospitalisations due to heart failure at 60 – 90 days follow-up compared to diuretic therapy, but readmissions due to all causes were lower in the diuretic therapy group. Ultrafiltration led to increased weight loss and lower serum creatinine up to 72 hours, but at 60 days these parameters both favoured diuretic therapy. Ultrafiltration was associated with more serious adverse events than diuretic therapy.
Economic considerations	Ultrafiltration is more costly than diuretic therapy. Any savings from reduced hospital care are unlikely to fully offset the additional cost of equipment and consumables. Furthermore, the economic analysis may have under-estimated the costs associated with ultrafiltration since it did not take into account nursing levels or the need for more than one ultrafiltration filter.

^g Diuretic resistance is defined as dose escalation beyond a person's previously recognised dose ceiling or a dose approaching the maximum recommended daily dose without incremental improvement in diuresis. From Diuretics and ultrafiltration in acute decompensated heart failure.

	24. Do not routinely offer ultrafiltration to people with acute heart failure.
Recommendations	25. Consider ultrafiltration for people with confirmed diuretic resistance^g.
Quality of evidence	The evidence was rated as moderate to very low according to the GRADE criteria. The two largest studies (UNLOAD and CARRESS-HF) were given the most emphasis in the discussion. These had different study populations: CARRESS recruited patients with acute decompensated heart failure with worsened renal function (defined as an increase of serum creatinine of at least 26.5µmol/l within 12 weeks before or 10 days after index admission); UNLOAD had no pre-defined renal function inclusion criteria. Both these studies and Hanna 2012 were set in the US, so may not be directly applicable to the UK NHS setting. In particular, lengths of stay for heart failure are shorter in the US, so these data were treated with caution. In addition some study results could not be pooled due to lack of detail provided in the study manuscripts.
Other considerations	<p>Ultrafiltration allows faster fluid removal. However, over longer time periods it is not associated with better outcomes in terms of mortality or total hospital readmissions. Ultrafiltration was safe for the majority of patients, although a higher rate of bleeding and venous access complications were seen. The adverse event reporting was not complete in all the studies.</p> <p>The two main trials used new technology allowing smaller gauge intravenous access. Therefore, the findings may not be directly applicable to centres using haemofiltration type machines, which need large bore central venous access. The CARRESS-HF study used an aggressive, highly structured diuretic regime which is not standard diuretic practice in the UK. The use of ultrafiltration following failure of diuretics has not been studied. In the studies salt and water intake was restricted to 2l/day of fluid and 2g/day sodium. The addition of ultrafiltration to diuretics did not produce a clinically important difference in outcomes.</p> <p>The GDG agreed that the current evidence does not support a routine strategy of ultrafiltration in patients with acute heart failure. Ultrafiltration is a more costly therapy and has not been evaluated in a UK setting. The GDG considered scenarios in which ultrafiltration may be justified. It was agreed that prior to its use specialist input was required to ensure that medical treatment has been optimised. The GDG acknowledged that it is used in carefully selected patients within the UK, for example, in those with particularly large volumes of fluid to remove or where diuretic resistance is a problem. In this context the definition of diuretic resistance as provided by Felker and Mentz 2012^{74,75} was agreed to be appropriate: "Dose escalation beyond a patient's previously recognised dose ceiling or a dose approaching the maximum recommended daily dose without incremental improvement in diuresis". It was recognised that some centres might add a thiazide diuretic before concluding that diuretic resistance had been observed.</p> <p>In the UK, ultrafiltration is used as a rescue type therapy in patients that need to offload fluid quickly in a level 2 care setting. It is currently provided by cardiologists, nephrologists and critical care physicians in a relatively small number of centres, possibly leading to inequality of access.</p> <p>The GDG agreed to make a research recommendation to study the effectiveness of ultrafiltration as used in the UK. This would allow a future update of the recommendation on ultrafiltration to be informed by UK data.</p>

9 Treatment after stabilisation

9.1 Timing of beta blocker therapy

The use of beta-blockers in the treatment of patients with heart failure due to left ventricular systolic dysfunction has resulted in significant reductions in morbidity and mortality. However, when these patients are admitted with an acute exacerbation of their chronic heart failure, beta-blockers may be discontinued for fear of the acute exacerbation being either caused or worsened by the negative inotropic effects of beta-blockers on the myocardium. These patients may then be labelled as intolerant of beta-blockers and may not be re-started on these agents. There is concern that this practice may be preventing patients receiving beta-blockers in either the short or long term who would otherwise benefit from this treatment. There is also uncertainty as to whether beta-blocker medications should be commenced prior to leaving hospital, or later, in people presenting with acute heart failure not already taking a beta-blocker. Guidance is therefore needed with regards the administration of beta-blockade in patients who have presented with either de-novo heart failure or an acute exacerbation of chronic heart failure.

Review question 1: In people with acute heart failure already on beta-blocker therapy should beta-blockers be reduced or discontinued, and if so should they be reinstated in hospital after stabilisation?

Review question 2: For people with confirmed acute heart failure not already on beta-blocker therapy should beta-blocker treatment commence in hospital after stabilisation or following discharge?

For full details see review protocols in Appendix C.

Table 71: PICO characteristics of review question (the number respond to characteristics of review question 1 or review question 2)

Population	Adults with acute heart failure (1) already on beta-blockers (2) not already on beta-blockers
Intervention	(1) Continuation (2) Commencing beta-blocker therapy in hospital
Comparison	(1) Discontinuation (or reduction of dose) of beta-blocker therapy (2) Commencing beta-blocker therapy after discharge
Outcomes	Mortality Major cardiovascular events Length of hospital stay Re-admission rates and re-admission to critical care units Quality of Life Change in renal function Rate of patients receiving beta-blocker treatment at follow-up Adverse events (hyperkalaemia, cough, symptomatic hypotension)
Study design	Systematic reviews, randomised controlled trials (no particular year or sample size restrictions) and observational studies (n>2000)

9.1.1 Clinical evidence

We searched for randomised controlled trials (RCTs), systematic reviews and observational studies that reported on beta-blocker therapy addressing timing of beta-blocker treatment. The review topics consider timings with respect to continuation or discontinuation of beta-blocker treatment on

admission to hospital for people who are already taking beta-blockers, and for patients without prescriptions on whether to commence beta-blocker therapy in hospital or after discharge. For the comparison of continuation and discontinuation, one RCT (Jeandeau et al, 2009 – B-CONVINCED¹⁰⁷) was identified as well as one observational study (Fonarow et al, 2008^{79,82} using data from the OPTIMIZE-HF registry). The question of when to commence beta-blocker treatment was addressed by one RCT (Gattis et al, 2004^{85,86} IMPACT-HF trial) and five observational studies^{9,41,67,81,86} from 4 large registries or audits. All these studies only provided evidence for timing of commencement of medication at discharge and it was not made clear whether those people who did not commence beta-blockers at discharge were administered the treatment at a later stage. It is therefore an indirect comparison and therefore does not completely satisfy the protocol. One of the studies⁶⁷ presented results for two subgroups of heart failure patients separately into ejection fraction of more than or less than 50%. See also the study selection flow chart in Appendix D and exclusion list in Appendix K.

Summary of included studies

The characteristics of included studies are briefly outlined in Table 72 for continuation compared against discontinuation, and Table 73 for commencement of beta-blockers – for details please see Appendix G.

Table 72: Summary of studies included in the review for continuation compared to discontinuation of beta-blocker therapy on admission to hospital

Study	Type of study	Population	Outcomes	Comments
Fonarow 2008 ^{79,82}	Data from the OPTIMIZE-HF – a US national registry	N=5,791 patients hospitalised for episodes of new or worsening HF as the primary discharge diagnosis. N=2,373 patients were eligible for beta-blockers at discharge: N=1,350 who were receiving beta-blockers and continued and N=79 in which therapy was withdrawn	Only mortality was analysed using a multivariate statistical method (i.e. accounting for patient differences)	Propensity score analysis and multivariate adjustment. Both are statistical techniques that account for important baseline differences between groups of patients who continued and people whose beta-blocker therapy was stopped.
Jondeau et al, 2009 ¹⁰⁷	B-CONVINCED randomised controlled trial	N=169 patients with congestive heart failure hospitalised for a decompensation episode	Mortality (in-hospital and at 3 months), improvement in dyspnoea and well-being, re-hospitalisation within 3 months, length of stay	The authors stated that in 50% of the patients, the average dose of the beta-blockers used was 50% of the recommended target dose level according to the European Society of Cardiology Guidelines.

Table 73: Summary of studies included in the review of commencing beta-blocker treatment in hospital or after discharge

Study	Type of study	Population	Outcomes	Comments
Amed et al,	Data from the	N=5,791 patients hospitalised	Mortality (6-year	Long-term follow-up

Study	Type of study	Population	Outcomes	Comments
2011 ⁹	OPTIMIZE-HF — a US national registry	with new-onset or worsening heart failure (HF) as the primary cause of admission	follow-up)	analysis of the OPTIMIZE-HF registry – multivariate adjusted analysis. Only available as a conference abstract.
Cleland et al, 2012 / 2013 ⁴¹	Longitudinal national heart failure audit report of with the latest covering 145 out of 150 NHS Trusts in England and Health Boards in Wales	In 2012/13 this involved N=36,788 index admissions and N=7,106 readmissions. For the 4 year follow-up results N=93953 records With regards to beta blockers at discharge N=20,099 received a prescription at discharge whereas N=6,056 did not when recorded at 30 day follow-up N=57,952 received a prescription at discharge whereas N=25,765 did not when recorded at 4-year follow-up.	All-cause mortality at 30 day and 4 year follow-up	A Cox proportional hazard model was employed which adjusted risks according to multiple variables selected from a literature review rather than based on statistical significance.
Ezekowitz et al, 2008 ⁶⁷	Data from the EFFECT registry, which prospectively enrolled HF patients admitted to 103 hospitals in Ontario, Canada	N=9,943 newly admitted patients with a primary discharge diagnosis of HF. After exclusions (such as died during hospital stay or medications not recorded) patients were subdivided into >50% ejection fraction (N=1,026) or <50% ejection fraction (N=1,898)	Composite endpoint of mortality or heart failure readmission to hospital (1 year)	Difficult to disentangle whether the results are due to mortality or readmission since it is a composite endpoint.
Fonarow et al, 2007 ^{79,81}	Data from the OPTIMIZE-HF — a US national registry	N=2,333 of patients were eligible for beta-blockers at discharge	Mortality at 60- to 90-day follow-up analysed using multivariate statistical methods	Risk and propensity adjusted model (to account for patient differences between groups)
Gattis et al 2004 ^{85,86}	IMPACT-HF randomised controlled trial	N=363 hospitalised patients with a primary diagnosis of HF and left ventricular ejection fraction ≤40% within the previous 12 months	Mortality, rehospitalisation, length of hospital stay	60-day mortality or rehospitalisation rate was 25% which the authors in part attributed to a rate of 10% of patients being discharged with symptoms of congestion (i.e. rales).

Table 74: GRADE evidence profile – quality of evidence and summary of findings. Review question 1: Beta-blocker continuation compared to discontinuation (or reduction of dose)

Quality assessment							Summary of findings				Quality	Importance
No of studies Follow-up	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Beta-blocker discontinued Number of event / Total N (%) or Mean (SD)	Beta-blocker continued	Effect			
									Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect or Mean Difference (95% CI)		
Mortality (RCT data) ^{85,107}												
1 In-hospital	RCT	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	2/78 (2.6%)	1/69 (1.4%)	RR 1.77 (0.16 to 19.09)	11 more per 1000 (from 12 fewer to 262 more)	LOW	CRITICAL
1 3 months	RCT	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	6/78 (7.7%)	6/69 (8.7%)	RR 0.88 (0.3 to 2.62)	10 fewer per 1000 (from 61 fewer to 141 more)		
Mortality (observational data) 60-90 days ^{79,80}												
1	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	19/79 (24.1%)	117/1350 (8.7%)	HR 2.34 (1.2 to 4.55)	104 more per 1000 (from 16 more to 251 more)	LOW	CRITICAL
Re-hospitalisation within 3 months - for heart failure ¹⁰⁷												
1	RCT	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ^(a)	none	24/78 (30.8%)	15/69 (21.7%)	RR 1.42 (0.81 to 2.47)	91 more per 1000 (from 41 fewer to 320 more)	MODERATE	IMPORTANT
Re-hospitalisation within 3 months - for arrhythmia ¹⁰⁷												
1	RCT	no serious	no serious inconsistency	no serious indirectness	very serious ^(a)	none	3/78 (3.8%)	2/69 (2.9%)	RR 1.33 (0.23 to	10 more per 1000 (from	LOW	IMPORTANT

Quality assessment							Summary of findings				Quality	Importance
No of studies Follow-up	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Beta-blocker discontinued Number of event / Total N (%) or Mean (SD)	Beta-blocker continued	Effect			
									Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect or Mean Difference (95% CI)		
		risk of bias							7.71)	22 fewer to 194 more)		
Re-hospitalisation within 3 months – for other reasons ¹⁰⁷												
1	RCT	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	9/78 (11.5%)	10/69 (14.5%)	RR 0.8 (0.34 to 1.84)	29 fewer per 1000 (from 96 fewer to 122 more)	LOW	IMPORTANT
Improvement in dyspnoea and well-being (day 8) - Physician-rated (blinded) ¹⁰⁷												
1	RCT	serious ^(b)	no serious inconsistency	no serious indirectness	no serious imprecision	none	74/78 (94.9%)	66/69 (95.7%)	RR 0.99 (0.92 to 1.07)	10 fewer per 1000 (from 77 fewer to 67 more)	MODERATE	IMPORTANT
Improvement in dyspnoea and well-being (day 8) - Self-rated ¹⁰⁷												
1	RCT	serious ^(b)	no serious inconsistency	no serious indirectness	no serious imprecision	none	74/78 (94.9%)	65/69 (94.2%)	RR 1.01 (0.93 to 1.09)	9 more per 1000 (from 66 fewer to 85 more)	MODERATE	IMPORTANT
Length of hospital stay ¹⁰⁷												
1	RCT	no serious risk of bias	no serious inconsistency	no serious indirectness	No serious imprecision	none	10.4 (9.7)	11.5 (8.3)	-	MD 1.10 lower (4.01 lower to 1.81 higher)	HIGH	IMPORTANT
Rate of patients taking beta-blockers since discharge at 3 months follow-up ¹⁰⁷												
1	RCT	no serious	no serious inconsistency	no serious indirectness	serious ^(a)	none	59/78 (75.6%)	62/69 (89.9%)	RR 0.84 (0.73 to	144 fewer per 1000	MODERATE	IMPORTANT

Quality assessment							Summary of findings				Quality	Importance
No of studies Follow-up	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Beta-blocker discontinued Number of event / Total N (%) or Mean (SD)	Beta-blocker continued	Effect			
									Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect or Mean Difference (95% CI)		
		risk of bias							0.98)	(from 18 fewer to 243 fewer)		

- (a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed one of the default MIDs (i.e. ranging from a noticeable effect to no effect). Outcomes were downgraded by two increments if 95% CI of the effect crossed both default MIDs (i.e. ranging from a noticeable improvement to a noticeable harm). Default MIDs are set at RRs / HRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the median control group standard deviation either side of the null line for continuous variables.
- (b) It was unclear how this outcome was rated either by the clinician or the patient. No validated scale was used.

Table 75: GRADE evidence profile - quality of evidence by and summary of findings. Review question 2: Beta-blockers in hospital versus early or possibly started after discharge for acute heart failure

Quality assessment							Summary of Finding			Quality	Importance
No. of	Design	Risk of	Inconsistency	Indirectness	Imprecision	Other	Beta-	Clearly or	Effect		

studies Follow -up time		bias				considerations	blockers in hospital Number of event / Total N (%) or Mean (SD)	possibly started after discharge Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect or Mean Difference (95% CI)		
Mortality (RCT) ^{85,86}												
1 60 days	RCT	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	6/185 (3.2%)	8/178 (4.5%)	RR 0.72 (0.26 to 2.04)	13 fewer per 1000 (from 33 fewer to 47 more)	LOW	CRITICAL
Mortality HR (observational data) ^{9,40,80}												
1 30 days	observational studies	no serious risk of bias	no serious inconsistency	serious ^(b)	no serious imprecision	none	602/19785 (3%)	266/5854 (4.5%)	HR 0.74 (0.62 to 0.88)	12 fewer per 1000 (from 5 fewer to 17 fewer)	VERY LOW	CRITICAL
1 60-90 days	observational studies	no serious risk of bias	no serious inconsistency	serious ^(b)	no serious imprecision	none	114/1959 (5.8%)	42/374 (11.2%)	HR 0.48 (0.31 to 0.74)	57 fewer per 1000 (from 28 fewer to 76 fewer)	VERY LOW	CRITICAL
1 4 year	observational studies	no serious risk of bias	no serious inconsistency	serious ^(b)	no serious imprecision	none	19944/57952 (34.4%)	12946/257 65 (50.2%)	HR 0.79 (0.76 to 0.89)	79 fewer per 1000 (from 40 fewer to 91 fewer)	VERY LOW	CRITICAL
1 6 years	observational studies	serious ^(c)	no serious inconsistency	serious ^(b)	no serious imprecision	none	2334/3382 (69%)	2435/3382 (72%)	HR 0.9 (0.85 to 0.95)	38 fewer per 1000 (from 18 fewer to 59 fewer)	VERY LOW	CRITICAL
Mortality or heart failure hospitalisation - Ejection fraction > 50% ⁶⁷												
1 1 year	observational studies	no serious risk of bias	no serious inconsistency	serious ^(e)	very serious ^(a)	none	n/a ^(e)	n/a ^(f)	HR 1.02 (0.75 to 1.39)	20 more per 1000(from 290 fewer to 330 more) ^(g)	VERY LOW	CRITICAL

Quality assessment							Summary of Finding				Quality	Importance
No. of studies Follow-up time	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Beta-blockers in hospital Number of event / Total N (%) or Mean (SD)	Clearly or possibly started after discharge Number of event / Total N (%) or Mean (SD)	Effect			
									Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect or Mean Difference (95% CI)		
Mortality or heart failure hospitalisation - Ejection fraction < 50% ⁶⁷												
1 year	observational studies	no serious risk of bias	no serious inconsistency	serious ^(e)	no serious imprecision	none	n/a ^(e)	n/a ^(f)	HR 0.8 (0.72 to 0.89)	220 fewer per 1000(from 330 fewer to 120 fewer) ^(g)	VERY LOW	CRITICAL
Withdrawal due to serious adverse events ^{85,86}												
1 60 days	RCT	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	7/555 (1.3%)	4/534 (0.75%)	Peto OR 1.67 (0.51 to 5.46)	10 more per 1000 (from 10 fewer to 20 more) ^(g)	LOW	IMPORTANT
Withdrawal due to serious adverse events - Hypotension ^{85,86}												
1 60 days	RCT	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	3/185 (1.6%)	1/178 (0.56%)	Peto OR 2.64 (0.37 to 18.88)	10 more per 1000 (from 10 fewer to 30 more) ^(h)	LOW	IMPORTANT
Withdrawal due to serious adverse events - Bradycardia ^{85,86}												
1 60 days	RCT	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	3/185 (1.6%)	0/178 (0%)	Peto OR 7.19 (0.74 to 69.61)	20 more per 1000 (from 0 more to 40 more) ^(h)	LOW	IMPORTANT
Withdrawal due to serious adverse events - Worsening heart failure ^{85,86}												

Quality assessment							Summary of Finding				Quality	Importance
No. of studies Follow-up time	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Beta-blockers in hospital Number of event / Total N (%) or Mean (SD)	Clearly or possibly started after discharge Number of event / Total N (%) or Mean (SD)	Effect			
									Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect or Mean Difference (95% CI)		
1 60 days	RCT	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	1/185 (0.54%)	3/178 (1.7%)	Peto OR 0.35 (0.05 to 2.51)	10 fewer per 1000 (from 30 fewer to 10 more) ^(h)	LOW	IMPORTANT
Re-hospitalisation ^{85,86}												
1 60 days	RCT	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ^(a)	none	40/185 (21.6%)	45/178 (25.3%)	RR 0.86 (0.59 to 1.24)	35 fewer per 1000 (from 104 fewer to 61 more)	MODERATE	IMPORTANT
Rate of patients taken beta-blockers since discharge ^{85,86}												
1 60 days	RCT	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ^(a)	none	165/185 (89.2%)	130/178 (73%)	RR 1.22 (1.10 to 1.35)	161 more per 1000 (from 73 more to 256 more)	MODERATE	IMPORTANT

(a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed one of the default MIDs (i.e. ranging from a noticeable effect to no effect). Outcomes were downgraded by two increments if 95% CI of the effect crossed both default MIDs (i.e. ranging from a noticeable improvement to a noticeable harm). Default MIDs are set at RRs / HRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the median control group standard deviation either side of the null line for continuous variables.

(b) These studies report on prescriptions at discharge only (i.e. early). It is unclear whether or not the group of patients without prescription later receive this medication (i.e. only indirectly addressing the protocol).

(c) In this observational study it was unclear why only a very limited number of variables were used in the multivariable analysis.

(d) This is long term follow-up data from a registry. Data is only presented in a conference abstract and the level of detail was insufficient to determine factors such as levels of attrition etc.

(e) Only a composite outcome of mortality or heart failure re-hospitalisation was reported.

- (f) The exact numbers of patients who had or had not received beta-blockers at discharge was not clear and therefore absolute numbers could not be derived and the risk difference was calculated.*
- (g) The exact numbers for the outcome were not provided and therefore the risk difference was calculated from the hazard ratio.*
- (h) Due to low number of events in either of the control or intervention arms the Peto OR was used (which is more robust in cases of small event rates) a risk difference was then calculated*

9.1.2 Economic evidence

Published literature

Review question 1: Should beta-blockers be reduced or discontinued?

No relevant economic evaluations were identified.

Review question 2: Should beta-blocker treatment commence in hospital after stabilisation or following discharge?

No relevant economic evaluations were identified.

New cost-effectiveness analysis

New analysis was not prioritised for this area.

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs were provided in Appendix N to aid consideration of cost-effectiveness.

Economic considerations

The economic consequences of the discontinuation or later commencement of beta-blockers in admitted acute heart failure patients should be considered over an appropriately long time horizon because the health effect of beta-blockers is attained over a long time period.

Differences in up-front costs between the strategies are very low (the cost of a 14-day course of Bisoprolol is £0.55) so they can be excluded in the consideration of cost-effectiveness. Downstream costs from any long-term gains in morbidity and mortality, such as the cost of acute admissions of heart failure, are the primary consideration.

9.1.3 Evidence statements

Clinical

Review question 1: Beta-blockers continuation compared to discontinuation

Mortality (RCT data)

Low quality evidence from one randomised controlled trial (RCT) (N=147) showed similar mortality rates for people who continued or discontinued with beta-blocker medication during hospital stay and at 3-month follow-up.

Mortality (observational studies)

Low quality evidence from one observational study (N=1,429 patients eligible for beta-blocker treatment) showed a reduced rate of mortality associated with beta-blocker continuation rather than discontinuation.

Rehospitalisation

Moderate to low quality evidence from one RCT (N= 147) showed similar rates of re-admission rates (for heart failure, arrhythmia and other reasons) in patients who continued compared to patients who discontinued beta-blocker treatment.

Improvement in dyspnoea and well-being

Moderate quality evidence from one RCT (N=147) showed similar rates of improvement in dyspnoea and well-being for people who continued compared to patients who discontinued beta-blocker treatment (at day 8). This was regardless of whether it was rated by a physician blinded to the treatment or by the patient.

Rate of patients receiving beta-blockers at 3 months

Moderate quality evidence from one RCT (N=147) showed higher rates of beta-blocker treatment at 3 months for people who continued compared to patients who discontinued beta-blocker treatment.

Length of hospital stay

High quality evidence from one RCT (N=147) indicated no clear difference in the length of hospital stay for people who continued beta-blocker therapy compared to patients who discontinued the beta-blocker treatment.

Review question 2: Beta-blockers in hospital compared to clearly or possibly prescribed after discharge

Mortality (RCT data)

Low quality evidence from one RCT (N=363) indicated similar rates of mortality for people receiving beta-blocker in hospital compared to those who started beta-blocker therapy after discharge.

Mortality (observational studies)

Very low quality evidence from three observational large scale registry studies (N ranged from 2,333 to 93,953) showed a reduction in mortality and longer length of survival associated with beta-blocker medication prescribed at discharge compared to no or possible later beta-blocker treatment. This was the case at all follow-up points from 30 days to 6 years.

Mortality or heart failure hospitalisation (observational study)

Very low quality evidence from one observational study (N=1,026 with ejection fraction >50% and N=1,898 <50% ejection fraction) showed that beta-blocker medication at discharge was associated with a reduction in mortality or rehospitalisation only in patients with an ejection fraction <50% whereas the rates of mortality were similar for people with an ejection fraction >50% regardless of beta-blocker prescription at discharge.

Withdrawal due to serious adverse events

Low quality evidence from one RCT (N=363) showed similar overall withdrawal rates due to adverse events for people given beta-blockers in hospital compared to those who received this treatment after discharge. When dividing this into individual categories of adverse events which were hypotension, bradycardia and worsening heart failure the numbers of events were too small to draw clear conclusions about benefits and harms from either strategy.

Rehospitalisation

Low quality evidence from one RCT (N=363) showed similar rates of re-hospitalisations (at 60 day follow-up) for people who received beta-blockers in hospital compared to people prescribed beta-blocker therapy after discharge.

Rate of patients receiving beta-blockers at 60 days

Moderate quality evidence from one RCT (N=363) showed that people who were prescribed a beta blocker whilst in hospital were more likely than those whose prescription was deferred until after discharge to be taking a beta blocker 60 days after discharge..

Economic

No relevant economic evaluations were identified.

9.1.4 Recommendations and link to evidence

Recommendations	<p>26. In a person presenting with acute heart failure who is already taking beta-blockers, continue the beta-blocker treatment unless they have a heart rate less than 50 beats per minute, second or third degree atrioventricular block, or shock.</p> <p>27. Start or restart beta-blocker treatment during hospital admission in people with acute heart failure due to left ventricular systolic dysfunction, once their condition has been stabilised – for example, when intravenous diuretics are no longer needed.</p> <p>28. Ensure that the person's condition is stable for typically 48 hours after starting or restarting beta-blockers and before discharging from hospital.</p> <p>29. Closely monitor the person's renal function, electrolytes, heart rate, blood pressure and overall clinical status during treatment with beta-blockers, aldosterone antagonists or angiotensin-converting enzyme inhibitors.</p>
Relative values of different outcomes	The GDG regarded mortality as the most important outcome, but also considered adverse events such as bradycardia, hypotension and withdrawal of beta-blockers. The GDG were also interested in beta-blocker use at follow-up and the incidence of rehospitalisation. The GDG recognised the importance of length of hospital stay but gave this less weight due to the potential contribution of factors not directly related to heart failure. The GDG considered 'Improvement of dyspnoea and wellbeing', but raised concerns about the validity of this scale so attached less importance to this outcome. Health-related quality of life was not reported in any of the studies.
Trade-off between clinical benefits and harms	<p>Continuation of beta-blockers in an acute heart failure admission is associated with their increased use at a higher dose at 3 months follow-up without an increase in adverse events.</p> <p>Similarly, in-hospital initiation of beta blockers is associated with higher rates of beta blockers prescription at 60 days after discharge and no increase in adverse events.</p>
Economic considerations	<p>No relevant economic evaluations were identified.</p> <p>Continuation or initiation of beta blockers in hospital is associated with minimal additional cost, no adverse effects and significant long term benefit (since it is associated with greater long term use of beta blockers). On this basis this strategy was considered cost effective.</p>

Quality of evidence	<p>The evidence on continuation of beta blockers was rated low or very low quality on the GRADE criteria. Two studies were included; one randomised controlled trial and one observational study. The primary outcome of the RCT was clinician or self-rated improvement in dyspnoea or wellbeing using an invalidated six point rating scale. The sample size was sufficient to detect differences in this outcome, but was not sufficient to detect differences in any of the other reported outcomes. The only outcome given a high GRADE rating was 'length of hospital stay', which was not associated with any clinically important difference between continuation and discontinuation. The observational study was a multivariable analysis of registry data in which the vast majority (1350 of 1429) of patients continued beta blocker medication on admission. Although multivariable statistical techniques were used to compare patients who continued with those who discontinued, it was not possible to attribute the differences in mortality to beta blocker continuation alone.</p> <p>The evidence on initiation of beta blocker treatment also ranged from low to very low. One randomised controlled trial and 4 non-randomised studies were included in this review. The primary outcome of the RCT was the rate of beta blocker prescription at 3 months, which was higher in the in-hospital initiation group. While the 4 observational studies all pointed towards reduced mortality in patients commenced on beta blockers in hospital, the strong risk of confounding meant that little weight could be attached to these outcomes. With the exception of the National Heart Failure audit, all the studies were based in North America. It is also worth considering that despite the corrections for case-mix in the National Audit data, there may be bias to patient selection for early re-starting of Beta Blockers.</p> <p>The randomised controlled trials were set in the US where there are important differences in treatment of heart failure such as shorter length of hospital stay.</p>
Other considerations	<p><u>Continuation of beta-blockers versus cessation in an acute heart failure admission due to left ventricular systolic dysfunction</u></p> <p>Current UK practice varies. The GDG considered that the evidence demonstrates a clear signal for lack of harm when continuing beta-blocker treatment and that such an approach is associated with improved longer term use. This is important, because in chronic heart failure use of beta blockers is associated with improved survival and reduced rates of hospitalisation.</p> <p>The RCT excluded patients with heart rate <50bpm, second or third degree atrioventricular block and shock. These criteria were felt to be appropriate reasons to stop beta blockers. However, there are individual clinical circumstances when the beta-blocker dose may be reduced rather than continued at the admission dose or stopped completely (e.g. relative bradycardia).</p> <p><u>In-hospital initiation or reintroduction of beta-blockers following admission with acute heart failure</u></p> <p>In-hospital introduction of beta-blockers is associated with increased use of beta blockers at follow-up without an increase in adverse events. Beta blockers should be started once the patient has been clinically stabilised. Typically, this might be when intravenous diuretics are no longer required. The GDG speculated that the relatively high 60 day mortality and rehospitalisation rate in the IMPACT-HF study (Gattis 2004), was in part due to the relatively short median length of stay of 5 days. In this study patients were discharged after a minimum of 12 hours following the first dose of beta blockers. With this in mind it was recommended that a period of patient observation of typically 48 hours is required following initiation of beta-blocker treatment to monitor for any clinical deterioration and to ensure tolerability. The GDG discussed the use of the word 'typically' at length and used it to deliberately give necessary flexibility to the clinician.</p>

	<p>Even if it is not possible to commence a beta-blocker due to intolerance or clinical circumstances during an acute admission, this should not preclude initiating therapy as an outpatient as soon as possible.</p> <p>The initial beta-blocker dose should be low and incrementally titrated upwards to minimise adverse effects. In-hospital initiation increases longer term use of beta-blockers which is recognised to have a number of beneficial effects in heart failure due to left ventricular systolic dysfunction (e.g. reduction in mortality, reverse LV remodelling, reduction in sudden cardiac death), but not in heart failure with preserved ejection fraction.</p> <p>There were concerns as to whether, as a consequence of staying on beta blockers, patients might be less likely to start or continue on angiotensin-converting enzyme inhibitors. There was some evidence from the BCONVINCED study that this may be the case (angiotensin-converting enzyme inhibitors were introduced in 5/69 patients in the continuation group; 12/78 in the discontinuation group), however as the numbers were small it is not possible to draw a definitive conclusion.</p> <p>The GDG consider that changes to treatment or introduction of new therapies would be made by the specialist team. The GDG discussed the need to monitor the patient's renal function, electrolytes, heart rate and blood pressure when introducing these treatments.</p>
--	--

9.2 Timing of ACE inhibitor and aldosterone antagonist therapy

There is a strong evidence-base for the use of angiotensin converting enzyme inhibitors (ACEI) in the treatment of patients with heart failure caused by left ventricular systolic dysfunction (HF-LVSD). In acute heart failure patients who have LVSD, and who are not already on an ACEI, one needs to establish the best timing for introducing these agents, whether early during the hospital admission or following discharge. This is particularly relevant when several pharmacological agents need to be introduced and some may fear side-effects or intolerance.

Review question 1: For people with confirmed acute heart failure not already on angiotensin converting enzyme (ACE)-inhibitor therapy, should ACE inhibitor therapy commence in hospital or following discharge?

The importance of the aldosterone antagonists, (spironolactone or eplerenone), in reducing the morbidity and mortality of patients with heart failure due to left ventricular systolic dysfunction (HF-LVSD) and in reducing hospital admissions of patients with heart failure with preserved left ventricular ejection fraction (HFPEF) is well-established. Several other agents have important roles to play in the treatment of patients with acute heart failure. There is uncertainty about the timing of introducing aldosterone antagonists to the treatment of patients with acute heart failure who are not already on these agents, during their hospital admission or after discharge.

Review question 2: For people with confirmed acute heart failure not already on aldosterone antagonists should aldosterone antagonist therapy commence in hospital after stabilization or following discharge?

A brief summary of the protocol is provided in Table 76 below. For full details see the review protocol in Appendix C.

Table 76: PICO characteristics of review question

Population	Adults with acute heart failure
Intervention/s	Commencing ACE inhibitor or aldosterone antagonist therapy in hospital
Comparison/s	Commencing ACE inhibitor or aldosterone antagonist therapy after discharge
Outcomes	Mortality Major cardiovascular events Length of hospital stay Re-admission rates and re-admission to critical care units Quality of Life Change in renal function Adverse events (hyperkalaemia, cough, symptomatic hypotension)
Study design	Systematic reviews (SRs), randomised controlled trials (RCTs) (no particular year or sample size restrictions) and observational studies (n>2000) so long as multivariate analysis is used

9.2.1 Clinical evidence

9.2.1.1 ACE inhibitor therapy

We searched for any study that reported on ACE inhibitor therapy that commenced in hospital compared against ACE inhibitor therapy that commenced after discharge. Three observational studies were identified^{40,41,67,67,149,149}. All of these studies^{40,67,149} compared people who commenced ACE inhibitor therapy at discharge with those who did not. However, the studies did not indicate whether those who did not commence the therapy at discharge were prescribed ACE inhibitor at a later stage. One of the studies presented results separately for two subgroups of heart failure patients: one with ejection fraction of more than 50% and the other with less than 50%. See also the study selection flow chart in Appendix D and exclusion list in Appendix K.

9.2.1.2 Aldosterone antagonist therapy

We searched for any study that compared aldosterone antagonist therapy that commenced in hospital with aldosterone antagonist therapy that commenced after discharge from hospital. Three observational studies were identified^{4,10,67}. One of the studies^{4,5} was a post hoc analysis of the EPHESUS randomised controlled trial^{175,176} (RCT), that compared earlier (within 7 days) with later (after 7 days) commencement of aldosterone antagonist therapy. The other two studies^{10,67} compared people who were prescribed aldosterone antagonists at discharge to those who were not. The studies did not indicate whether those who were not prescribed aldosterone antagonists at discharge were prescribed it at a later stage. One of the studies^{67,67} presented results for two subgroups of heart failure patients separately according to ejection fraction > 50% or < 50%. See also the study selection flow chart in Appendix D and exclusion list in Appendix K.

Summary of included studies

The characteristics of included studies are outlined in Table 77 for ACE inhibitor timing and in Table 78 for aldosterone antagonist timing— for details please see Appendix G.

Table 77: Summary of studies included in the ACE inhibitor review

Study	Type of study	Population	Outcomes	Comments
Cleland et al, 2012/13 ⁴¹	Longitudinal national heart failure audit report of with the	In 2012/13 this involved N=36,788 index admissions and N=7,106 readmissions. For the 4 year follow-up	Mortality (30 day and 4 year follow-up)	A Cox proportional hazard model was employed which adjusted risks

Study	Type of study	Population	Outcomes	Comments
	latest covering 145 out of 150 NHS Trusts in England and Health Boards in Wales	results N=93953 records With regards to ACE inhibitors at discharge N=20,550 received a prescription at discharge whereas N=5,082 did not when recorded at 30 day follow-up (ACEi or ARB). With regards to ACE inhibitors at discharge N=54,451 received a prescription at discharge whereas N=27,397 did not when recorded at 4-year follow-up.		according to multiple variables selected from a literature review rather than based on statistical significance.
Ezekowitz et al, 2008 ⁶⁷	Data from the EFFECT registry, which prospectively enrolled HF patients admitted to 103 hospitals in Ontario, Canada	N=9,943 after exclusions (such as died during hospital stay or medications not recorded) patients were subdivided into >50% ejection fraction (N=1,026) or <50% ejection fraction (N=1,898)	Composite endpoint of mortality or heart failure readmission to hospital (1 year)	Difficult to disentangle whether the results are due to mortality or readmission since it is a composite endpoint.
Mujib et al, 2013 ¹⁴⁹	Data from the OPTIMIZE-HF -a US national registry	N=4,189 patients hospitalised with new-onset or worsening HF who were eligible for new prescriptions for ACE inhibitors	Mortality and heart failure hospitalisation (2.4 year follow-up)	Propensity score analysis was used as a statistical techniques that accounts for important baseline differences between groups of patients who received an ACE inhibitor prescription at discharge and those who did not (leading to a cohort of 1337 pairs of patients who were similar on 114 baseline characteristics).

Table 78: Summary of studies included in the aldosterone antagonist review

Study	Type of study	Population	Outcomes	Comments
Adamopoulos et al, 2009 ^{4,5}	Post hoc analysis of data from the EPHEUS study (RCT) of	N=1,369 starting within 7 days and N=1,950 starting ≥7 days Patients with acute	All-cause mortality, sudden cardiac death and mortality	All endpoints were adjudicated by blinded clinical events committee. Adjusted hazard ratios were presented which used 25

Study	Type of study	Population	Outcomes	Comments
	aldosterone antagonist effectiveness	myocardial infarction complicated by clinical HF and LVSD (LVEF≤40%)	from cardiovascular (CV) causes or hospitalisation for CV events Mean follow-up length 16 months	patient characteristics in a multivariate analysis.
Ahmed et al, 2011 ^{9,10}	Data from the OPTIMIZE-HF — a US national registry	Of N=10,429 patients hospitalised with new-onset heart failure with preserved ejection fraction (LVEF ≥ 40%) aged ≥65 of which N=866 received a discharge prescription of aldosterone antagonists.	Mortality (2.4 year follow-up)	Propensity score analysis was used as a statistical techniques that accounts for important baseline differences between groups of patients who received an aldosterone antagonists prescription at discharge and those who did not (leading to a cohort of 864 pairs of patients who were similar on 116 baseline characteristics).
Ezekowitz et al, 2008 ⁶⁷	Data from the EFFECT registry, which prospectively enrolled HF patients admitted to 103 hospitals in Ontario, Canada	N=9,943 after exclusions (such as died during hospital stay or medications not recorded) patients were subdivided into >50% ejection fraction (N=1,026) or <50 ejection fraction (N=1,898)	Composite endpoint of mortality or heart failure readmission to hospital (1 year)	Difficult to disentangle whether the results are due to mortality or readmission since it is a composite endpoint.

Table 79: GRADE evidence profile - quality of evidence by and summary of findings. ACE-inhibitors in hospital versus clearly or possibly started after discharge for acute heart failure.

Quality assessment							Summary of Finding				Quality	Importance
No of studies Follow-up time	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	ACE inhibitors in hospital Number of event / Total N (%) or Mean (SD)	Clearly or possibly started after discharge Number of event / Total N (%) or Mean (SD)	Effect			
									Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect or Mean Difference (95% CI)		
All-cause mortality ^{41,149}												
1 30 day	observational studies	no serious risk of bias	no serious inconsistency	serious ^(b)	no serious imprecision	none	769/20550 (3.7%)	513/5082 (10.1%)	HR 0.50 (0.42 to 0.59)	49 fewer per 1000 (from 40 fewer to 57 fewer)	VERY LOW	CRITICAL
1 2.4 years	observational studies	no serious risk of bias	no serious inconsistency	serious ^(b)	no serious imprecision	none	930/1337 (69.6%)	951/1337 (71.1%)	HR 0.96 (0.88 to 1.05)	15 fewer per 1000 (from 46 fewer to 17 more)	VERY LOW	CRITICAL
1 4 years	observational studies	no serious risk of bias	no serious inconsistency	serious ^(b)	no serious imprecision	none	19084/54451 (35%)	12880/27397 (47%)	HR 0.67 (0.65 to 0.69)	124 fewer per 1000 (from 115 fewer to 132 fewer)	VERY LOW	CRITICAL
Mortality or heart failure hospitalisation - Ejection fraction > 50% ⁶⁷												
1 1 year	observational studies	no serious risk of bias	no serious inconsistency	serious ^(c)	very serious ^(c)	none	n/a ^(d)	n/a ^(d)	HR 1.04 (0.79 to 1.37)	40 more per 1000 (from 240 fewer to 310	VERY LOW	CRITICAL

										more) ^(e)		
Mortality or heart failure hospitalisation - Ejection fraction < 50%⁶⁷												
1 1 year	observational studies	no serious risk of bias	no serious inconsistency	serious ^(c)	no serious imprecision	none	n/a ^(d)	n/a ^(d)	HR 0.85 (0.77 to 0.94)	160 fewer per 1000 (from 260 fewer to 60 fewer) ^(e)	VERY LOW	CRITICAL
Heart Failure hospitalisation¹⁴⁹												
1 2.4 years	observational studies	no serious risk of bias	no serious inconsistency	serious ^(b)	no serious imprecision	none	558/1337 (41.7%)	564/1337 (42.2%)	HR 0.97 (0.89 to 1.06)	10 fewer per 1000 (from 36 fewer to 19 more)	VERY LOW	IMPORTANT
All cause hospitalisation¹⁴⁹												
1 2.4 years	observational studies	no serious risk of bias	no serious inconsistency	serious ^(b)	no serious imprecision	none	1165/1337 (87.1%)	1155/1337 (86.4%)	HR 0.97 (0.89 to 1.05)	8 fewer per 1000 (from 33 fewer to 13 more)	VERY LOW	IMPORTANT

(a) It was unclear how this outcome was rated either by the clinician or the patient. No validated scale was used.

(b) All studies report on prescriptions at discharge only (i.e. early). It is unclear whether or not the group of patients without prescription later receive this medication (i.e. only indirectly addressing the protocol).

(c) Only a composite outcome of mortality or heart failure re-hospitalisation was reported.

(d) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed one of the default MIDs (i.e. ranging from a noticeable effect to no effect). Outcomes were downgraded by two increments if 95% CI of the effect crossed both default MIDs (i.e. ranging from a noticeable improvement to a noticeable harm). Default MIDs are set at RRs / HRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the median control group standard deviation either side of the null line for continuous variables.

(e) The exact numbers of patients who had or had not received beta-blockers at discharge was not clear and therefore absolute numbers could not be derived and a risk difference was calculated.

Table 80: GRADE evidence profile - quality of evidence by and summary of findings. Aldosterone antagonists in hospital versus clearly or possibly started after discharge for acute heart failure

Quality assessment							Summary of Finding				Quality	Importance	
No of studies Follow-up time	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Aldosterone antagonist s in hospital Number of event / Total N (%) or Mean (SD)	Clearly or possibly started after discharge Number of event / Total N (%) or Mean (SD)	Effect				
									Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect or Mean Difference (95% CI)			
All-cause mortality ^{4,10}													
16 months	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	157/1369 (11.5%)	321/1950 (16.5%)	HR 0.74 (0.6 to 0.91)	40 fewer per 1000 (from 14 fewer to 62 fewer)	LOW	CRITICAL	
16 years	observational studies	no serious risk of bias	no serious inconsistency	serious indirectness ^(a)	no serious imprecision	none	613/864 (70.9%)	605/864 (70%)	HR 0.99 (0.91 to 1.08)	4 fewer per 1000 (from 34 fewer to 28 more)	VERY LOW	CRITICAL	
Sudden cardiac mortality ^{4,5}													
16 months	observational studies	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	51/1369 (3.7%)	111/1950 (5.7%)	HR 0.71 (0.51 to 0.99)	16 fewer per 1000 (from 1 fewer to 27 fewer)	LOW	CRITICAL	
Mortality or heart failure hospitalisation - Ejection fraction > 50% ⁶⁷													
1 year	observational studies	no serious risk of bias	no serious inconsistency	serious indirectness ^(a, b)	very serious ^(c)	none	n/a ^(d)	n/a ^(d)	HR 0.97 (0.67 to 1.4)	30 fewer per 1000 (from 400 fewer to 340 more) ^(d)	VERY LOW	CRITICAL	
Mortality or heart failure hospitalisation - Ejection fraction < 50% ⁶⁷													
1 year	observational studies	no serious risk of bias	no serious inconsistency	serious indirectness ^(a, c)	no serious imprecision	none	n/a ^(d)	n/a ^(d)	HR 0.8 (0.66 to 0.97)	220 fewer per 1000 (from 420 fewer to 30 fewer) ^(d)	VERY LOW	CRITICAL	
Mortality or heart failure hospitalisation ^{4,5}													

1 16 months	observational studies	no serious risk of bias	no serious inconsistency	serious indirectness (c)	no serious imprecision	none	329/1369 (24%)	573/1950 (29.4%)	HR 0.82 (0.71 to 0.95)	200 fewer per 1000 (from 340 fewer to 50 fewer)	VERY LOW	CRITICAL
-------------------	--------------------------	----------------------------------	-----------------------------	--------------------------------	---------------------------	------	-------------------	---------------------	---------------------------	---	----------	----------

- (a) It is unclear whether the group without a prescription for aldosterone antagonists at discharge received this medication later or not, it therefore does not fully match the protocol and is such, considered to be indirect.
- (b) Only a composite outcome of mortality or heart failure re-hospitalisation was reported since this composite was not specified in the review protocol for this review it is considered to be indirect.
- (c) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed one of the default MIDs (i.e. ranging from a noticeable effect to no effect). Outcomes were downgraded by two increments if 95% CI of the effect crossed both default MIDs (i.e. ranging from a noticeable improvement to a noticeable harm). Default MIDs are set at RRs / HRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the median control group standard deviation either side of the null line for continuous variables.
- (d) The exact numbers of patients who had or had not received aldosterone antagonists at discharge was not clear and therefore absolute numbers could not be derived. A risk difference was therefore calculated from the hazard ratio.

9.2.2 Economic evidence

9.2.2.1 ACE inhibitors

Published literature

No relevant economic evaluations were identified.

See also the study selection flow chart in Appendix E.

New cost-effectiveness analysis

New analysis was not prioritised for this area.

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs were provided to aid consideration of cost-effectiveness. These are given in Appendix N.

Economic considerations

The difference in up-front costs between in-hospital and community-based commencement of ACE inhibitors is very low (the cost of a 28-day course of Captopril is £2.91) so their acquisition can be excluded in the consideration of cost-effectiveness. The downstream costs from any long term gains in morbidity and mortality, such as the cost of heart failure acute admissions, are the primary economic considerations.

9.2.2.2 Aldosterone antagonists

Published literature

No relevant economic evaluations were identified which compared the commencement of aldosterone antagonists in the pre-discharge setting after stabilisation with commencement post-discharge setting, for patients not already on an aldosterone antagonist.

See also the study selection flow chart in Appendix E.

New analysis was not prioritised for this area.

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs were provided to aid consideration of cost-effectiveness. These are given in Appendix N.

Economic considerations

The difference in up-front costs between in-hospital and community-based commencement of aldosterone antagonists is low (the cost of a 28-day courses of spironolactone and eplerenone are £1.98 and £42.72 respectively) so their acquisition can be given a low weight in the consideration of cost-effectiveness of best time to commence. The downstream costs from any long term gains in morbidity and mortality, such as the cost of heart failure acute admissions, are the primary economic considerations.

9.2.3 Evidence statements

9.2.3.1 Clinical – ACE inhibitor

Mortality

Very low quality evidence from one observational audit report (N=29,882) showed that ACE inhibitor prescription at the 30 days was associated with a lower rate of mortality and longer length of survival compared to discharge without an ACE inhibitor prescription. At 2.4-year follow-up, low quality evidence from another observational study with 1,337 pairs of patients matched on many baseline characteristics indicated no clear difference in rates of mortality and length of survival between patients who were and were not prescribed an ACE inhibitor at discharge. At 4-year follow-up, very low quality evidence from one observational audit report (N=93,953) showed ACE-inhibitor prescriptions at discharge was associated with a lower rate of mortality.

Mortality or heart failure hospitalisation

Very low quality evidence from one observational study (N=1,026 with ejection fraction >50% and N=1,898 <50% ejection fraction) showed that ACE inhibitor medication at discharge was associated with a reduction in mortality or rehospitalisation only in patients with an ejection fraction <50% whereas the rates of mortality were similar for people with an ejection fraction >50% regardless of ACE-inhibitor prescription at discharge.

Re-hospitalisation

After 2.4 years follow-up low quality evidence from one observational study with 1,337 pairs of matched patients indicated no clear difference in rates of heart failure or all-cause hospitalisation and length of time out of hospital between patients with or without an ACE inhibitor prescription at discharge.

9.2.3.2 Economic – ACE inhibitor

No relevant economic evaluations were identified.

9.2.3.3 Clinical – aldosterone antagonists

Mortality

Low quality evidence from a post hoc analysis of a randomised controlled trial (RCT) (N=3,319) showed that eplerenone when administered earlier led to a lower rate of all-cause and sudden cardiac mortality and longer length of survival compared to eplerenone commenced 7 days later (when followed up for an average of 16 months). Very low quality evidence from an observational study (N=864 pairs of patients with new-onset heart failure matched on 116 baseline characteristics) at 6-year follow-up showed similar rates of mortality for those people who received aldosterone antagonists at discharge from hospital and those without a discharge prescription.

Composite outcome: mortality or heart failure hospitalisation

Very low quality evidence from one observational study (N=1,026 with ejection fraction >50% and N=1,898 <50% ejection fraction) showed that aldosterone antagonist medication at discharge was associated with a reduction in mortality or heart failure rehospitalisation only in patients with an ejection fraction <50%, whereas the rates of mortality were similar for people with an ejection fraction >50% regardless of aldosterone antagonist prescription at discharge. Low quality evidence from one post hoc analysis of an RCT (N=3,319) showed that eplerenone when administered earlier led to a lower rate of mortality or re-hospitalisation for cardiovascular events compared to eplerenone commenced 7 days later (when followed up for an average 16 months).

9.2.3.4 Economic - aldosterone antagonists

No relevant economic evaluations were identified.

9.2.4 Recommendations and link to evidence

Recommendations	30. Offer an angiotensin-converting enzyme inhibitor (or angiotensin receptor blocker if there are intolerable side effects) and an aldosterone antagonist during hospital admission to people with acute heart failure and reduced left ventricular ejection fraction. If the angiotensin-converting enzyme inhibitor (or angiotensin receptor blocker) is not tolerated an aldosterone antagonist should still be offered.
Relative values of different outcomes	The GDG gave most weight to mortality, but also considered a composite outcome of mortality or heart failure hospitalisation and all-cause hospitalisation. The included studies did not provide any data on the other outcomes that had been listed in the protocol.
Trade-off between clinical benefits and harms	Earlier initiation of either medication was associated with lower mortality. Earlier initiation of ACE inhibitor was also associated with improvement in the combined end-point of mortality and hospitalisation due to heart failure. One registry study found that whether or not ACE inhibitor was prescribed at discharge was not associated with long-term risk of rehospitalisation rates (at 2.4 year follow-up). Aldosterone antagonist prescription within 7 days was associated with lower rates of all-cause and sudden cardiac mortality at 16 months as well as the combined end-point of mortality or heart failure rehospitalisations at 16 months compared to aldosterone antagonist initiation after 7 days.
Economic considerations	No relevant economic evidence was identified. Both ACE inhibitors and aldosterone antagonists are established as cost-effective versus placebo in patients with chronic heart failure. Given the association of earlier ACE inhibitor/aldosterone antagonist initiation with improved survival in any patient with LVSD (acute or chronic), and the low unit cost of these drugs, earlier initiation (E.g. in hospital) is unlikely to reduce their cost-effectiveness.
Quality of evidence	The quality of evidence was rated as low to very low using the GRADE criteria. The post-hoc analysis of a randomised controlled trial of earlier versus later initiation of eplerenone was in a post-myocardial infarction population, and the other evidence was observational.
Other considerations	The GDG considered that it was possible to draw conclusions in relation to the question of the optimal timing for these drugs from two systematic reviews that had not met the protocol criteria. The ACE inhibitor MI Collaborative Group ³ found that ACE inhibitor treatment initiated within 36 hours led to improved survival in patients with myocardial infarction complicated by heart failure. Earlier administration within the 36 hours was not associated with any further improvements in survival. Flather and colleagues ⁷⁷ provided data which allowed an indirect comparison of outcome of trials where ACEi had been initiated early (3 to 16 days) and where ACEi had been initiated later (greater than one month). Mortality reductions were similar in both groups of trials. Given that neither systematic review found evidence of harm associated with earlier administration, the GDG concluded that it was safe to prescribe ACE inhibitors early. The GDG recognised that if the drugs were not started in hospital, there was a risk that they would not be started at all.

10 Valvular surgery and percutaneous intervention

10.1 Aortic stenosis

Aortic stenosis is an increasing cause of heart failure, particularly with changes in the epidemiological circumstances posed by an ageing population. Aortic stenosis can lead to acute heart failure, which may initially respond to diuretic therapy, but carries a particularly poor prognosis unless the obstruction to systemic flow is relieved. The relief of aortic stenosis is classically surgical by aortic valve replacement. This operation is associated with significant prognostic benefit. However, the increasing age of patients at presentation and the presence of co-morbidities pose increasing risk to the patient from surgical intervention. This can make surgery an unacceptable option in some cases. For these patients there is a recently developed percutaneous option: transcatheter aortic valve implantation (TAVI). It is important to consider the indications and the timing of the interventions available for patients with acute heart failure and aortic stenosis.

Review question: For people with aortic stenosis are percutaneous or surgical valvular interventions more clinically or cost effective compared to best medical therapy or each other?

For full details see review protocol in Appendix C.

Table 81: PICO characteristics of review question

Population	Adults with aortic stenosis
Intervention/s	<ul style="list-style-type: none"> • Percutaneous repair • Surgical treatment
Comparison/s	or medical management
Outcomes	Mortality Major cardiovascular events Dyspnoea Echocardiographic Criteria: Ejection Fraction Length of index hospital stay and re-admission rates including critical care units Quality of life Adverse events (perioperative vascular AEs)
Study design	Randomised controlled trials Systematic reviews

10.1.1 Clinical evidence

We searched for randomised controlled studies (RCT) and systematic reviews investigating the clinical effectiveness of percutaneous or surgical treatments compared to each other or medical therapy. Two trials (Smith et al, 2011 PARTNER trial²⁰⁵ / Leon et al, 2010 PARTNER B¹²⁰ and Nielsen et al, 2012 STACCATO trial¹⁵⁹) were included in the review. These trials compared percutaneous (transcatheter aortic valve replacement – using the SAPIEN heart valve system Edwards Lifesciences) with surgical management. One of the trials (Leon et al, 2010 PARTNER B cohort¹²⁰) also included a comparison between percutaneous and medical therapy. The PARTNER trial included two different cohorts: whilst cohort A consisted of patients at high risk but operable, cohort B consisted of patients who could not undergo surgery. Outcomes of the two PARTNER cohorts were reported at different time points, in nine publications, which were consulted in this review as well as the accompanying supplementary information (Hancock-Howard et al, 2013⁹⁸, Kodali et al, 2012¹¹⁴, Leon et al, 2010¹²⁰, Makkar et al, 2012¹³¹, Miller et al, 2012^{143,144}, Reynolds et al, 2012a^{186,187}, Reynolds et al, 2011¹⁸⁷,

Nielsen et al, 2012 STACCATO trial ¹⁵⁹	Transcatheter aortic valve implantation Or Surgical aortic valve replacement	70 patients with significant valvular aortic stenosis (valve area <1 cm ² ; age initially ≥70 later ≥75 yrs; condition accessible both by surgical or transcatheter treatment; expected survival >1 year following successful treatment.	Trial terminated early
---	--	---	------------------------

Reynolds et al, 2012^{187,188} and Smith et al, 2011²⁰⁵). No studies were found that compared surgical with medical therapy. Study characteristics are summarised in table 2. Evidence from these are summarised in the clinical GRADE evidence profiles below (table 3 for percutaneous compared to medical treatment and table 4 for percutaneous compared to surgical intervention). See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and lists of excluded studies in Appendix K.

Summary of included studies

Table 82: Summary of studies included in the review

Study	Intervention / Comparison	Population	Comments
Leon et al, 2011 PARTNER B ¹²⁰ / Smith et al, 2011 PARTNER trial ²⁰⁵	Transcatheter aortic valve replacement with a balloon expandable bovine pericardial valve (either transfemoral or a transapical approach using the SAPIEN heart-valve system (Edwards Lifesciences) Or Surgical replacement Or Standard therapy (including balloon aortic valvuloplasty)	PARTNER cohort A: 699 high risk patients with severe aortic stenosis. Severe aortic stenosis was defined as an aortic-valve area of <0.8 cm ² plus either a mean valve gradient of ≥40 mmHg or a peak velocity of ≥4.0 m/s. Age mean (sd) yrs: percutaneous 83.6 (6.8); surgery 84.5 (6.4) PARTNER cohort B: 358 patients with severe aortic stenosis whom surgeons considered not to be suitable candidates for surgery. Age mean (sd) yrs: percutaneous 83.1 (8.6); medical care 83.2 (8.3) For detailed inclusions / exclusions please see the Appendix G/K.	Trial sponsored by industry. There were a multitude of exclusion criteria which might make it questionable whether this is a representative population. Patients do not necessarily have acute heart failure. Some baseline differences were noted.

Table 83: GRADE profile for percutaneous vs. medical management of aortic stenosis (one trial – PARTNER B cohort^{98,120,131,189})

Quality assessment							Summary of Findings				Quality	Importance
							TAVR	Surgery	Effect			
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)		
All-cause mortality - hazard ratio ¹³¹												
12 years	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	77/179 (43%)	117/179 (65.4%)	HR 0.56 (0.43 to 0.61)	206 fewer per 1000 (from 177 fewer to 288 fewer)	MODERATE	CRITICAL
Mortality from cardiac causes - Hazard ratio ¹³¹												
12 years	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	50/179 (27.9%)	100/179 (55.9%)	HR 0.44 (0.32 to 0.61)	256 fewer per 1000 (from 166 fewer to 328 fewer)	MODERATE	CRITICAL
Minor stroke ^{119,120}												
130 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	3/179 (1.7%)	1/179 (0.56%)	RR 3 (0.32 to 28.57)	11 more per 1000 (from 4 fewer to 154 more)	VERY LOW	IMPORTANT
112 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	4/179 (2.2%)	1/179 (0.56%)	RR 4 (0.45 to 35.44)	17 more per 1000 (from 3 fewer to 192 more)		
Major stroke ^{119,120}												
230	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	9/179 (5%)	2/179 (1.1%)	RR 4.5 (0.99 to	39 more per 1000 (from 0 fewer to	LOW	CRITICAL

Quality assessment							Summary of Findings					
							TAVR	Surgery	Effect			
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
days			y	s					20.54)	218 more)		
1 12 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious(b)	none	14/179 (7.8%)	7/179 (3.9%)	RR 2 (0.83 to 4.84)	39 more per 1000 (from 7 fewer to 150 more)		
1 Major and minor stroke s at 24 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	22/179 (12.3%)	8/179 (4.5%)	RR 2.75 (1.26 to 6.01)	78 more per 1000 (from 12 more to 224 more)	MODERATE	
Transient ischemic attack ^{119,120}												
1 30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/179 (0%)	0/179 (0%)	not pooled	not pooled	MODERATE	CRITICAL
1 12 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	1/179 (0.56%)	0/179 (0%)	Peto OR 7.39 (0.15 to 372.38)	10 more per 1000 (from 10 fewer to 20 more) ^(c)	VERY LOW	
Myocardial infarction ^{119,120}												
1	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/179 (0%)	0/179 (0%)	not pooled	not pooled	MODERATE	CRITICAL

Quality assessment							Summary of Findings					
							TAVR	Surgery	Effect			
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
30 days			y	s	n							
1 12 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	1/179 (0.56%)	1/179 (0.56%)	RR 1.00 (0.06 to 15.86)	0 fewer per 1000 (from 5 fewer to 83 more)	VERY LOW	
1 24 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	2/179 (1.1%)	2/179 (1.1%)	RR 1 (0.14 to 7.02)	0 fewer per 1000 (from 10 fewer to 67 more)		
Major vascular complication ^{119,120}												
1 30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	29/179 (16.2%)	2/179 (1.1%)	RR 14.5 (3.51 to 59.86)	151 more per 1000 (from 28 more to 658 more)	MODERATE	CRITICAL
2 12 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	30/179 (16.8%)	4/179 (2.2%)	RR 7.5 (2.7 to 20.85)	145 more per 1000 (from 38 more to 444 more)		
Other adverse events - Renal replacement therapy ^{119,131}												
1 30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	2/179 (1.1%)	3/179 (1.7%)	RR 0.67 (0.11 to 3.94)	6 fewer per 1000 (from 15 fewer to 49 more)	VERY LOW	IMPORTANT
1	RCT	serious ^(a)	no serious	no serious	very	none	5/179	9/179	RR 0.56	22 fewer per		

Quality assessment							Summary of Findings					
							TAVR	Surgery	Effect			
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
24 mths		^{a)}	inconsistency	indirectness	serious ^(b)		(2.8%)	(5%)	(0.19 to 1.63)	1000 (from 41 fewer to 32 more)		
Other adverse events - Major bleeding ^{119,131}												
2-30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	30/179 (16.8%)	7/179 (3.9%)	RR 4.29 (1.93 to 9.5)	129 more per 1000 (from 36 more to 332 more)	MODERATE	IMPORTANT
1-24 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	48/179 (26.8%)	25/179 (14%)	RR 1.92 (1.24 to 2.97)	128 more per 1000 (from 34 more to 275 more)	LOW	
Other adverse events - Endocarditis ^{119,131}												
1-30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/179 (0%)	0/179 (0%)	not pooled	not pooled	MODERATE	IMPORTANT
1-24 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	3/179 (1.7%)	1/179 (0.56%)	RR 3 (0.32 to 28.57)	11 more per 1000 (from 4 fewer to 154 more)	VERY LOW	
Other adverse events - New-onset atrial fibrillation ^{119,131}												
1-30	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	1/179 (0.56%)	2/179 (1.1%)	RR 0.5 (0.05 to 5.46)	6 fewer per 1000 (from 11 fewer to 50 more)	VERY LOW	IMPORTANT

Quality assessment							Summary of Findings					
							TAVR	Surgery	Effect			
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
12 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	1/179 (0.56%)	3/179 (1.7%)	RR 0.33 (0.04 to 3.17)	11 fewer per 1000 (from 16 fewer to 36 more)		
Other adverse events - New pacemaker ^{119,131}												
130 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	6/179 (3.4%)	9/179 (5%)	RR 0.67 (0.24 to 1.83)	17 fewer per 1000 (from 38 fewer to 42 more)	VERY LOW	IMPORTANT
124 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	10/179 (5.6%)	14/179 (7.8%)	RR 0.71 (0.33 to 1.57)	23 fewer per 1000 (from 52 fewer to 45 more)		
Rehospitalisation (2 years) – Hazard Ratio ¹³¹												
1	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	53/179 (29.6%)	95/179 (53.1%)	HR 0.41 (0.30 to 0.56)	264 fewer per 1000 (from 185 fewer to 328 fewer)	MODERATE	IMPORTANT
Quality of Life SF-12 - Physical (Better indicated by high values) – adjusted mean differences from growth curve analysis ¹⁸⁷												
11 mths	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	34.6 (10.3)	30.2 (7.3)	-	MD 4.5 higher (2.5 to 6.5 higher)	VERY LOW	IMPORTANT
112	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	34.9 (11.1)	29.7 (8.5)	-	MD 5.7 higher (2.8 to 8.6 higher)		

Quality assessment							Summary of Findings					
							TAVR	Surgery	Effect			
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
mths		^{a)}	y	s								
Quality of Life SF-12 - Mental (Better indicated by higher values)- adjusted mean differences from growth curve analysis ¹⁸⁷												
1 1 mths	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	47.9 (11.0)	48.5 (10.9)	-	MD 0.6 higher (1.6 lower to 2.8 higher)	LOW	IMPORTANCE
1 12 mths	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	53.3 (10.0)	46.6 (11.7)	-	MD 6.4 higher (3.5 to 9.3 higher)	LOW	

^(a)Outcomes were downgraded by one increment if the weighted average number of serious methodological limitations across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality. Even though due to the differences in procedures participant blinding was impossible this was not considered to be a major limitation for adverse events. However for the quality of life outcomes it was considered to be a more serious risk of bias since this is a subjective rating.

^(b)Outcomes were downgraded by one increment if the upper or lower 95% CI crossed one of the default MIDs (i.e. ranging from a noticeable effect to no effect). Outcomes were downgraded by two increments if 95% CI of the effect crossed both default MIDs (i.e. ranging from a noticeable improvement to a noticeable harm). Default MIDs are set at RRs / HRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the median control group standard deviation either side of the null line for continuous variables.

^(c)Due to zero number of events in either of the control arms the Peto OR was used (which is more robust in cases of small event rates) a risk difference was then calculated.

Table 84: GRADE profile for percutaneous vs. surgical management of aortic stenosis (two trials – PARTNER^{114,144,186,188,205} and STACCATO¹⁵⁹)

Quality assessment							Summary of Findings					
							TAVR	Surgery	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
All-cause mortality - hazard ratio ¹¹⁴												
13 years	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	116/348 (33.3%)	114/351 (32.5%)	HR 0.90 (0.71 to 1.14)	27 fewer per 1000 (from 81 fewer to 36 more)	LOW	CRITICAL
Mortality from cardiac causes ²⁰⁵												
230 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	12/382 (3.1%)	10/387 (2.6%) Control risk 1.4%	RR 1.21 (0.54 to 2.7)	3 more per 1000 (from 6 fewer to 24 more)	VERY LOW	CRITICAL
212 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	50/382 (13.1%)	40/387 (10.3%) Control risk 5.7%	RR 1.26 (0.86 to 1.86)	15 more per 1000 (from 8 fewer to 49 more)	LOW	
124 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	67/348 (19.3%)	59/351 (16.8%)	RR 1.15 (0.83 to 1.57)	25 more per 1000 (from 29 fewer to 96 more)	LOW	
Stroke - Hazard ratio ¹¹⁴												
13 years	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	24/348 (6.9%)	14/351 (4%)	HR 1.22 (0.67 to 2.23)	9 more per 1000 (from 13 fewer to 47 more)	VERY LOW	CRITICAL
Minor stroke ²⁰⁵												

Quality assessment							Summary of Findings					
							TAVR	Surgery	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
1 30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	3/348 (0.86%)	1/351 (0.28%)	RR 3.03 (0.32 to 28.95)	6 more per 1000 (from 2 fewer to 80 more)	VERY LOW	IMPORTANT
1 12 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	3/348 (0.86%)	2/351 (0.57%)	RR 1.51 (0.25 to 9.00)	3 more per 1000 (from 4 fewer to 46 more)		
Major stroke ^{159,205}												
2 30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	16/382 (4.2%)	8/387 (2.1%) control risk 2.4%	RR 2.03 (0.88 to 4.69)	25 more per 1000 (from 3 fewer to 89 more)	LOW	CRITICAL
2 12 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	20/382 (5.2%)	9/387 (2.3%) Control risk 2.4%	RR 2.26 (1.04 to 4.89)	30 more per 1000 (from 1 more to 93 more)		
Transient ischemic attack ^{159,205}												
2 30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	4/382 (1%)	1/387 (0.26%) Control risk 0.1%	RR 3.07 (0.49 to 19.33)	2 more per 1000 (from 1 fewer to 18 more)	VERY LOW	CRITICAL
2 12 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	8/382 (2.1%)	4/387 (1%) Control	RR 1.92 (0.62 to 5.95)	6 more per 1000 (from 2 fewer to 30 more)		

Quality assessment							Summary of Findings					
							TAVR	Surgery	Effect			
No of studies							Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)		
Follo w-up time	Desi gn	Risk of bias	Inconsiste ncy	Indirectne ss	Imprecisi on	Other considerations					Quality	Importanc e
								risk 0.6%				
1 24 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	10/348 (2.9%)	5/351 (1.4%)	RR 2.02 (0.7 to 5.84)	15 more per 1000 (from 4 fewer to 69 more)		
Myocardial infarction ^{159,205}												
2 30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	0/382 (0%)	2/387 (0.5%) Control risk0.3%	RR 0.2 (0.01 to 4.19)	2 fewer per 1000 (from 3 fewer to 10 more)	VERY LOW	CRITICAL
2 12 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	1/382 (0.3%)	2/387 (0.5%) Control risk 0.3%	RR 0.5 (0.05 to 5.54)	2 fewer per 1000 (from 3 fewer to 14 more)		
1 24 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	0/348 (0%)	4/351 (1.1%)	RR 0.11 (0.01 to 2.07)	10 fewer per 1000 (from 11 fewer to 12 more)		
Major vascular complication ^{159,205}												
2 30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	45/382 (11.8%)	11/387 (2.8%) Control risk 1.6%	RR 4.01 (2.13 to 7.54)	48 more per 1000 (from 18 more to 105 more)	MODERATE	CRITICAL
2 12	RCT	serious ^(a)	no serious inconsisten	no serious indirectnes	no serious	none	46/382 (12%)	12/387 (3.1%)	RR 3.77 (2.05 to	47 more per 1000 (from 18 more to		

Quality assessment							Summary of Findings					
							TAVR	Surgery	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
mths			cy	s	imprecision			Control risk 1.7%	6.92)	101 more)		
1 24 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	40/348 (11.5%)	13/351 (3.7%)	RR 3.1 (1.69 to 5.7)	78 more per 1000 (from 26 more to 174 more)		
Other adverse events - Acute renal injury ^{159,205}												
2 30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ²	none	11/382 (2.9%)	10/387 (2.6%) Control risk 1.4%	RR 1.11 (0.49 to 2.53)	2 more per 1000 (from 7 fewer to 21 more)	VERY LOW	IMPORTANT
1 24 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ²	none	20/348 (5.7%)	21/351 (6%)	RR 0.96 (0.53 to 1.74)	2 fewer per 1000 (from 28 fewer to 44 more)		
Other adverse events - Major bleeding ^{114,159,205}												
2 30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	33/382 (8.6%)	68/387 (17.6%) Control risk 10.9%	RR 0.49 (0.33 to 0.72)	56 fewer per 1000 (from 31 fewer to 73 fewer)	MODERATE	IMPORTANT
1 24 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	60/348 (17.2%)	95/351 (27.1%)	RR 0.64 (0.48 to 0.85)	97 fewer per 1000 (from 41 fewer to 141 fewer)	LOW	

Quality assessment							Summary of Findings					
							TAVR	Surgery	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
Other adverse events - Endocarditis ²⁰⁵												
1 30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	0/348 (0%)	1/351 (0.28%)	RR 0.34 (0.01 to 8.22)	2 fewer per 1000 (from 3 fewer to 21 more)	VERY LOW	IMPORTANT
1 24 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	4/348 (1.1%)	3/351 (0.85%)	RR 1.34 (0.3 to 5.96)	3 more per 1000 (from 6 fewer to 42 more)		
Other adverse events - New-onset atrial fibrillation ²⁰⁵												
1 30	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	30/348 (8.6%)	56/351 (16%)	RR 0.54 (0.36 to 0.82)	73 fewer per 1000 (from 29 fewer to 102 fewer)	LOW	IMPORTANT
1 12 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	42/348 (12.1%)	60/351 (17.1%)	RR 0.71 (0.49 to 1.02)	50 fewer per 1000 (from 87 fewer to 3 more)		
Other adverse events - New pacemaker ^{159,205}												
2 30 days	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	15/382 (3.9%)	13/387 (3.4%) Control risk 3.1%	RR 1.17 (0.56 to 2.42)	5 more per 1000 (from 14 fewer to 44 more)	VERY LOW	IMPORTANT
1 24 mths	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	very serious ^(b)	none	23/348 (6.6%)	19/351 (5.4%)	RR 1.22 (0.68 to 2.2)	12 more per 1000 (from 17 fewer to 65 more)		

Quality assessment							Summary of Findings					
							TAVR	Surgery	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
Rehospitalisation - Follow-up 24 mths ²⁰⁵												
1	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	74/348 (21.3%)	60/351 (17.1%)	RR 1.24 (0.92 to 1.69)	41 more per 1000 (from 14 fewer to 118 more)	LOW	IMPORTANT
Length of index hospital stay (Better indicated by lower values) ¹⁵⁹												
1	RCT	serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	8.8 (6.7) n=34	7.6 (2.4) n=36	-	MD 1.2 higher (1.18 lower to 3.58 higher)	LOW	IMPORTANT
EQ-5D - Transfemoral (Better indicated by high values) ^{187,188} – change from baseline scores												
11 mths	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	0.08 (0.28) n=192	0.02 (0.25) n=154	-	MD 0.06 higher (0 to 0.12 higher)	LOW	IMPORTANT
112 mths	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	no serious imprecision	none	0.09 (0.26) n=160	0.08 (0.23) n=129	-	MD 0.01 higher (0.05 lower to 0.07 higher)		
EQ-5D - Transapical (Better indicated by higher values) ^{187,188} – change from baseline scores												
11 mths	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	-0.02 (0.26) n=74	0.01 (0.19) n=58	-	MD 0.03 lower (0.11 lower to 0.05 higher)	VERY LOW	IMPORTANT
112	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	no serious	none	0.06 (0.20)	0.05 (0.26)	-	MD 0.01 higher (0.07 lower to 0.09	LOW	

Quality assessment							Summary of Findings					
							TAVR	Surgery	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of event / Total N (%) or Mean (SD)	Number of event / Total N (%) or Mean (SD)	Hazard Ratio, Relative Risk, Peto Odds Ratio, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)	Quality	Importance
mths		a)	cy	s	imprecision		n=61	n=54		higher)		
Quality of Life - SF 36 Mental composite (Better indicated by higher values) ¹⁵⁹ – final scores												
13 mths	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	53 (14) n=27	50 (17) n=32	-	MD 3.00 higher (1.91 lower to 10.91 higher)	VERY LOW	IMPORTANT
Quality of Life - SF 36 Physical composite (Better indicated by higher values) ¹⁵⁹ – final scores												
13 mths	RCT	very serious ^(a)	no serious inconsistency	no serious indirectness	serious ^(b)	none	42 (14) n=27	43 (15) n=32	-	MD 1.00 lower (8.41 lower to 6.41 higher)	VERY LOW	IMPORTANT

^(a) Outcomes were downgraded by one increment if the weighted average number of serious methodological limitations across studies was one, and downgraded by two increments if the weighted average number of serious methodological limitation across studies were two or more. Methodological limitations comprised one or more of the following: unclear allocation concealment, the lack of blinding, inadequate allowance for drop-outs in the analysis, selective outcome reporting or unadjusted baseline inequality. Even though due to the differences in procedures participant blinding was impossible this was not considered to be a major limitation for adverse events. However for the quality of life outcomes it was considered to be a more serious risk of bias since this is a subjective rating.

^(b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed one of the default MIDs (i.e. ranging from a noticeable effect to no effect). Outcomes were downgraded by two increments if 95% CI of the effect crossed both default MIDs (i.e. ranging from a noticeable improvement to a noticeable harm). Default MIDs are set at RRs / HRs of 0.75 and 1.25 for dichotomous outcomes, and at 0.5 of the median control group standard deviation either side of the null line for continuous variables.

10.1.2 Economic evidence

Published literature

Four studies were included that compared transcatheter aortic valve implantation (TAVI) with either surgical aortic valve replacement (SAVR) or medical management (MM) in patients with aortic stenosis. These are summarised in the economic evidence profile below (**Table 85**) and the economic evidence tables in Appendix H.

Four economic evaluations relating to this review question were identified but were excluded due to the availability of more applicable evidence. These are summarised in Appendix L, with reasons for exclusion given.

See also the study selection flow chart in Appendix E.

Patients at operative high risk but eligible for surgery

Table 85: TAVI versus Surgery – Economic study characteristics

Study	Applicability	Limitations	Other comments
Fairbairn 2013 ⁶⁹ (UK)	Directly Applicable ^(a)	Potentially Serious Limitations ^(b)	Cost-utility analysis 10-year time horizon Use of PARTNER A data

(a) Recent evaluations and costs and effects are measured in a UK NHS context

(b) Main limitations: The increase in QALYs for TAVI versus SAVR in this evaluation is not supported by the findings of the underlying outcomes study, PARTNER

Table 86: TAVI versus Surgery – Economic summary of findings

Study	Incremental cost	Incremental effects	ICER	Uncertainty
Fairbairn 2013 ⁶⁹ (UK)	TAVI = £1,350 decrease	TAVI = 0.063 QALYs gained	TAVI dominates SAVR	Probabilistic sensitivity analysis estimated a 65% probability of TAVI being cost effective compared to SAVR. Univariate deterministic analysis showed that dominance was robust

Abbreviations: TAVI = transcatheter aortic valve implantation; SAVR = surgical aortic valve replacement; MM = medical management; CI = 95% confidence interval; ICER = incremental cost-effectiveness ratio; QALYs = quality-adjusted life years; WTP = willingness to pay

Patients at high risk and ineligible for surgery/inoperable

Table 87: TAVI versus Medical Management – Economic study characteristics

Study	Applicability	Limitations	Other comments
Orlando 2013 ¹⁶⁴ (UK)	Directly Applicable ^(a)	Potentially Serious Limitations ^(b)	Cost-utility analysis 25-year time horizon Use of 2-year PARTNER B data
Watt 2012 ^{224, 225} (UK)	Directly Applicable ^(a)	Potentially Serious Limitations ^(b)	Cost-utility analysis 10-year time horizon Use of 2-year PARTNER B data

Study	Applicability	Limitations	Other comments
Murphy 2013 ^{150,151} (UK)	Directly Applicable ^(a)	Potentially Serious Limitations ^(b)	Cost-utility analysis Lifetime time horizon Uses 1-year mortality data from PARTNER B and Bleiziffer (Germany)

(a) Recent evaluations: costs and effects are measured in a UK NHS context

(b) Main limitations: unbalanced patient characteristics in the PARTNER trial; uncertainty surrounding the mortality benefit applied into the long-term

Table 88: TAVI versus Medical Management – Economic summary of findings

Study	Incremental cost	Incremental effects	ICER	Uncertainty
Orlando 2013 ¹⁶⁴ (UK)	TAVI = £24,147 increase	TAVI = 1.87 QALYs gained	= £12,900 per QALY gained	Not reported
Watt 2012 ^{224,225} (UK)	TAVI = £25,200 increase	TAVI = 1.56 QALYs gained	= £16,200 per QALY gained	The ICER is sensitive to the time horizon adopted (10-years) 100% probability of TAVI being more cost-effective than MM at the £20,000/QALY WTP threshold
Murphy 2013 ^{150,151} (UK)	TAVI = £8,415 increase	TAVI = 0.488 QALYs gained	= £35,956 per QALY gained	ICER 95% CI: £24,768, £65,103 18% probability of TAVI being more cost-effective than MM at the £30,000/QALY WTP threshold (66% at £40,000)

Abbreviations: TAVI = transcatheter aortic valve implantation; SAVR = surgical aortic valve replacement; MM = medical management; CI = 95% confidence interval; ICER = incremental cost-effectiveness ratio; QALYs = quality-adjusted life years; WTP = willingness to pay.

New cost-effectiveness analysis

New analysis was not prioritised for this area.

10.1.3 Evidence statements

Clinical

Percutaneous versus medical treatment

Mortality

- All-cause mortality – Moderate quality evidence from one randomised control trial (N=258) showed that there was a lower rate of mortality and longer length of survival associated with percutaneous treatment compared to medical treatment over a 2-year follow-up period.
- Mortality from cardiac causes – Moderate quality evidence from one randomised control trial (N=258) showed that there was a lower rate of mortality and longer length of survival associated with percutaneous treatment compared to medical treatment over a 2-year follow-up period.

Stroke

- Minor stroke – Very low quality evidence from one randomised controlled trial (N=258) showed overall a low rate of minor strokes in either group. However, the rate was higher in the

percutaneous treatment group compared to medical care both at 30-day and 1-year follow-up. There is very large uncertainty around this effect and it is unclear whether this constitutes a clear clinical harm associated with percutaneous treatment.

- Major stroke - Very low quality evidence from one randomised controlled trial (N=258) suggests a higher rate of major strokes in the percutaneous treatment group compared to medical care both at 30-days and 1-year follow-up. This may indicate a clinical risk associated with percutaneous treatment but with some uncertainty.
- All strokes at 2-year follow-up – Moderate quality evidence from one randomised controlled trial (N=258) showed a higher rate of strokes associated with percutaneous treatment. This indicates a clinical risk associated with percutaneous treatment over medical treatment.

Transient ischemic attack

The overall short and long term rate of transient ischemic attacks was very low in one trial comprising 258 patients (one reported case over the 1-year period in the percutaneous group) in either the percutaneous or medical group. This was moderate to very low evidence indicating no clear difference between treatment groups.

Myocardial infarction

The overall short and long term rate of myocardial infarctions was very low in one trial comprising 258 patients (<1% in the 1-year period and 1% after 2-year follow-up) in either the percutaneous or medical group. This was moderate to very low evidence indicating no clear difference between treatment groups.

Major vascular complications

Moderate quality evidence from one randomised controlled trial comprising 258 patients shows that there is a higher rate of major vascular complications in the percutaneous treatment group. This constitutes a clinical risk associated with percutaneous treatment

Quality of life (as measured by SF-12)

Very low quality evidence from one randomised control trial (N=258) showed that patients who had received percutaneous treatment rated their quality of life higher in short and long term (1 month and 1 year) for the physical (1 month and 1 year) and mental (1 year) components. However, the rating of the short-term mental quality of life scale did not show a clear difference between patients who had received percutaneous compared to patients who received medical treatment.

Rehospitalisation

Moderate quality evidence from one randomised controlled trial (N=258) showed a lower rate of rehospitalisation and a longer time out of hospital associated with percutaneous treatment compared to medical treatment.

Other severe adverse events (30 day and 2 year follow-up):

Renal replacement therapy

Very low quality evidence from one randomised controlled trial (N=258) showed similar rates of renal replacement therapy in both the percutaneous and medical treatment groups.

Major bleeding

Low to moderate quality evidence from one randomised controlled trial (N=258) showed that major bleeding is more common in the percutaneous group both long and short term. This indicates a risk associated with percutaneous treatment.

Endocarditis

Evidence from one randomised controlled trial (N=258) showed no endocarditis in either group after 30 days indicating no difference between groups (moderate quality). After one year the rate was higher in the percutaneous group but due to uncertainty around this effect it is unclear whether this indicates a clear risk associated with percutaneous treatment (very low quality).

New onset atrial fibrillation

Very low quality evidence from one randomised controlled trial (N=258) showed higher rates of new onset atrial fibrillation in the medical treatment group both short and long term. However it is unclear whether this indicates a clear clinical benefit associated with percutaneous treatment.

New pacemaker

Very low quality evidence from one randomised controlled trial (N=258) controlled trial showed similar rates of new pacemaker placement at short and longer term follow-up.

Percutaneous versus surgical treatment

Mortality

- All-cause mortality – Low quality evidence from one randomised controlled (N=699) trial shows similar rates of mortality and length of survival in both the percutaneous and the surgery groups.
- Mortality from cardiac causes – Low to very low quality evidence from two randomised controlled trials comprising 769 patients (at 30 day and 1 year follow-up) showed similar rates of mortality from cardiac causes. One of the studies (N=699) also provided low quality evidence for mortality from cardiac causes at 2-year follow-up and also showed no clear difference in rates of mortality in the percutaneous and surgery groups.

Stroke

- All strokes – Very low quality evidence from one study at two-year (N=699) follow up showed there were similar rates of all strokes and a similar length of time without a stroke in the percutaneous and surgery groups.
- Minor stroke – Very low quality evidence from one randomised controlled trial (N=699) showed similar rates of minor strokes both at 30-day and 1-year follow-up in the percutaneous and surgery groups.
- Major stroke – Low quality evidence from two randomised controlled trials (N=769) showed a higher rate of major strokes associated with percutaneous treatment which constitutes a clinical risk associated with percutaneous treatment over surgical treatment.

Transient ischemic attack

Very low quality evidence from two randomised controlled studies comprising 769 patients (at 30 days and 1 year) showed higher rates of transient ischemic attacks associated with the percutaneous group, but the difference was small and it is unclear whether this constitutes a clear risk associated with percutaneous treatment. Very low quality of one randomised controlled trial (N=699) also

provided 2-year follow-up results which showed a higher rate associated with percutaneous treatment which may indicate a clinical risk.

Myocardial infarction

Very low quality evidence from two randomised controlled trials (N=769) at 30-day and 1-year follow up showed similar rates of myocardial infarctions in the two treatment groups which was also the case at 2-year follow-up (evidence only from one of the two studies – N=699).

Major vascular complications

Very low quality evidence from two randomised controlled trials (N=769) at 30-day and 1-year follow up showed a higher rate of major cardiovascular events in the percutaneous group which was also the case at 2-year follow-up (evidence only from one of the two studies). This indicates a risk associated with percutaneous treatment.

Quality of life (as measured by EQ-5D and SF-36)

Low to very low quality evidence from one randomised controlled trial showed no clear difference in quality of life between treatment groups both at 1 month and 1 year after treatment (N=699). The second randomised controlled trial provided very low quality evidence using a different quality of life measurement and also indicated no differences between the treatment groups both in a physical and mental component of quality of life (N=70).

Rehospitalisation

Low quality of one randomised controlled trial showed that there was a lower rate of rehospitalisation associated with the surgery group (N=699). However, the results are a bit uncertain and it is unclear whether this constitutes a clear clinical benefit associated with surgery.

Length of hospital stay

Low quality evidence from one randomised controlled trial (N=70) showed a 1 day different length of index stay in the percutaneous group, but there was uncertainty around this effect and it is unclear whether this is a clear clinical benefit in favour of surgical treatment.

Other severe adverse events (30-day and 2-year follow-up):

Renal replacement therapy

Very low quality evidence from two randomised controlled trials (N=769) at 30-day follow-up and from one of them at 2-year follow-up (N=699) indicated similar rates of renal injuries in the two treatment groups.

Major bleeding

Moderate quality evidence from two randomised controlled trials (N=769) at 30-day follow-up and low quality evidence from one of the trials (N=699) at 2-year follow-up showed higher rates of major bleeding associated with the surgery group. This indicates a clinical risk associated with surgery.

Endocarditis

Very low quality evidence from one randomised controlled trial (N=699) at 30-day and 2-year follow-up indicates similar low rates of endocarditis in either treatment group.

New onset atrial fibrillation

Low quality evidence from one randomised controlled trial at (N=699) 30-day and 1-year follow-up showed higher rates of new onset atrial fibrillation associated with the surgery group. This seems to indicate a clinical risk associated with surgery over percutaneous treatment.

New pacemaker

Very low quality evidence from two randomised controlled trials (N=769) at 30-day follow up and from one of the trials (N=699) at 2-year follow up indicated similar rates of new pacemaker placements in the two treatment groups.

Economic

Patients at high risk but eligible for surgery

One study found that TAVI was cost effective compared to surgery in patients with acute heart failure with high operative risk. This study was assessed as directly applicable with potentially serious limitations.

Patients ineligible for surgery

Two cost-utility analyses found that TAVI was cost-effective compared to medical management for treating aortic stenosis in patients ineligible for surgery (the ICERs were £16,200 and £12,900 per QALY gained). These studies were assessed as directly applicable with potentially serious limitations.

One study found that TAVI was not cost-effective compared to medical management (ICER: £35,956 per QALY gained). This study was assessed as directly applicable with potentially serious limitations.

10.1.4 Recommendations and links to evidence

Recommendations	<p>31. Offer surgical aortic valve replacement to people^h with heart failure due to severe aortic stenosis assessed as suitable for surgery.</p> <p>32. Consider transcatheter aortic valve implantation (TAVI) in selected people^h, with heart failure caused by severe aortic stenosis, who are assessed as unsuitable for surgical aortic valve replacement. Details of all people undergoing TAVI should be entered into the UK Central Cardiac Audit database.</p> <p>33. For guidance on coronary revascularisation see Chronic heart failure (NICE clinical guideline 108).</p>
Relative values of different outcomes	The GDG considered all cause and cardiac mortality to be the most important outcomes for both comparisons (surgery vs percutaneous; percutaneous vs medical therapy). The GDG also looked at other adverse events associated with each treatment, in particular, stroke was a critical outcome in relation to percutaneous intervention, and quality of life. Long term follow-up data on TAVI are not available.
Trade-off between clinical benefits and harms	<p>TAVI versus surgery</p> <p>There was no difference in all cause and cardiac mortality between TAVI and</p>

^h For information about patient selection, see Transcatheter aortic valve implantation for aortic stenosis (NICE interventional procedure guidance 421).

<p>Recommendations</p>	<p>31.Offer surgical aortic valve replacement to people^h with heart failure due to severe aortic stenosis assessed as suitable for surgery.</p> <p>32.Consider transcatheter aortic valve implantation (TAVI) in selected people^h, with heart failure caused by severe aortic stenosis, who are assessed as unsuitable for surgical aortic valve replacement. Details of all people undergoing TAVI should be entered into the UK Central Cardiac Audit database.</p> <p>33.For guidance on coronary revascularisation see Chronic heart failure (NICE clinical guideline 108).</p>
	<p>surgery.</p> <p>Major bleeding and new onset atrial fibrillation were less common in the TAVI group, but were outweighed by lower rates of major stroke and vascular complications in the surgery group.</p> <p>TAVI versus medical care</p> <p>Lower rates of mortality and longer survival associated with TAVI were the key benefit of this treatment. Risks of major stroke were higher in the TAVI group compared to the medical treatment group (at 30 days and 1 year follow-up times). Short and long term physical health-related quality of life as well as short term mental health-related quality of life were rated more highly by people who received TAVI. Readmission rates were higher in the medical care group.</p>
<p>Economic considerations</p>	<p>TAVI versus surgery</p> <p>The GDG judged that TAVI was not cost effective compared to surgery in patients eligible for surgery. One recent UK economic evaluation (Fairbairn2013) found TAVI to be cost-effective compared to surgery but this analysis relied upon the extrapolation of a short follow-up period to 10 years, and the findings of the clinical review did not provide confidence in the evaluation's conclusion. Randomised clinical trials of the most recent technology for TAVI over a longer follow-up are required.</p> <p>TAVI versus medical care</p> <p>The GDG judged TAVI to be cost effective versus medical care based on three UK economic evaluations, two of which found TAVI to be cost effective, though the third did not.</p>
<p>Quality of evidence</p>	<p>The evidence was rated moderate to very low quality. Even though the industry-sponsored PARTNER study is a landmark trial for this type of percutaneous intervention, it has several exclusion criteria which makes generalisation to all people with aortic stenosis difficult. The trial included mostly stable patients, so is of uncertain relevance to acute heart failure patients. The GDG did not extensively discuss the data from the STACCATO trial as this trial stopped earlier than anticipated due to the considerably higher rate of adverse events in TAVI group.</p>
<p>Other considerations</p>	<p>The GDG was aware of NICE Interventional Procedure Guidance 421 on TAVI for aortic stenosis. This guidance considers TAVI to be an option for patients unsuitable for surgery, but for patients who can undergo surgery, TAVI should only be offered in the context of research. The GDG considered this and the presented evidence and made recommendations in line with this guidance.</p> <p>The GDG agreed with the NICE IPG421 that patient selection for TAVI should take into account patient comorbidities and procedural risks and be performed by a multi-disciplinary team including expertise from interventional cardiology, cardiac surgery, cardiac anaesthetics and cardiac imaging. Patients need to satisfy a number of anatomical criteria in order to be eligible for TAVI.</p>

Recommendations	<p>31. Offer surgical aortic valve replacement to people^h with heart failure due to severe aortic stenosis assessed as suitable for surgery.</p> <p>32. Consider transcatheter aortic valve implantation (TAVI) in selected people^h, with heart failure caused by severe aortic stenosis, who are assessed as unsuitable for surgical aortic valve replacement. Details of all people undergoing TAVI should be entered into the UK Central Cardiac Audit database.</p> <p>33. For guidance on coronary revascularisation see Chronic heart failure (NICE clinical guideline 108).</p>
	<p>The GDG recognised the importance of fully informed patient choice with regard to TAVI, so that decisions to undergo the procedure are made in knowledge of the potential risks and benefits.</p> <p>The GDG noted that the existing guidance on coronary revascularisation provided in NICE Chronic Heart Failure guideline (CG108) is applicable to the acute population and agreed that reference should be made to this guideline.</p> <p>The GDG noted that the timing of surgery should be individualised and often would not occur during the index admission.</p>

10.2 Mitral regurgitation

Mitral regurgitation is frequently present in patients with heart failure whether they have heart failure with preserved ejection fraction or heart failure due to left ventricular systolic dysfunction. The mitral regurgitation in these patients is generally a secondary phenomenon due to either the haemodynamic or the structural disturbances caused by these types of heart failure. The mitral regurgitation in the majority of these patients is not severe and the treatment of the cause is frequently sufficient.

However, when the mitral regurgitation is severe, or more importantly when such severe mitral regurgitation is of an acute onset, a different approach to the diagnosis and management becomes necessary. Conventionally, mitral valve intervention would be performed via an 'open heart' surgical approach and is well validated, but the advent of new technology means other options may be available, particularly when the risk of conventional surgery is deemed prohibitive. This review examines the mitral valve surgery and percutaneous mitral valve interventions in people with heart failure.

Review question: For people with heart failure with mitral regurgitation, are surgical valvular or percutaneous interventions more clinically or cost effective compared to best medical therapy or each other?

For full details see review protocol in Appendix C.

Table 89: PICO characteristics of review question

Population	Adults with mitral regurgitation
-------------------	----------------------------------

Population	Adults with mitral regurgitation
Intervention/s	<ul style="list-style-type: none"> • Percutaneous repair • Surgical treatment
Comparison/s	Medical care
Outcomes	Mortality Major cardiovascular events Dyspnoea Echocardiographic criteria: Ejection Fraction Length of index hospital stay and re-admission rates including critical care units Quality of life Adverse events (perioperative vascular events)
Study design	Randomised controlled trial

10.2.1 Clinical evidence

We searched for randomised controlled studies investigating the clinical effectiveness of percutaneous or surgical treatments compared to each other or medical care. One trial (the EVEREST II - Feldman et al, 2011^{71,72} with 4-year follow-up data reported in Mauri et al, 2013^{135,136}) was included in the review. This trial compared percutaneous (MitraClip, Abbott Vascular) with surgical management. No trials were identified for the comparisons, percutaneous versus medical care or surgical versus medical care. Evidence from the trial is summarised in the clinical GRADE evidence profile below (Table 3). See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and lists of excluded studies in Appendix K.

Summary of included studies

Study characteristics are briefly outlined in Table 2.

Table 90: Summary of studies included in the review

Study	Intervention / Comparison	Population	Comments
Feldman et al, 2011 and the 4-year follow-up results reported by Mauri et al, 2013 – EVEREST-II ^{72,135}	<p>Percutaneous repair using the MitraClip device which is a 4-mm-wide cobalt-chromium implant with two arms that are opened and closed with the use of the delivery-system handle and the procedure is performed under general anaesthesia with the use of fluoroscopic and transoesophageal echocardiographic guidance.</p> <p>Randomised to percutaneous treatment and surgery in a 2:1 ratio.</p>	<p>N=279</p> <p>All patients had grade 3+ or 4+ chronic mitral regurgitation. Those who were symptomatic were required to have a left ventricular ejection fraction (LVEF) of more than 25% and a left ventricular end-systolic diameter of 55 mm or less. Asymptomatic patients were required to have at least one of the following: an LVEF of 25 to 60%, a left ventricular end-systolic diameter of 40 mm to 55 mm, new atrial fibrillation, or pulmonary hypertension. (For full inclusion and</p>	<p>There were a multitude of exclusion criteria which might make it questionable whether this is a representative population. Patients do not necessarily have acute heart failure.</p>

Study	Intervention / Comparison	Population	Comments
	Compared to mitral valve surgery	exclusion criteria please see Appendix G/K)	

Table 91: GRADE profile for Percutaneous vs. Surgical management of mitral regurgitation (Feldman et al, 2011 – EVEREST II^{71,72} and four year follow-up reported in Mauri et al, 2013^{135,136})

Quality assessment							Summary of Findings				Quality	Importance
							No. of studies	Design	Risk of bias	Inconsistency		
Relative Risk, Peto OR, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)											
		Follow-up time										
All-cause mortality ^{72,135}												
1 30 days	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	2/180 (1.1%)	2/94 (2.1%)	RR 0.52 (0.07 to 3.65)	10 fewer per 1000 (from 20 fewer to 56 more)	LOW	CRITICAL
1 12 mths	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	11/181 (6.1%)	5/89 (5.6%)	RR 1.08 (0.39 to 3.02)	4 more per 1000 (from 34 fewer to 113 more)		
1 2 yrs	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	20/172 (11.6%)	10/83 (12.1%)	RR 0.97 (0.47 to 1.97)	4 fewer per 1000 (from 64 fewer to 117 more)		
1 4 yrs	randomised trials	no serious risk of	no serious inconsistency	no serious indirectness	very serious ^(a)	none	28/161 (17.4%)	13/73 (17.8%)	RR 0.98 (0.54 to 1.77)	4 fewer per 1000 (from 82 fewer to 137 more)		

Quality assessment							Summary of Findings				Quality	Importance
							No. of studies	Design	Risk of bias	Inconsistency		
Follow-up time	Relative Risk, Peto OR, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)										
		bias						%)		more)		
Re-operation for failed surgical repair/replacement - 30 days ^{71,72}												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	0/180 (0%)	1/94 (1.1%)	RR 0.17 (0.01 to 4.25)	9 fewer per 1000 (from 11 fewer to 35 more)	LOW	CRITICAL
Urgent or emergency cardiovascular surgery - 30 days ^{71,72}												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	4/180 (2.2%)	4/94 (4.3%)	RR 0.52 (0.13 to 2.04)	20 fewer per 1000 (from 37 fewer to 44 more)	LOW	CRITICAL
Surgery for mitral-valve dysfunction - 12 mths ^{72,135}												
1 12 mths	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	37/181 (20.4%)	2/89 (2.2%)	RR 9.1 (2.24 to 36.89)	182 more per 1000 (from 28 more to 807 more)	HIGH	CRITICAL
1 4 yrs	randomised trials	no serious risk of	no serious inconsistency	no serious indirectness	no serious imprecision	none	40/161 (24.8%)	4/73 (5.5)	RR 4.53 (1.68 to	193 more per 1000 (from 37 more to 614		

Quality assessment							Summary of Findings				Quality	Importance
							No. of studies	Design	Risk of bias	Inconsistency		
Follow-up time	Relative Risk, Peto OR, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)										
		bias						%)	12.2)	more)		
Grade 3+ or 4+ mitral regurgitation ^{72,135}												
12 mths	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	38/181 (21%)	20/89 (20.2%)	RR 1.04 (0.63 to 1.71)	8 more per 1000 (from 75 fewer to 144 more)	LOW	IMPORTANT
14 yrs	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	35/161 (21.7%)	18/73 (24.7%)	RR 0.88 (0.54 to 1.45)	30 fewer per 1000 (from 113 fewer to 111 more)		
Any major adverse events ^{71,72}												
130 days	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	15/180 (15%)	45/94 (47.9%)	RR 0.31 (0.21 to 0.47)	330 fewer per 1000 (from 254 fewer to 378 fewer)	HIGH	CRITICAL
Any major adverse events excluding transfusion ^{71,72}												
130	randomised trials	no serious risk of	no serious inconsistency	no serious indirectness	very serious ^(a)	none	9/180 (5%)	9/94 (9.6)	RR 0.52 (0.21 to	46 fewer per 1000 (from 76 fewer to 26	LOW	CRITICAL

Quality assessment							Summary of Findings				Quality	Importance
							Percutaneous Number of event / Total N (%) or Mean (SD)	Surgery Number of event / Total N (%) or Mean (SD)	Effect			
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			Relative Risk, Peto OR, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)		
Follow-up time												
days		bias						%)	1.27)	more)		
Adverse events - Myocardial Infarction ^{71,72}												
1 30 days	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/180 (0%)	0/94 (0%)	not pooled	not pooled	HIGH	CRITICAL
Adverse events - Major stroke ^{71,72}												
1 30 days	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	2/180 (1.1%)	2/94 (2.1%)	Peto OR 0.49 (0.06 to 3.94)	10 fewer per 1000 (from 40 fewer to 20 more) ^(b)	LOW	CRITICAL
Adverse events - Renal failure ^{71,72}												
1 30 days	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	1/180 (0.6%)	0/94 (0%)	Peto OR 4.58 (0.07 to 284.51)	10 more per 1000 (from 10 fewer to 30 more) ^(b)	LOW	CRITICAL
Adverse events - Deep wound infection ^{71,72}												
1	randomised	no	no serious	no serious	no serious	none	0/180	0/94	not pooled	not pooled	HIGH	IMPORT

Quality assessment							Summary of Findings				Quality	Importance
							No. of studies	Design	Risk of bias	Inconsistency		
Follow-up time	Relative Risk, Peto OR, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)										
30 days	ed trials	serious risk of bias	inconsistency	indirectness	imprecision		(0%)	(0%)				ANT
Adverse events - Invasive ventilation >48hrs ^{71,72}												
1 30 days	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/180 (0%)	4/94 (4.2%)	RR 0.06 (0 to 1.06)	40 fewer per 1000 (from 90 fewer to 0 more) ^(b)	HIGH	CRITICAL
Adverse events - Gastrointestinal complications requiring surgery ^{71,72}												
1 30 days	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ^(a)	none	2/180 (1.1%)	0/94 (0%)	Peto OR 4.58 (0.24 to 85.77)	10 more per 1000 (from 10 fewer to 30 more) ^(b)	MODERATE	IMPORTANT
Adverse events - New onset permanent atrial fibrillation ^{71,72}												
1 30 days	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ^(a)	none	2/180 (1.1%)	0/94 (0%)	Peto OR 4.58 (0.24 to 85.77)	10 more per 1000 (from 10 fewer to 30 more) ^(b)	MODERATE	CRITICAL

Quality assessment							Summary of Findings				Quality	Importance
							No. of studies	Design	Risk of bias	Inconsistency		
Relative Risk, Peto OR, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)											
		Follow-up time										
Adverse events - Septicaemia ^{71,72}												
1 30 days	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/180 (0%)	0/94 (0%)	not pooled	not pooled	HIGH	IMPORTANT
Adverse events - Transfusion of >=2 units of blood ^{71,72}												
1 30 days	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	24/180 (13%)	42/94 (44.2%)	RR 0.3 (0.19 to 0.46)	313 fewer per 1000 (from 241 fewer to 362 fewer)	HIGH	IMPORTANT
Quality of Life - SF-36 - Physical Component - 30 days (Better indicated by higher values) ^{71,72} – change from baseline scores												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ^(a)	none	3.1 (9.4) 147	-4.1 (13.3) 64	-	MD 7.2 higher (3.6 to 10.8 higher)	Moderate	CRITICAL
Quality of Life - SF-36 - Mental Component - 30 days (Better indicated by higher values) ^{71,72} – change from baseline scores												
1	randomised trials	no serious	no serious inconsistency	no serious indirectness	serious ^(a)	none	4.4 (11.3) 148	1.8 (13.4)	-	MD 2.6 higher (1.15 lower to	Moderate	CRITICAL

Quality assessment							Summary of Findings					
							No. of studies	Design	Risk of bias	Inconsistency		
Follow-up time												Relative Risk, Peto OR, Mean (SD), (95% CI)
		risk of bias) 64		6.35 higher)		
Quality of Life - SF-36 - Physical Component - 12 mths (Better indicated by higher values) ^{71,72}												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	4.4 (9.8) 132	4.4 (10.4) 60	-	MD 0 higher (3.12 lower to 3.12 higher)	HIGH	CRITICAL
Quality of Life - SF-36 - Mental Component - 12 mths (Better indicated by higher values) ^{71,72}												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	5.7 (9.9) 133	3.8 (10.3) 60	-	MD 1.9 higher (1.2 lower to 5 higher)	HIGH	CRITICAL
Change in ejection fraction % (follow-up 12 months; Better indicated by higher values) ^{71,72}												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious ^(a)	none	-2.8 (7.2) 144	-6.8 (10.1) 66	-	MD 4.00 higher (1.29 to 6.71 higher)	MODERATE	IMPORTANT
Breathlessness - NYHA functional class III or IV ^{72,135}												

Quality assessment							Summary of Findings				Quality	Importance
							No. of studies	Design	Risk of bias	Inconsistency		
Follow-up time	Relative Risk, Peto OR, Mean (SD), (95% CI)	Absolute effect / Mean difference (MD)										
12 months	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	4/181 (2.2%)	12/89 (13.5%)	RR 0.16 (0.05 to 0.49)	113 fewer per 1000 (from 69 fewer to 128 fewer)	HIGH	IMPORTANT
4 years	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ^(a)	none	9/161 (5.6)	5/73 (6.8)	RR 0.82 (0.28 to 2.35)	12 fewer per 1000 (from 49 fewer to 92 more)	LOW	IMPORTANT

^(a) If the CI for the meta-analysis was consistent with an appreciable clinically beneficial effect (or harmful as the case may be) of the intervention and an effect that indicates no clear clinical advantage, imprecision was graded as serious (these are default values set at a RR 0.75 and 1.25 or +/- 0.5*median control group SD); if the CI was consistent with both appreciable clinical benefit and an appreciable clinical harm, then imprecision was graded as very serious.

^(b) Due to the low rate of adverse events Peto OR were used (which is a more robust statistics with low or no events in either arm) and the risk difference was calculated rather than the absolute effect.

10.2.2 Economic evidence

Published literature

One economic evaluation was identified with the relevant comparison and has been included in this review. This is summarised in the economic evidence profile below (Table 92) and the economic evidence tables in Appendix H.

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

Table 92: Summary of economic studies included in the review

Study	Applicability	Limitations	Other comments	Incremental cost per patient (£)	Incremental effects (incremental risk of death at 90 days)	Cost effectiveness (£ per death averted (all-cause))	Uncertainty
Mealing 2013	Directly applicable ^(a)	Potentially serious limitations ^(b)	Based on outcomes from a small registry (EVEREST II high risk registry) 5 year time horizon	MitraClip versus MM= £26 989 increase	MitraClip versus MM= 1.22 QALYs gained	£22 153 per QALY gained	A probabilistic sensitivity analysis found the probability of MitraClip being cost effective versus MM to be 37% and 93% at £20 000 and £30 000 thresholds respectively. Univariate deterministic analysis showed the ICER to be sensitive to a number of parameters

Abbreviations: CI = 95% confidence interval; ICER = incremental cost-effectiveness ratio; MM = medical management; QALYs = quality-adjusted life years

(a) Recent evaluations and costs and effects are measured in a UK NHS context

(b) Main limitations: Based on a small single armed study and compared to a small concurrent case control; both with only 12 months follow-u, which necessitated parametric extrapolation for outcomes through years two to five.

Unit Costs

The cost of the MitraClip implant and percutaneous procedure is estimated to be £20,000.^{140,141}

New cost-effectiveness analysis

New analysis was not prioritised for this question.

10.2.3 Evidence statements

Clinical

Mortality

Low quality evidence from one randomised controlled trial (N=279) showed no clear differences in rates of mortality at 30-day, and at 1-, 2- and 4-year follow-up in the percutaneous and surgery groups.

Surgery or reoperation

At 30 days low quality evidence from one randomised controlled trial (N=279) reported that there was one re-surgery in the surgery group and none in the percutaneous group. Therefore, it shows no clear differences between treatment groups. For urgent or emergency cardiovascular surgery, low quality evidence shows lower rates of urgent surgery associated with the percutaneous group. This may represent a clinical risk associated with surgical treatment, but rates of events were low. High quality evidence from the same trial at 1 and 4 years showed that percutaneous treatment was associated with higher rates of surgery for mitral-valve dysfunction compared to surgical treatment indicating a clinical risk.

Grade 3+ or 4+ mitral regurgitation (severity of mitral regurgitation after 1 or 4 years)

Low quality evidence from one randomised controlled trial (N=279) showed no clear differences between surgery and percutaneous treatment in the rate of people with severe mitral regurgitation after 1 and 4 years.

Major adverse events – 30 day follow-up

In total high quality evidence from one randomised controlled trial (N=279) showed that percutaneous treatment was effective in lowering the rate of any major adverse events when considering any major adverse events excluding major blood transfusions. Percutaneous treatment was also associated with fewer major adverse events but to a lesser extent. When considering individual types of adverse events, rates for myocardial infarction, major stroke, renal failure, deep wound infection, septicaemia, new onset atrial fibrillation and gastrointestinal complications requiring surgery were low (with 0-2 events reported in either arm). Therefore no clear difference was observed (low quality evidence). Higher rates of invasive ventilation of >48 hrs and transfusions of 2 or more units of blood were associated with surgery compared to percutaneous treatment (high quality evidence).

Quality of life – SF-36 (30 days and 1 year)

Moderate quality evidence from one randomised controlled trial (N=279) suggests patients rated their physical quality of life higher in the percutaneous treatment at 30 days. However, high quality

evidence did not show an advantage of one treatment over the other for the mental component of quality of life at 30 day and on both physical and mental quality of life rating after 1 year.

Change in percentage ejection fraction

Moderate quality evidence from one randomised controlled trial (N=279) showed an improvement in ejection fraction associated with the percutaneous group compared to surgery, but it is uncertain whether this effect is large enough to constitute a clear clinical benefit.

Breathlessness - NYHA functional class III or IV (follow-up 12 months)

High quality evidence from one randomised controlled trial (N=279) showed an improvement in the rate of severe breathlessness associated with the percutaneous group compared to surgery. However, this was no longer different between the groups at the end of a 4 year follow-up period (low quality evidence).

Economic

One cost–utility analysis found that that percutaneous treatment was cost effective compared to standard medical care at a threshold of £30,000 per QALY gained for treating mitral regurgitation but was not cost-effective at a threshold of £20,000 per QALY gained (ICER: £22,153 per QALY gained). This analysis was assessed as directly applicable with potentially serious limitations.

10.2.4 Recommendations and links to evidence

Recommendations	34. Consider surgical mitral valve repair or replacement for people with heart failure due to severe mitral regurgitation assessed as suitable for surgery.
Relative values of different outcomes	The GDG attached most weight to mortality, but also took into account severity of mitral regurgitation at twelve months and two years, risks associated with the interventions, quality of life and need for additional procedures, such as re-operation or emergency surgery. The GDG considered the level of breathlessness after 12 months to be less important, and decided not to consider change in ejection fraction at 12 months.
Trade-off between clinical benefits and harms	Percutaneous treatment and surgery had similar rates of mortality and severity of mitral regurgitation (Grade 3+ or 4+ mitral regurgitation) at 12 months and longer term follow-up to 2 years. Percutaneous treatment was associated with lower need for invasive ventilation and transfusions of more than 2 units, but a higher need for re-operations. There was a difference in short-term (30 day) physical quality of life favouring percutaneous treatment, but this did not persist to 12 months.
Economic considerations	One recent economic evaluation was included in the economic review based on observational evidence. The GDG gave greater weight to the randomised controlled trial evidence included in the clinical review. The GDG concluded that percutaneous intervention was not cost effective, in that it was more costly than surgery with no sustained clinical benefit.
Quality of evidence	There was no randomised control trial evidence for percutaneous management compared to medical care or surgical management compared to medical care.. The EVEREST-trial, 2011 ^{71,72}) compared percutaneous management to surgery with four year follow-up results reported in 2013 ^{135,136} . The evidence from this trial was rated as high to low quality according to the GRADE criteria, depending on the

Recommendations	34. Consider surgical mitral valve repair or replacement for people with heart failure due to severe mitral regurgitation assessed as suitable for surgery.
	outcome measure. EVEREST II trial (Feldman et al, 2011 ^{71,72}) was relatively small, and most of the trial participants had chronic (rather than acute) mitral regurgitation. The study had extensive exclusion criteria so the results were applicable to only a subgroup of the overall population of people with mitral regurgitation.
Other considerations	<p>The GDG did not formally review the observational data on medical management and percutaneous repair which have been summarised in a number of systematic reviews. These data were not considered to offer robust evidence and the GDG agreed that randomised controlled trial data was necessary to give the least biased information about the effectiveness of percutaneous procedures. In the absence of RCT evidence, no recommendations were made for percutaneous management for people unsuitable for surgery. The GDG advocated the use of percutaneous management only in the context of a randomised controlled trial.</p> <p>In the absence of RCT evidence, no recommendations were made for percutaneous management for people unsuitable for surgery. The GDG considered the option of making a recommendation stating that percutaneous management should only be used in the context of a research trial. However, concerns were raised that this would not allow clinicians to make decisions for particular patients who may benefit from percutaneous repair. The GDG agreed not to make a recommendation, rather than advocate use of this procedure only in the context of a randomised controlled trial.</p> <p>The GDG did not consider that trial evidence supported the use of percutaneous repair over surgery, and therefore agreed to make a consensus recommendation in favour of surgery.</p> <p>The technology of percutaneous treatments is still evolving and it was noted that the device used in the clinical trial that was reviewed was subsequently withdrawn from the market and only reintroduced following modifications.</p>

11 Mechanical assist devices

In some patients with severe acute heart failure associated with severe haemodynamic compromise, medications alone are inadequate to support life and mechanical circulatory assistance may be considered. There is a range of different mechanical assist devices available and these can be grouped by their intended duration of use (short, intermediate, long term) and their means of insertion (percutaneous or surgical). The most commonly used percutaneous short term assistance device is the intra-aortic balloon pump which is widely available and relatively easy to place. Recent advances in technology have meant that there is now a number of different types of percutaneous short term devices with different advantages and disadvantages. Their impact on clinical outcomes in patients with acute heart failure is uncertain and their use may depend on regional expertise. In addition, there are a number of ventricular assist devices designed for long term use and are fully implantable. These devices require surgical placement and may allow the patient to leave hospital whilst either 'bridging to recovery' or 'bridging to transplantation'. At present in the UK, these devices are not funded for use in long term therapy for people who are not thought suitable for heart transplantation. The surgically inserted devices are of varied generations and it is difficult to consider these devices uniformly as they are not identical to each other in their performance. Similarly, comparison with the percutaneous devices may be difficult due to the different technology and clinical aim. At present, the use of implantable left ventricular assist devices is mainly restricted to expert transplant centres and there is variation in who is considered suitable for transfer and assessment for this advanced therapy.

Review question: For people with acute heart failure which, of the following, is the most clinically / cost effective intervention: (1) intra-aortic balloon counterpulsation (2) left ventricular assist devices or (3) medical care alone?

For full details see review protocol in Appendix C.

Table 93: PICO characteristics of review question

Population	Adults with acute heart failure (AHF)
Intervention/s	I. Intra-aortic balloon pump II. Left ventricular assist devices
Comparison/s	Medical care alone or each other
Outcomes	Mortality Length of hospital stay Admission to critical care units Re-admission rates Number of patients requiring invasive ventilation Quality of life Adverse events
Study design	Systematic reviews (SRs) Randomised controlled trials (RCTs)

11.1.1 Clinical evidence

A systematic literature search was undertaken to identify systematic reviews (SRs) or randomised controlled trials (RCTs) comparing the effectiveness of intra-aortic balloon pump (IABP) versus medical care alone, IABP versus left ventricular assist devices (LVAD), and LVAD versus medical care alone. Four sets of studies have been included in this review^{161,190,212,218}. Evidence from these is

summarised in the clinical GRADE evidence profile (Evidence profile Table 82). See also the study selection flow chart in Appendix D, forest plots in Appendix I, study evidence tables in Appendix G and lists of excluded studies in Appendix K. One Cochrane SR²¹⁸ with six included studies^{15,34,162,180,199,210} was identified. The Cochrane SR was restricted to patients with myocardial infarction (MI) and cardiogenic shock, and this only covers a minority of the core population of this present protocol. One of the excluded studies in the Cochrane review, O'Rourke (1981)¹⁶¹ satisfied our inclusion criteria and therefore it was included in this present review. An update search was carried out and one further study satisfied our inclusion criteria: Thiele (2012)^{210,212}, the IABP-SHOCK II trial. Additionally, the long-term follow-up data were extracted for this trial from another publication^{210,211}.

In the studies by O'Rourke (1981)¹⁶¹ and Thiele (2012)^{210,212}, IABP was compared against medical care. The Cochrane SR²¹⁸ compared IABP against non-IABP or other mechanical assist devices in their effects on mortality and morbidity. Although the Cochrane review applied the term, "LVAD" to cover all of the VADs compared against IABP this is not entirely accurate. The ventricular assist devices (VADs) used in two of the included studies^{34,210} were percutaneous VADs (TandemHeart™ Paracorporeal Ventricular Assist Device, Thoratec, US) that provide partial circulatory support by withdrawing oxygenated blood from the left atrium and propelling it back to one or both femoral arteries via arterial cannulas. The VAD used in the other included study¹⁹⁹ is also a percutaneous device (Impella® LP 2.5, Abiomed Europe GmbH, Germany) that provides partial circulatory support to the left side of the heart. One study was identified for the comparison between LVAD and medical care: Rose (2001)^{190,191}, the Randomized Evaluation of Mechanical Assistance for the Treatment of Congestive Heart Failure (REMATCH) study. Although its population and intervention are not directly applicable to those of our protocol, in the absence of evidence for this category of comparison, the decision was made to include this study. In the REMATCH study^{190,191}, the intervention assessed was an electrically powered implantable LVAD (HeartMate® XVE, Thoratec, US) that provided full circulatory support by pumping blood from the heart in a pulsatile flow in synchrony with the functioning cardiac cycle.

In undertaking meta-analyses for this present review, the studies were grouped into those that compared IABP against medical care alone, IABP against LVAD and LVAD against medical care alone. The authors of the Cochrane review obtained individual patient data from the studies that they included and carried out time-to-event analyses. Since these are statistically more robust analyses the time-to-event data from the SHOCK-II trial and the REMATCH study were added to our meta-analyses.

Summary of included studies

The characteristics of the included studies are briefly outlined in the below. Details of the studies can be found in Appendix G.

Table 94: Summary of the included studies

Study	Intervention / comparison	Population	Outcomes	Comments
Unverzagt et al., 2011 ²¹⁸	IABP vs. medical care IABP vs. LVAD <u>LVAD types</u> Percutaneous: - TandemHeart™	Adult patients with acute MI complicated by cardiogenic shock who have undergone PCI, CABG or thrombolysis	All-cause mortality (in-hospital, at 30 days, at 6 months); length of hospital stay; adverse events	This Cochrane systematic review is updated here. It included 6 RCTs: 3 comparing IABP vs. medical care and 3 comparing IABP vs. LVAD.

Study	Intervention / comparison	Population	Outcomes	Comments
	(Thoratec, USA) in Thiele 2005 ²¹⁰ and Burkhoff 2006 ³⁴ - Impella® LP 2.5 (Abiomed Europe GmbH, Germany) in Seyfarth 2008 ¹⁹⁹	N = 242 (Germany)		
O'Rourke et al., 1981 ¹⁶¹	IABP vs. medical care	Adult patients with early MI complicated by AHF N = 30 (New Zealand)	All-cause mortality (in-hospital)	This study was previously excluded from the Cochrane systematic review but has been re-included since it matched our protocol. All patients in this study received pulmonary arterial cannulation using a Swan-Ganz catheter after randomization.
Thiele et al., 2012 – the IABP-SHOCK II trial ²¹⁰ Thiele et al., 2013 – the IABP-SHOCK II long-term follow-up ^{210,211}	IABP vs. Medical care	Adult patients with acute MI complicated by cardiogenic shock and with plan of early revascularisation N = 600 (Germany)	All-cause mortality (30 days); serious adverse events All-cause mortality (6 and 12 months); cardiac mortality; serious adverse events; health-related quality of life	This study was identified in the update search. It is substantially larger than all previous studies combined, therefore contributes most weight to the analysis. Long-term data from 6 and 12 months follow-up were available for all surviving participants and these were extracted from a separate publication.
Rose et al., 2001 – the REMATCH trial ^{190,191}	LVAD vs. Medical care <u>LVAD type</u> First generation implantable LVAD, HeartMate® XVE (Thoratec, US)	Adults with chronic end-stage heart failure and contraindications to transplantation* N = 129 (US)	All-cause mortality (by the final analysis point**); length of hospital stay; serious adverse events*** ; quality of life	*The study population consists of a mixture of those with chronic and acute decompensated heart failure: data for these two populations are not analysed separately. ** “The final analysis point” has not been defined clearly but it is known that it took place after enrolment ended at the pre-specified 92 nd death. ***Adverse events

Study	Intervention / comparison	Population	Outcomes	Comments
				were considered to be serious if they caused death or permanent disability, were life-threatening, or required/prolonged hospitalization.

Table 95: GRADE evidence profile: quality of evidence and summary of findings (IABP vs. medical care)

Quality assessment							No. of events / Total no. of participants in the group (%) or Mean (SD)		Summary of findings		Quality	Importance
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	IABP	Medical care	Relative effect: Hazard ratio (HR); relative risk (RR), Peto odds ratio (OR) (95% CI)	Absolute effect / mean difference (MD) (95% CI)		
All-cause in-hospital mortality rate ^{15,161,180}												
3	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	24/64 (37.5%)	18/46 (39.1%)	RR 0.97 (0.61 to 1.54)	13 fewer per 1000 (from 171 fewer to 237 more)	LOW	CRITICAL
All-cause long-term mortality rate (1 to 36 months with mean of 15 months; 12 months) ^{161,211}												
2	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	156/313 (49.8%)	155/312 (49.7%)	RR 1 (0.85 to 1.17)	0 fewer per 1000 (from 53 fewer to 60 more)	HIGH	CRITICAL
Survival at 6 months ^{162,180}												
2	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	14/29 (48.3%)	13/28 (46.4%)	HR 1.05 (0.40 to 2.76)	16 more per 1000 (from 243 fewer to 357 more)	LOW	CRITICAL
Survival at 12 months ^{210,211}												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious ^(a)	None	155/299 (51.8%)	152/296 (51.4%)	HR 1.01 (0.78 to 1.31)	3 more per 1000 (from 84 fewer to 97 more)	MODERATE	CRITICAL
Cardiac mortality at 12 months ^{210,211}												

Quality assessment							No. of events / Total no. of participants in the group (%) or Mean (SD)		Summary of findings		Quality	Importance
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	IABP	Medical care	Relative effect: Hazard ratio (HR); relative risk (RR), Peto odds ratio (OR) (95% CI)	Absolute effect / mean difference (MD) (95% CI)		
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	150/299 (50.2%)	148/296 (50%)	RR 1 (0.85 to 1.18)	0 fewer per 1000 (from 75 fewer to 90 more)	HIGH	CRITICAL
Serious adverse events - Myocardial infarction - In-hospital ^{162,180,212}												
3	RCT	No serious risk of bias	Serious ^(b)	No serious indirectness	Serious ^(a)	None	11/331 (3.3%)	5/329 (1.5%)	OR 2.11 (0.78 to 5.68)	14 more per 1000 (from 3 fewer to 57 more)	LOW	IMPORTANT
Serious adverse events – Myocardial infarction - Over 12 months ^{210,211}												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious ^(a)	None	13/144 (9.0%)	5/144 (3.5%)	OR 2.57 (0.99 to 6.67)	50 more per 1000 (from 0 fewer to 160 more)	MODERATE	IMPORTANT
Serious adverse events – Stroke – In-hospital ^{162,180,212}												
3	RCT	No serious risk of bias	Serious ^(b)	No serious indirectness	Very serious ^(a)	None	4/331 (1.2%)	5/329 (1.5%)	OR 0.76 (0.2 to 2.85)	4 fewer per 1000 (from 12 fewer to 27 more)	VERY LOW	IMPORTANT
Serious adverse events – Stroke – Over 12 months ^{210,211}												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	3/144 (2.1%)	2/144 (1.4%)	OR 1.5 (0.26 to 8.77)	7 more per 1000 (from 10 fewer to 97 more)	LOW	IMPORTANT
Serious adverse events – Major bleeding – In-hospital ^{161,162,212}												

Quality assessment							No. of events / Total no. of participants in the group (%) or Mean (SD)		Summary of findings		Quality	Importance
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	IABP	Medical care	Relative effect: Hazard ratio (HR); relative risk (RR), Peto odds ratio (OR) (95% CI)	Absolute effect / mean difference (MD) (95% CI)		
3	RCT	No serious risk of bias	Serious ^(b)	No serious indirectness	Very serious ^(a)	None	13/326 (4%)	14/324 (4.3%) 4.4%	OR 0.93 (0.43 to 2)	3 fewer per 1000 (from 25 fewer to 40 more)	VERY LOW	IMPORTANT
Serious adverse events – Infections – In-hospital ^{180,212}												
2	RCT	No serious risk of bias	Serious ^(b)	No serious indirectness	Serious ^(a)	None	47/319 (14.7%)	61/319 (19.1%) 10.2%	OR 0.72 (0.48 to 1.1)	26 fewer per 1000 (from 50 fewer to 9 more)	LOW	IMPORTANT
Serious adverse events – Limb ischaemia – In-hospital ^{162,180}												
2	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	1/31 (3.2%)	1/31 (3.1%) 5%	OR 0.96 (0.06 to 15.46)	2 fewer per 1000 (from 47 fewer to 399 more)	LOW	IMPORTANT
Length of hospital stay (better indicated by lower values) ¹⁸⁰												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	18.3 (14.5)	29.4 (28.6)	NA	MD 11.1 lower (24.96 lower to 2.76 higher)	LOW	IMPORTANT

(a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed one of the default MIDs (i.e. ranging from a noticeable effect to no effect). Outcomes were downgraded by two increments if 95% CI of the effect crossed both default MIDs (i.e. ranging from a noticeable improvement to a noticeable harm). Default MIDs are set at RRs or HRs of 0.75 to 1.25 for dichotomous outcomes, or at 0.5 of the median control group standard deviation either side of the line of no effect for continuous variables.

(b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I-squared value of 50 to 74% or chi-squared p-value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I-squared value of 75% or more).

Health-related quality of life at 6 and 12 months

Health-related quality of life of the study participants in Thiele 2013^{210,211} was measured using EuroQoL EQ-5D-3L questionnaire. However, numerical data were not shown in the publication but were presented as charts. For this reason, the outcome could not be included in the GRADE evidence profile above. Nonetheless, the charts provided clearly indicate that there were no differences between the IABP and control groups in terms of any of the five dimensions of quality of life (mobility, self-care, usual activities, pain / discomfort, and anxiety / depression). The quality of this evidence was rated as very low due to the absence of numerical data.

Table 96: GRADE evidence profile: quality of evidence and summary of findings (IABP vs. LVAD)

Quality assessment							No. of events / Total no. of participants in the group (%) or Mean (SD)		Summary of findings		Quality	Importance
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	IABP	LVAD	Relative effect: Hazard ratio (HR); relative risk (RR), Peto odds ratio (OR) (95% CI); mean difference (MD) (SD)	Absolute effect / mean difference (MD) (95% CI)		
All-cause mortality distribution (from individual patient data and one trial from the update search) - follow-up 30 days ^{34,199,210}												
3	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	19/43 (44.2%)	19/45 (42.2%)	HR 1.02 (0.54 to 1.93)	6 more per 1000 (from 166 fewer to 231 more)	LOW	CRITICAL
All-cause mortality (from individual patient data) - follow-up 6 months ¹⁹⁹												
1	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	6/13 (46.2%)	8/13 (61.5%)	HR 0.72 (0.24 to 2.13)	118 fewer per 1000 (from 410 fewer to 254 more)	LOW	CRITICAL
All-cause in-hospital mortality ^{199,210}												
2	RCT	No serious	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	15/33 (45.5%)	16/34 (47.1%)	RR 0.97 (0.58 to 1.62)	15 fewer per 1000	LOW	CRITICAL

Quality assessment							No. of events / Total no. of participants in the group (%) or Mean (SD)		Summary of findings		Quality	Importance
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	IABP	LVAD	Relative effect: Hazard ratio (HR); relative risk (RR), Peto odds ratio (OR) (95% CI); mean difference (MD) (SD)	Absolute effect / mean difference (MD) (95% CI)		
		risk of bias						48.4%		(from 203 fewer to 300 more)		
Serious adverse events (cardiovascular) - myocardial infarction ^{199,210}												
2	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious ^(a)	None	1/33 (3%)	1/34 (2.9%)	OR 1.02 ^(c) (0.06 to 16.39)	1 more per 1000 (from 37 fewer to 360 more)	LOW	CRITICAL
Serious adverse events (cardiovascular) - stroke ^{199,210}												
2	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	0/33 (0%)	0/34 (0%)	NA	NA	HIGH	CRITICAL
Serious adverse events (other) - major bleeding ^{34,199,210}												
3	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	10/47 (21.3%)	27/53 (50.9%)	RR 0.41 (0.24 to 0.71)	299 fewer per 1000 (from 164 to 370 fewer)	HIGH	IMPORTANT
								42.1%				
Serious adverse events (other) – infections ^{34,199,210}												
3	RCT	No serious	Very serious ^(b)	No serious indirectness	Very serious ^(a)	None	15/47 (31.9%)	17/53 (32.1%)	RR 0.95 (0.59 to 1.55)	13 fewer per 1000	VERY LOW	IMPORTANT

Quality assessment							No. of events / Total no. of participants in the group (%) or Mean (SD)		Summary of findings		Quality	Importance
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	IABP	LVAD	Relative effect: Hazard ratio (HR); relative risk (RR), Peto odds ratio (OR) (95% CI); mean difference (MD) (SD)	Absolute effect / mean difference (MD) (95% CI)		
		risk of bias						15.8%		(from 92 fewer to 104 more)		
Serious adverse events (other) - limb ischaemia ^{34,199,210}												
3	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious ^(a)	None	2/47 (4.3%)	12/53 (22.6%) 21.1%	RR 0.27 (0.08 to 0.87)	154 fewer per 1000 (from 53 to 192 fewer)	MODERATE	IMPORTANT
Length of hospital stay (days) [better indicated by lower values] ^{199,210}												
2	RCT	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious	None	13.6 (3.6) 12.4 (13.6)	18.2 (10.6) 13.1 (13.4)	NA	MD 3.08 lower (9.98 lower to 3.83 higher)	MODERATE	IMPORTANT

(a) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed one of the default MIDs (i.e. ranging from a noticeable effect to no effect). Outcomes were downgraded by two increments if 95% CI of the effect crossed both default MIDs (i.e. ranging from a noticeable improvement to a noticeable harm). Default MIDs are set at RRs or HRs of 0.75 to 1.25 for dichotomous outcomes, or at 0.5 of the median control group standard deviation either side of the line of no effect for continuous variables.

(b) Outcomes were downgraded by one increment if the degree of inconsistency across studies was deemed serious (I-squared value of 50 to 74% or chi-squared p-value of 0.05 or less). Outcomes were downgraded by two increments if the degree of inconsistency was deemed very serious (I-squared value of 75% or more).

(c) Due to the low event rate a Peto OR was calculated and the risk difference was used for the absolute effect.

Table 97: GRADE evidence profile: quality of evidence and summary of findings (LVAD vs. medical care)

Quality assessment	No. of events /	Summary of findings	Quality	Importance
--------------------	-----------------	---------------------	---------	------------

							Total no. of participants in the group (%); Rate per patient-year; Mean (SD)					
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	LVAD	Medical care	Relative effect: Hazard ratio (HR); relative risk (RR), Peto odds ratio (OR) (95% CI); mean difference (MD) (SD); rate ratio	Absolute effect / mean difference (MD) (95% CI)		
All-cause mortality at 2 years ^{190,191}												
1	RCT	No serious risk of bias	No serious inconsistency	Very serious ^(a)	Serious ^(b)	None	41/68 (60.3)	54/61 (88.5)	HR 0.7 (0.47 to 1.06)	105 fewer per 1000 (from 247 fewer to 14 more)	VERY LOW	CRITICAL
Number of cardiac deaths in 2 years ^{190,191}												
1	RCT	No serious risk of bias	No serious inconsistency	Very serious ^(a)	No serious imprecision	None	16/68 (23.5)	53/61 (86.9)	RR 0.27 (0.17 to 0.42)	191 fewer per 1000 (from 152 fewer to 217 fewer)	LOW	CRITICAL
Serious adverse events – all serious adverse events ^{190,191}												
1	RCT	No serious risk of bias	No serious inconsistency	Very serious ^(a)	No serious imprecision	None	6.45	2.75	Rate ratio ^(c) 2.35 (1.86 to 2.97)	NA	LOW	CRITICAL
Serious adverse events – neurologic dysfunction ^{(d)190,191}												
1	RCT	No serious risk of bias	No serious inconsistency	Very serious ^(a)	Very serious ^(b)	None	0.39	0.09	Rate ratio ^(c) 4.35 (1.31 to 14.44)	NA	VERY LOW	IMPORTANT
Serious adverse events – supraventricular arrhythmia ^{190,191}												

Quality assessment							No. of events / Total no. of participants in the group (%); Rate per patient-year; Mean (SD)		Summary of findings		Quality	Importance
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	LVAD	Medical care	Relative effect: Hazard ratio (HR); relative risk (RR), Peto odds ratio (OR) (95% CI); mean difference (MD) (SD); rate ratio	Absolute effect / mean difference (MD) (95% CI)		
1	RCT	No serious risk of bias	No serious inconsistency	Very serious ^(a)	Very serious ^(b)	None	0.12	0.03	Rate ratio ^(c) 3.92 (0.47 to 32.69)	NA	VERY LOW	IMPORTANT
Serious adverse events – sepsis ^{190,191}												
1	RCT	No serious risk of bias	No serious inconsistency	Very serious ^(a)	Serious ^(b)	None	0.60	0.30	Rate ratio ^(c) 2.03 (0.99 to 4.16)	NA	VERY LOW	IMPORTANT
Serious adverse events – renal failure ^{190,191}												
1	RCT	No serious risk of bias	No serious inconsistency	Very serious ^(a)	Serious ^(b)	None	0.25	0.18	Rate ratio ^(c) 1.42 (0.54 to 3.73)	NA	VERY LOW	IMPORTANT
Serious adverse events – cardiac arrest ^{190,191}												
1	RCT	No serious risk of bias	No serious inconsistency	Very serious ^(a)	Very serious ^(b)	None	0.12	0.18	Rate ratio ^(c) 0.65 (0.21 to 14.3)	NA	VERY LOW	IMPORTANT
Serious adverse events – non-perioperative myocardial infarction ^{190,191}												
1	RCT	No serious risk of bias	No serious inconsistency	Very serious ^(a)	Very serious ^(b)	None	0.02	0.03	Rate ratio ^(c) 0.65 (0.04 to 10.56)	NA	VERY LOW	IMPORTANT

Quality assessment							No. of events / Total no. of participants in the group (%); Rate per patient-year; Mean (SD)		Summary of findings		Quality	Importance
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	LVAD	Medical care	Relative effect: Hazard ratio (HR); relative risk (RR), Peto odds ratio (OR) (95% CI); mean difference (MD) (SD); rate ratio	Absolute effect / mean difference (MD) (95% CI)		
Serious adverse events – ventricular arrhythmia ^{190,191}												
1	RCT	No serious risk of bias	No serious inconsistency	Very serious ^(a)	Serious ^(b)	None	0.25	0.56	Rate ratio ^(c) 0.45 (0.22 to 0.92)	NA	VERY LOW	IMPORTANT
Quality of life – Minnesota Living with Heart Failure questionnaire score (better indicated by lower values) ^{190,191}												
1	RCT	Serious ^(d)	No serious inconsistency	Very serious ^(a)	Serious ^(b)	None	41 (22)	58 (21)	NA	MD 17 lower (from 36.06 lower to 2.06 higher)	VERY LOW	CRITICAL
Quality of life – SF36: physical function (better indicated by higher values) ^{190,191}												
1	RCT	Serious ^(d)	No serious inconsistency	Very serious ^(a)	Serious ^(b)	None	46 (19)	21 (21)	NA	MD 25 higher (from 6.49 to 43.51 higher)	VERY LOW	CRITICAL
Quality of life – SF36: emotional role (better indicated by higher values) ^{190,191}												
1	RCT	Serious ^(d)	No serious inconsistency	Very serious ^(a)	No serious imprecision	None	64 (45)	17 (28)	NA	MD 47 higher (from 18.01 to 75.99 higher)	VERY LOW	CRITICAL

- (a) The study population largely consists of chronic heart failure patients and it is not possible to separate those with acute symptoms from the rest of the study participants. The evidence here therefore is not directly applicable to the protocol population. Similarly, the intervention was also not directly applicable for use in acute patients. The LVAD here was intended as destination therapy rather than for short-term use. NICE/NHS currently does not fund destination therapies.
- (b) Outcomes were downgraded by one increment if the upper or lower 95% CI crossed one of the default MIDs (i.e. ranging from a noticeable effect to no effect). Outcomes were downgraded by two increments if 95% CI of the effect crossed both default MIDs (i.e. ranging from a noticeable improvement to a noticeable harm). Default MIDs are set at RRs or HRs of 0.75 to 1.25 for dichotomous outcomes, or at 0.5 of the median control group standard deviation either side of the line of no effect for continuous variables.
- (c) Incidences of serious adverse events were measured in rate per patient-year. The relative effects of the two groups were calculated as rate ratio.
- (d) Neurologic dysfunction includes stroke, transient ischaemic attacks and toxic or metabolic encephalopathy.
- (e) The proportion of participants alive at 1 year who underwent assessment of quality of life was small. There was also a considerable difference in the proportions of participants who completed the assessment between the LVAD and medical care groups.

11.1.2 Economic evidence

Published literature

Two economic evaluations were identified with the relevant comparison and have been included in this review. These are summarised in the economic evidence profile below (Table 98) and the economic evidence tables in Appendix H.

One economic evaluation relating to this review question was identified but was excluded due to a combination of limited applicability and methodological limitations. This is summarised in Appendix L, with reasons for exclusion given.

No relevant economic evaluations were identified that compared intra-aortic balloon counter-pulsation with left ventricular assist devices, or intra-aortic balloon counter-pulsation with standard medical care.

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

Table 98: Economic evidence profile: Left Ventricular Assist Devices as Bridge to Transplantation versus Usual Medical Care

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost effectiveness	Uncertainty
Sutcliffe 2013 UK	Partially applicable ^(a)	Potentially serious limitations ^(b)	UK Health Technology Assessment Second and third generation LVADs BTT intent	LVADS = £127 391 increase	LVADs = 2.38 QALYs increase	£53 527/ QALY gained End of Life ICER: £36 190/ QALY gained ^(c)	CI: £31 802 to £94 853 (pa) The ICER was robust under a range of univariate tests. To bring the ICER to £30 000 per QALY would require a reduction in device cost of 76%

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost effectiveness	Uncertainty
Moreno 2012 UK NHS	Partially applicable ^(a)	Potentially serious limitations ^(d)	Second generation LVAD (HeartMate II) Single US uncontrolled multi-centre study (Pagani 2009) Control: patients on the US transplant registry 2000-2005	LVADs = £142 495 increase	LVADs = 0.55 QALYs increase	£258 922/ QALY gained	CI: £140 000 to £980 000 (pa) When the device acquisition cost is reduced to £0 and the bridging period is extended to 18-months the ICER=£24 063 per QALY gained

Abbreviations: BTT: bridge to transplant; CI: 95% confidence interval; HT: heart transplant; LVAD: left-ventricular assist device; MM: medical management; pa: probabilistic analysis; QALYs: quality-adjusted life years; UK: United Kingdom; VAD: ventricular assist device

(a) The study evaluated an indirectly relevant population whereby the use of the LVAD was intended for bridging to transplant. The review question considers the intent to bridge to recovery.

(b) Study was structurally limited by the absence of a directly comparable control group. Survival estimates are not from randomised controlled sources.

(c) Calculated under the assumption of life lived in full health for patients in the intervention (LVAD) group .i.e. QALYs gained are substituted by life years gained

New cost-effectiveness analysis

New analysis was not prioritised for this area.

Unit costs

In the absence of recent UK cost-effectiveness analysis, relevant unit costs were provided to aid consideration of cost-effectiveness. These are given in Appendix N.

Economic considerations

All included economic evaluations of left ventricular assist devices drew on the outcomes of patients in the United States. It is important to consider the generalizability of these health outcomes to the NHS setting both in respect of survival and resource utilisation.

11.1.3 Evidence statements

Clinical

Intra-aortic balloon pump (IABP) vs. Medical care

All-cause in-hospital mortality

Low quality evidence from 3 RCTs with 110 participants showed similar rates of mortality before discharge from hospital between the participants receiving IABP and those receiving medical care.

All-cause long-term mortality

High quality evidence from 2 RCTs with 625 participants showed similar rates of long-term mortality (1 RCT with a mean of 15 months follow-up and another RCT with a 12 month follow-up) between the participants receiving IABP and those receiving medical care.

Survival after discharge

Low quality evidence based on individual patient data from 2 RCTs with 67 participants showed similar lengths and rates of survival at 6 months between the participants receiving IABP and those receiving medical care.

Moderate quality evidence based on individual patient data from 1 RCT with 595 participants showed similar lengths and rates of survival at 12 months between the participants receiving IABP and those receiving medical care.

Deaths from cardiovascular events

High quality evidence from 1 RCT with 595 participants showed similar rates of deaths from cardiovascular events between the participants receiving IABP and those receiving medical care.

Serious adverse events – myocardial infarction

Low quality evidence from 3 RCTs with 660 participants showed that IABP was associated with a higher incidence of in-hospital myocardial infarction than medical care.

Moderate quality evidence from 1 RCT with 288 participants showed that IABP was associated with a higher incidence of myocardial infarction at 12 month follow-up than medical care.

Serious adverse events – stroke

Very low quality evidence from 3 RCTs with 660 participants showed that the incidence rates of in-hospital myocardial infarction were similar between IABP and medical care.

Low quality evidence from 1 RCT with 288 participants showed that the incidence rates of myocardial infarction were similar between IABP and medical care.

Serious adverse events - other

Very low quality evidence from 3 RCTs with 650 participants showed similar rates of in-hospital bleeding between IABP and medical care.

Low quality evidence from 2 RCTs with 638 participants showed that IABP was associated with a lower rate of in-hospital infections than medical care, however, there is uncertainty in drawing a clear conclusion.

Low quality evidence from 2 RCTs with 62 participants showed similar rates of in-hospital limb ischaemia between IABP and medical care.

Length of hospital stay

Low quality evidence from 1 RCT with 40 participants showed that IABP was associated with a shorter mean length of hospital stay compared to medical care, however, there is uncertainty in drawing a clear conclusion.

Health-related quality of life

Very low quality evidence from 1 RCT with 600 participants showed that the health-related quality of life measures at 6 and 12 months follow-up were similar between IABP and medical care groups. However, this set of outcomes could not be assessed by the GRADE criteria due to the absence of numerical data for each intervention group.

The following GDG prioritised outcomes were not reported in the studies included: re-admission rates, admission to critical care and number of patients requiring invasive ventilation.

Intra-aortic balloon pump (IABP) vs. Left ventricular assist devices (LVAD)

All-cause mortality

Low quality evidence from 2 RCTs with 67 participants, 3 RCTs with 88 participants and 1 RCT with 26 participants showed that the mortality rates before discharge from hospital, and length and rates of survival at 30 days and 6 months, respectively, were similar between the participants who received IABP and LVAD.

Serious adverse events (cardiovascular)

Low quality evidence from 2 RCTs with 67 participants showed that the incidence of myocardial infarctions was similar between the IABP and LVAD groups, and in addition, there were no events of stroke in either group.

Serious adverse events (other)

High quality evidence from 3 RCTs with 100 participants showed that IABP was associated with a lower incidence of major bleeding than LVAD. Very low quality evidence from 3 RCTs with 100 participants showed that the incidences of infections were similar in both intervention groups. Moderate quality evidence from 3 RCTs with 100 participants showed that the incidence of limb ischaemia was lower in IABP than in LVAD.

Length of hospital stay

Moderate quality evidence from 2 RCTs with 37 participants showed that IABP was associated with a shorter mean length of hospital stay compared to LVAD, however, there is uncertainty in drawing a clear conclusion.

The following GDG prioritised outcomes were not reported in the studies included: health-related quality of life, re-admission rates, admission to critical care and number of patients requiring invasive ventilation.

Left ventricular assist devices (LVAD) vs. Medical care

All-cause mortality

Low quality evidence from one RCT with 129 participants showed that lengths and rates of survival were better in the LVAD group than it was in the medical care group.

Cardiovascular deaths

Moderate quality evidence from one RCT with 129 participants showed that LVAD was associated with a lower incidence of cardiovascular deaths than medical care.

Serious adverse events

Very low to low quality evidence from one RCT with 129 participants showed variable comparative effectiveness of LVAD and medical care for different categories of serious adverse events. When all serious adverse events were combined, medical care was associated with a lower incidence of serious adverse events than LVAD. The quality of evidence was moderate.

Quality of life

Very low to low quality evidence based on scores from the questionnaires, the Minnesota Living with Heart Failure and SF-36 (physical function and emotional role) obtained from one RCT with 29 participants showed that LVAD was associated with better quality of life than medical care.

The following GDG prioritised outcomes were not reported in the studies included: length of hospital stay, re-admission rates, admission to critical care and number of patients requiring invasive ventilation.

Economic

Left ventricular assist devices vs. Medical care

Two cost–utility analyses found that LVADs as a bridge to transplant was not cost effective compared to usual medical care for patients with end stage heart failure (ICERs: £53,527 per QALY gained, and £258,922 per QALY gained, respectively). Both were assessed as partially applicable with potentially serious limitations.

11.1.4 Recommendations and link to evidence

Recommendations	<p>35. At an early stage, the specialist should have a discussion with a centre providing mechanical circulatory support about:</p> <ul style="list-style-type: none"> • people with potentially reversible severe acute heart failure or • people who are potential candidates for transplantation.
Relative values of different outcomes	The GDG considered mortality and health-related quality of life to be the most important outcomes, but also discussed serious adverse events and length of hospital stay.
Trade-off between clinical benefits and harms	<p><i>Intra-aortic balloon pump (IABP) versus medical care</i></p> <p>Short-term circulatory support with an IABP for cardiogenic shock complicating myocardial infarction showed no survival advantage over medical care.</p>

Recommendations	<p>35. At an early stage, the specialist should have a discussion with a centre providing mechanical circulatory support about:</p> <ul style="list-style-type: none"> • people with potentially reversible severe acute heart failure or • people who are potential candidates for transplantation.
	<p><i>IABP versus percutaneous left ventricular assist device (LVAD)</i></p> <p>Short-term circulatory support with a percutaneous LVAD showed no survival advantage over IABP and a higher rate of bleeding and limb ischaemia.</p> <p><i>LVAD versus medical care</i></p> <p>Long-term use of an LVAD in patients ineligible for transplantation was associated with better survival and health-related quality of life than medical care, but at the expense of increased adverse events such as bleeding, stroke and device related infection.</p>
Economic considerations	<p><i>IABP versus medical care</i></p> <p>The GDG considered that IABP would be more expensive than medical care so in the absence of any robust evidence of clinical benefit that it was unlikely to be cost-effective.</p> <p><i>IABP versus percutaneous LVAD</i></p> <p>No economic evidence was identified.</p> <p><i>LVADs versus medical care</i></p> <p>No randomised controlled trials were identified comparing LVADs with medical care in patients for whom implantation would be intended as a bridge to recovery. One recent UK economic evaluation found second and third generation LVADs were not cost-effective when compared with medical care when used as a bridge to transplant even when considered as life-extending treatment at the end of life.*</p> <p>In this circumstance, the cost per QALY is calculated under the assumption of survival in full-health for the intervention strategy (i.e. maximum utility whilst alive).</p> <p>There was no clear benefit to use of IABP over LVAD or medical care in terms of mortality and length of hospital stay. Although incidences of major bleeding and limb ischaemia were higher for percutaneous LVAD than IABP there were no differences in incidence of myocardial infarction, stroke or infections. The GDG noted that for all three sets of comparisons the evidence was not generalisable to the overall acute heart failure population. Therefore, it was not possible for the GDG to set out a recommendation based on the reported benefits and harms of the available evidence.</p>
Quality of evidence	<p>For all three comparisons, the identified evidence was of limited direct relevance to the management of patients with acute heart failure:</p> <p><i>IABP versus medical care</i></p> <p>The evidence from randomised controlled trials relates only to patients who had myocardial infarction complicated by cardiogenic shock.</p> <p><i>IABP versus percutaneous LVAD</i></p> <p>The evidence base is limited to patients with myocardial infarction complicated by cardiogenic shock.</p>

Recommendations	<p>35. At an early stage, the specialist should have a discussion with a centre providing mechanical circulatory support about:</p> <ul style="list-style-type: none"> • people with potentially reversible severe acute heart failure or • people who are potential candidates for transplantation.
	<p><i>LVADs versus medical care</i></p> <p>The one randomised controlled trial, REMATCH, examined the role of LVADs as a destination therapy (i.e. for long-term use), mostly in people with chronic end stage heart failure and with contraindications to transplantation.</p>
Other considerations	<p>The clinical evidence reviewed was largely in post-myocardial infarction patients receiving short-term support devices. The economic review concerned implantable VADs intended for intermediate to longer term support. There is fairly widespread availability of IABPs but access to other support devices is limited to specialist centres. Although there is good evidence that IABP or other short-term percutaneous support devices should not be used routinely in cardiogenic shock following myocardial infarction their role in specific cases and in cardiogenic shock of other aetiologies is unknown. The GDG noted that long-term mechanical support is an established therapy in some other countries (e.g. USA) but is currently not funded in the UK. In the future, with improvements in technology and reductions in cost, these therapies may become cost-effective. The GDG highlighted the existence of more contemporary mechanical assist devices and the emergence of data for the use of extracorporeal membrane oxygenation (ECMO) in acute heart failure patients (NICE IPG482) and the recognised the need for further research in these areas.</p> <p>The GDG concluded that they were not able to make specific recommendations on the use of mechanical circulatory support in acute heart failure patients. Due to the complexity of the clinical circumstances, and the variety of devices available which differ in mode of insertion, clinical outcomes and cost, the GDG gave emphasis to the importance of early involvement and prompt discussion with centres that have expertise in providing mechanical circulatory support.</p>

12 Reference list

- 1 A controlled trial to improve care for seriously ill hospitalized patients. The study to understand prognoses and preferences for outcomes and risks of treatments (SUPPORT). The SUPPORT Principal Investigators. *JAMA*. 1995; 274(20):1591-1598
- 2 Aaser E, Gullestad L, Tollofsrud S, Lundberg J, Hall C, Djoseand O et al. Effect of bolus injection versus continuous infusion of furosemide on diuresis and neurohormonal activation in patients with severe congestive heart failure. *Scandinavian Journal of Clinical and Laboratory Investigation*. 1997; 57(4):361-367
- 3 ACE Inhibitor Myocardial Infarction Collaborative Group. Indications for ACE inhibitors in the early treatment of acute myocardial infarction: systematic overview of individual data from 100,000 patients in randomized trials. *Circulation*. 1998; 97(22):2202-2212
- 4 Adamopoulos C, Ahmed A, Fay R, Angioi M, Filippatos G, Vincent J et al. Timing of eplerenone initiation and outcomes in patients with heart failure after acute myocardial infarction complicated by left ventricular systolic dysfunction: insights from the EPHESUS trial. *European Journal of Heart Failure*. 2009; 11(11):1099-1105
- 5 Adamopoulos S, Parissis JT, Iliodromitis EK, Paraskevaidis I, Tsiapras D, Farmakis D et al. Effects of levosimendan versus dobutamine on inflammatory and apoptotic pathways in acutely decompensated chronic heart failure. *American Journal of Cardiology*. 2006; 98(1):102-106
- 6 Adelaide Health Technology Assessment. B-type natriuretic peptide assays in the diagnosis of heart failure. Part A: in the hospital emergency setting. Part B: in the non-hospital setting. Adelaide. Adelaide Health Technology Assessment (AHTA), 2007. Available from: [http://www.msac.gov.au/internet/msac/publishing.nsf/Content/2CDBC3816FDE8D20CA2575AD0082FD8E/\\$File/1087%20-%20Carbon-labelled%20urea%20breath%20tests%20Report.pdf](http://www.msac.gov.au/internet/msac/publishing.nsf/Content/2CDBC3816FDE8D20CA2575AD0082FD8E/$File/1087%20-%20Carbon-labelled%20urea%20breath%20tests%20Report.pdf)
- 7 Afaq MA, Shoraki A, Ivanov O, Srinivasan J, Bernstein L, Zarich SW. Validity of Amino Terminal pro-Brain Natriuretic Peptide in a Medically Complex Elderly Population. *Journal of Clinical Medicine Research*. 2011; 3(4):156-163
- 8 Agmy GM, Ghanem MK. CPAP versus BiPAP in acute cardiogenic pulmonary edema: experience with 129 patients (NCT00912158). 2009. Available from: <http://www.clinicaltrials.gov/> [Last accessed: 28 November 2012]
- 9 Ahmed A, Zhang Y, Mujib M, Arora T, Feller M, Kilgore M et al. Use of beta-blockers and reduction in all-cause mortality in hospitalized Medicare beneficiaries with acute diastolic heart failure: A propensity-matched study of the OPTIMIZE-HF. *Journal of the American College of Cardiology*. 2011; 57(14 Suppl.):E222
- 10 Ahmed A, Zhang Y, Mujib M, Arora T, Kilgore M, Aban I et al. Lack of association between discharge prescription of aldosterone antagonists and outcomes in medicare beneficiaries with heart failure and preserved ejection fraction: A propensity-matched study of the OPTIMIZE-HF. *Circulation*. 2011; 124(21 Suppl. 1):A8493
- 11 Alibay Y, Beauchet A, El Mahmoud R, Schmitt C, Brun-Ney D, Benoit MO et al. Plasma N-terminal pro-brain natriuretic peptide and brain natriuretic peptide in assessment of acute dyspnea. *Biomedicine and Pharmacotherapy*. 2005; 59(1-2):20-24

- 12 Allen LA, O'Connor CM. Management of acute decompensated heart failure. *Canadian Medical Association Journal*. 2007; 176(6):797-805
- 13 Allen LA, Rogers JG, Warnica JW, Disalvo TG, Tasissa G, Binanay C et al. High mortality without ESCAPE: the registry of heart failure patients receiving pulmonary artery catheters without randomization. *Journal of Cardiac Failure*. 2008; 14(8):661-669
- 14 Allen LA, Turer AT, Dewald T, Stough WG, Cotter G, O'Connor CM. Continuous versus bolus dosing of Furosemide for patients hospitalized for heart failure. *American Journal of Cardiology*. 2010; 105(12):1794-1797
- 15 Arias EA, Gonzalez-Chon O, Garcia-Lopez SM, Chacon MA, Noriega-Iriondo F, Vega RE et al. [Impact of the intra-aortic balloon pump in the mortality due to cardiogenic shock secondary to acute myocardial infarction]. *Archivos De Cardiologia De Mexico*. 2005; 75(3):260-266
- 16 Arques S, Roux E, Sbragia P, Gelisse R, Ambrosi P, Pieri B et al. Comparative accuracy of color M-mode and tissue Doppler echocardiography in the emergency diagnosis of congestive heart failure in chronic hypertensive patients with normal left ventricular ejection fraction. *American Journal of Cardiology*. 2005; 96(10):1456-1459
- 17 Arques S, Roux E, Sbragia P, Pieri B, Gelisse R, Luccioni R et al. Usefulness of bedside tissue Doppler echocardiography and B-type natriuretic peptide (BNP) in differentiating congestive heart failure from noncardiac cause of acute dyspnea in elderly patients with a normal left ventricular ejection fraction and permanent, nonvalvular atrial fibrillation: insights from a prospective, monocenter study. *Echocardiography*. 2007; 24(5):499-507
- 18 Auerbach AD, Hamel MB, Califf RM, Davis RB, Wenger NS, Desbiens N et al. Patient characteristics associated with care by a cardiologist among adults hospitalized with severe congestive heart failure. SUPPORT Investigators. Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments. *Journal of the American College of Cardiology*. 2000; 36(7):2119-2125
- 19 Auerbach AD, Hamel MB, Davis RB, Connors AF, Regueiro C, Desbiens N et al. Resource use and survival for patients hospitalized with congestive heart failure: differences in care by specialty of the attending physician. *Annals of Internal Medicine*. 2000; 132(3):191-200
- 20 Badawy SS, Fahmy A. Efficacy and cardiovascular tolerability of continuous veno-venous hemodiafiltration in acute decompensated heart failure: a randomized comparative study. *Journal of Critical Care*. 2012; 27:106.e7-106.e13
- 21 Barcarse E, Kazanegra R, Chen A, Chiu A, Clopton P, Maisel A. Combination of B-type natriuretic peptide levels and non-invasive hemodynamic parameters in diagnosing congestive heart failure in the emergency department. *Congestive Heart Failure*. 2004; 10(4):171-176
- 22 Bart BA, Boyle A, Bank AJ, Anand I, Olivari MT, Kraemer M et al. Ultrafiltration versus usual care for hospitalized patients with heart failure: the Relief for Acutely Fluid-Overloaded Patients With Decompensated Congestive Heart Failure (RAPID-CHF) trial. *Journal of the American College of Cardiology*. 2005; 46(11):2043-2046
- 23 Bart BA, Goldsmith SR, Lee KL, Redfield MM, Felker GM, O'Connor CM et al. Cardiorenal rescue study in acute decompensated heart failure: rationale and design of CARRESS-HF, for the Heart Failure Clinical Research Network. *Journal of Cardiac Failure*. 2012; 18(3):176-182

- 24 Bart BA, Goldsmith SR, Lee KL, Givertz MM, O'Connor CM, Bull DA et al. Ultrafiltration in Decompensated Heart Failure with Cardiorenal Syndrome. *New England Journal of Medicine*. 2012; 367(24):2296-2304
- 25 Bayes-Genis A, Santalo-Bel M, Zapico-Muniz E, Lopez L, Cotes C, Bellido J et al. N-terminal probrain natriuretic peptide (NT-proBNP) in the emergency diagnosis and in-hospital monitoring of patients with dyspnoea and ventricular dysfunction. *European Journal of Heart Failure*. 2004; 6(3):301-308
- 26 Behnes M, Brueckmann M, Ahmad-Nejad P, Lang S, Wolpert C, Elmas E et al. Diagnostic performance and cost effectiveness of measurements of plasma N-terminal pro brain natriuretic peptide in patients presenting with acute dyspnea or peripheral edema. *International Journal of Cardiology*. 2009; 135(2):165-174
- 27 Berdague P, Caffin PY, Barazer I, Vergnes C, Sedighian S, Letrillard S et al. Use of N-terminal prohormone brain natriuretic peptide assay for etiologic diagnosis of acute dyspnea in elderly patients. *American Heart Journal*. 2006; 151(3):690-698
- 28 Bersten AD, Holt AW, Vedig AE, Skowronski GA, Baggoley CJ. Treatment of severe cardiogenic pulmonary edema with continuous positive airway pressure delivered by face mask. *New England Journal of Medicine*. 1991; 325(26):1825-1830
- 29 Blonde-Cynober F, Morineau G, Estrugo B, Fillie E, Aussel C, Vincent J-P. Diagnostic and prognostic value of brain natriuretic peptide (BNP) concentrations in very elderly heart disease patients: Specific geriatric cut-off and impacts of age, gender, renal dysfunction, and nutritional status. *Archives of Gerontology and Geriatrics*. 2011; 52(1):106-110
- 30 Boom NK, Lee DS, Tu JV. Comparison of processes of care and clinical outcomes for patients with newly hospitalized for heart failure attended by different physician specialists. *American Heart Journal*. 2012; 163(2):252-259
- 31 Bradley SM, Levy WC, Veenstra DL. Cost-consequences of ultrafiltration for acute heart failure: a decision model analysis. *Circulation: Cardiovascular Quality and Outcomes*. 2009; 2(6):566-573
- 32 Breidthardt T, Laule K, Strohmeyer AH, Schindler C, Meier S, Fischer M et al. Medical and economic long-term effects of B-type natriuretic peptide testing in patients with acute dyspnea. *Clinical Chemistry*. 2007; 53(8):1415-1422
- 33 Brezins M, Benari B, Papo V, Cohen A, Bursztein S, Markiewicz W. Left ventricular function in patients with acute myocardial infarction, acute pulmonary edema, and mechanical ventilation: relationship to prognosis. *Critical Care Medicine*. 1993; 21(3):380-385
- 34 Burkhoff D, Cohen H, Brunckhorst C, O'Neill WW, TandemHeart Investigators Group. A randomized multicenter clinical study to evaluate the safety and efficacy of the TandemHeart percutaneous ventricular assist device versus conventional therapy with intraaortic balloon pumping for treatment of cardiogenic shock. *American Heart Journal*. 2006; 152(3):469
- 35 Chen AA, Wood MJ, Krauser DG, Baggish AL, Tung R, Anwaruddin S et al. NT-proBNP levels, echocardiographic findings, and outcomes in breathless patients: results from the ProBNP Investigation of Dyspnoea in the Emergency Department (PRIDE) echocardiographic substudy. *European Heart Journal*. 2006; 27(7):839-845

- 36 Chen HH, Anstrom KJ, Givertz MM, Stevenson LW, Semigran MJ, Goldsmith SR et al. Low-dose dopamine or low-dose nesiritide in acute heart failure with renal dysfunction: the ROSE acute heart failure randomized trial. *JAMA : the Journal of the American Medical Association*. 2013; 310(23):2533-2543
- 37 Chenevier-Gobeaux C, Claessens YE, Voyer S, Desmoulins D, Ekindjian OG. Influence of renal function on N-terminal pro-brain natriuretic peptide (NT-proBNP) in patients admitted for dyspnoea in the Emergency Department: comparison with brain natriuretic peptide (BNP). *Clinica Chimica Acta*. 2005; 361(1-2):167-175
- 38 Chenevier-Gobeaux C, Guerin S, Andre S, Ray P, Cynober L, Gestin S et al. Midregional pro-atrial natriuretic peptide for the diagnosis of cardiac-related dyspnea according to renal function in the emergency department: a comparison with B-type natriuretic peptide (BNP) and N-terminal proBNP. *Clinical Chemistry*. 2010; 56(11):1708-1717
- 39 Chung T, Sindone A, Foo F, Dwyer A, Paoloni R, Janu MR et al. Influence of history of heart failure on diagnostic performance and utility of B-type natriuretic peptide testing for acute dyspnea in the emergency department. *American Heart Journal*. 2006; 152(5):949-955
- 40 Cleland J, Dargie H, Hardman S, McDonagh T, and Mitchell P. National Heart Failure Audit. April 2011-March 2012. London. National Institute for Cardiovascular Outcomes Research (NICOR), 2012. Available from:
<http://www.ucl.ac.uk/nicor/audits/heartfailure/additionalfiles/pdfs/annualreports/annual12.pdf>
- 41 Cleland J, Dargie H, Hardman S, McDonagh T, and Mitchell P. National Heart Failure Audit. April 2012 - March 2013. London. National Institute for Cardiovascular Outcomes Research (NICOR), 2013. Available from:
<http://www.ucl.ac.uk/nicor/audits/heartfailure/additionalfiles/pdfs/annualreports/NHFA13medium.pdf>
- 42 Cleland JGF. Health economic consequences of the pharmacological treatment of heart failure. *European Heart Journal*. 1998; 19(Suppl. P):32-39
- 43 Cleland JGF, Ghosh J, Freemantle N, Kaye GC, Nasir M, Clark AL et al. Clinical trials update and cumulative meta-analyses from the American College of Cardiology: WATCH, SCD-HeFT, DINAMIT, CASINO, INSPIRE, STRATUS-US, RIO-Lipids and cardiac resynchronisation therapy in heart failure. *European Journal of Heart Failure*. 2004; 6(4):501-508
- 44 Cohn JN, Franciosa JA, Francis GS, Archibald D, Tristani F, Fletcher R et al. Effect of short-term infusion of sodium nitroprusside on mortality rate in acute myocardial infarction complicated by left ventricular failure: results of a Veterans Administration cooperative study. *New England Journal of Medicine*. 1982; 306(19):1129-1135
- 45 Costanzo MR, Guglin ME, Saltzberg MT, Jessup ML, Bart BA, Teerlink JR et al. Ultrafiltration Versus Intravenous Diuretics for Patients Hospitalized for Acute Decompensated Heart Failure. *Journal of the American College of Cardiology*. 2007; 49(6):675-683
- 46 Costanzo MR, Saltzberg MT, Jessup M, Teerlink JR, Sobotka PA, Ultrafiltration Versus Intravenous Diuretics for Patients Hospitalized for Acute Decompensated Heart Failure Investigators. Ultrafiltration is associated with fewer rehospitalizations than continuous diuretic infusion in patients with decompensated heart failure: results from UNLOAD. *Journal of Cardiac Failure*. 2010; 16(4):277-284

- 47 Cotter G, Metzkor E, Kaluski E, Faigenberg Z, Miller R, Simovitz A et al. Randomised trial of high-dose isosorbide dinitrate plus low-dose furosemide versus high-dose furosemide plus low-dose isosorbide dinitrate in severe pulmonary oedema. *Lancet*. 1998; 351(9100):389-393
- 48 Cotter G, Weissgarten J, Metzkor E, Moshkovitz Y, Litinski I, Tavori U et al. Increased toxicity of high-dose furosemide versus low-dose dopamine in the treatment of refractory congestive heart failure. *Clinical Pharmacology and Therapeutics*. 1997; 62(2):187-193
- 49 Crane SD, Elliott MW, Gilligan P, Richards K, Gray AJ. Randomised controlled comparison of continuous positive airways pressure, bilevel non-invasive ventilation, and standard treatment in emergency department patients with acute cardiogenic pulmonary oedema. *Emergency Medicine Journal*. 2004; 21(2):155-161
- 50 Cuffe MS, Califf RM, Adams KF, Bourge RC, Colucci W, Massie B et al. Rationale and design of the OPTIME CHF trial: outcomes of a prospective trial of intravenous milrinone for exacerbations of chronic heart failure. *American Heart Journal*. 2000; 139(1 Pt. 1):15-22
- 51 Cuffe MS, Califf RM, Adams KF, Benza R, Bourge R, Colucci WS et al. Short-term intravenous milrinone for acute exacerbation of chronic heart failure: a randomized controlled trial. *JAMA*. 2002; 287(12):1541-1547
- 52 Dao Q, Krishnaswamy P, Kazanegra R, Harrison A, Amirnovin R, Lenert L et al. Utility of B-type natriuretic peptide in the diagnosis of congestive heart failure in an urgent-care setting. *Journal of the American College of Cardiology*. 2001; 37(2):379-385
- 53 Davis M, Espiner E, Richards G, Billings J, Town I, Neill A et al. Plasma brain natriuretic peptide in assessment of acute dyspnoea. *Lancet*. 1994; 343(8895):440-444
- 54 De Backer D, Biston P, Devriendt J, Madl C, Chochrad D, Aldecoa C et al. Comparison of dopamine and norepinephrine in the treatment of shock. *New England Journal of Medicine*. 2010; 362(9):779-789
- 55 deFilippi CR, Seliger SL, Maynard S, Christenson RH. Impact of renal disease on natriuretic peptide testing for diagnosing decompensated heart failure and predicting mortality. *Clinical Chemistry*. 2007; 53(8):1511-1519
- 56 Delclaux C, L'Her E, Alberti C, Mancebo J, Abroug F, Conti G et al. Treatment of acute hypoxemic nonhypercapnic respiratory insufficiency with continuous positive airway pressure delivery by a face mask. *JAMA*. 2000; 284(18):2352-2360
- 57 Dokainish H. Tissue Doppler imaging in the evaluation of left ventricular diastolic function. *Current Opinion in Cardiology*. 2004; 19(5):437-441
- 58 Dokainish H, Zoghbi WA, Lakkis NM, Quinones MA, Nagueh SF. Comparative accuracy of B-type natriuretic peptide and tissue Doppler echocardiography in the diagnosis of congestive heart failure. *American Journal of Cardiology*. 2004; 93(9):1130-1135
- 59 Dormans TP, van Meyel JJ, Gerlag PG, Tan Y, Russel FG, Smits P. Diuretic efficacy of high dose furosemide in severe heart failure: bolus injection versus continuous infusion. *Journal of the American College of Cardiology*. 1996; 28(2):376-382

- 60 Dubourg O, Gueret P, Ferrier A, Farcot JC, Terdjman M, Rigaud M et al. Controlled, eight-hour haemodynamic study of a sustained-release formulation of isosorbide dinitrate in moderate left ventricular failure. *European Journal of Clinical Pharmacology*. 1984; 27(3):259-263
- 61 Ducros L, Logeart D, Vicaud E, Henry P, Plaisance P, Collet JP et al. CPAP for acute cardiogenic pulmonary oedema from out-of-hospital to cardiac intensive care unit: a randomised multicentre study. *Intensive Care Medicine*. 2011; 37(9):1501-1509
- 62 Eckstein J, Potocki M, Murray K, Breidthardt T, Ziller R, Mosimann T et al. Direct comparison of mid-regional pro-atrial natriuretic peptide with N-terminal pro B-type natriuretic peptide in the diagnosis of patients with atrial fibrillation and dyspnoea. *Heart*. 2012; 98(20):1518-1522
- 63 Elkayam U, Kulick D, McIntosh N, Roth A, Hsueh W, Rahimtoola SH. Incidence of early tolerance to hemodynamic effects of continuous infusion of nitroglycerin in patients with coronary artery disease and heart failure. *Circulation*. 1987; 76(3):577-584
- 64 Elkayam U, Akhter MW, Singh H, Khan S, Usman A. Comparison of effects on left ventricular filling pressure of intravenous nesiritide and high-dose nitroglycerin in patients with decompensated heart failure. *American Journal of Cardiology*. 2004; 93(2):237-240
- 65 Elkayam U, Bitar F, Akhter MW, Khan S, Patrus S, Derakhshani M. Intravenous nitroglycerin in the treatment of decompensated heart failure: potential benefits and limitations. *Journal of Cardiovascular Pharmacology and Therapeutics*. 2004; 9(4):227-241
- 66 ESCAPE Investigators and ESCAPE Study Coordinators. Evaluation study of congestive heart failure and pulmonary artery catheterization effectiveness: the ESCAPE trial. *JAMA*. 2005; 294(13):1625-1633
- 67 Ezekowitz JA, Lee DS, Tu JV, Newman AM, McAlister FA. Comparison of one-year outcome (death and rehospitalization) in hospitalized heart failure patients with left ventricular ejection fraction >50% versus those with ejection fraction <50%. *American Journal of Cardiology*. 2008; 102(1):79-83
- 68 Fabbian F, De Giorgi A, Pala M, Tiseo R, Portaluppi F. Elevated NT-proBNP levels should be interpreted in elderly patients presenting with dyspnea. *European Journal of Internal Medicine*. 2011; 22(1):108-111
- 69 Fairbairn TA, Meads DM, Hulme C, Mather AN, Plein S, Blackman DJ et al. The cost-effectiveness of transcatheter aortic valve implantation versus surgical aortic valve replacement in patients with severe aortic stenosis at high operative risk. *Heart*. United Kingdom 2013; 99(13):914-920
- 70 Fedullo AJ, Swinburne AJ, Wahl GW, Bixby K. Acute cardiogenic pulmonary edema treated with mechanical ventilation. Factors determining in-hospital mortality. *Chest*. 1991; 99(5):1220-1226
- 71 Feldman T, Mauri L, Kar S, Rinaldi M, Fail P, Hermiller J et al. EVEREST II Randomized Clinical Trial: A critical assessment of the safety data in patients treated with the MitraClip. *Journal of the American College of Cardiology*. 2010; 56(13 Suppl.):B23
- 72 Feldman T, Foster E, Glower DD, Kar S, Rinaldi MJ, Fail PS et al. Percutaneous repair or surgery for mitral regurgitation. *New England Journal of Medicine*. 2011; 364(15):1395-1406

- 73 Felker GM, Lee KL, Bull DA, Redfield MM, Stevenson LW, Goldsmith SR et al. Diuretic strategies in patients with acute decompensated heart failure. *New England Journal of Medicine*. 2011; 364(9):797-805
- 74 Felker GM, Mentz RJ. Diuretics and ultrafiltration in acute decompensated heart failure. *Journal of the American College of Cardiology*. 2012; 59(24):2145-2153
- 75 Felker GM, Benza RL, Chandler AB, Leimberger JD, Cuffe MS, Califf RM et al. Heart failure etiology and response to milrinone in decompensated heart failure: results from the OPTIME-CHF study. *Journal of the American College of Cardiology*. 2003; 41(6):997-1003
- 76 Ferrer M, Esquinas A, Leon M, Gonzalez G, Alarcon A, Torres A. Noninvasive ventilation in severe hypoxemic respiratory failure: a randomized clinical trial. *American Journal of Respiratory and Critical Care Medicine*. 2003; 168(12):1438-1444
- 77 Flather MD, Yusuf S, Kober L, Pfeffer M, Hall A, Murray G et al. Long-term ACE-inhibitor therapy in patients with heart failure or left-ventricular dysfunction: a systematic overview of data from individual patients. ACE-Inhibitor Myocardial Infarction Collaborative Group. *Lancet*. 2000; 355(9215):1575-1581
- 78 Fleischer D, Espiner EA, Yandle TG, Livesey JH, Billings J, Town I et al. Rapid assay of plasma brain natriuretic peptide in the assessment of acute dyspnoea. *New Zealand Medical Journal*. 1997; 110(1039):71-74
- 79 Fonarow GC. The role of in-hospital initiation of cardioprotective therapies to improve treatment rates and clinical outcomes. *Reviews in Cardiovascular Medicine*. 2002; 3(Suppl. 3):S2-S10
- 80 Fonarow GC, Abraham WT, Albert NM, Stough WG, Gheorghiade M, Greenberg BH et al. Carvedilol use at discharge in patients hospitalized for heart failure is associated with improved survival: An analysis from Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF). *American Heart Journal*. 2007; 153(1):82.e1-82.e11
- 81 Fonarow GC, Abraham WT, Albert NM, Stough WG, Gheorghiade M, Greenberg BH et al. Prospective evaluation of beta-blocker use at the time of hospital discharge as a heart failure performance measure: results from OPTIMIZE-HF. *Journal of Cardiac Failure*. 2007; 13(9):722-731
- 82 Fonarow GC, Abraham WT, Albert NM, Stough WG, Gheorghiade M, Greenberg BH et al. Influence of beta-blocker continuation or withdrawal on outcomes in patients hospitalized with heart failure: findings from the OPTIMIZE-HF program. *Journal of the American College of Cardiology*. 2008; 52(3):190-199
- 83 Frontin P, Bounes V, Houzé-Cerfon CH, Charpentier S, Houzé-Cerfon V, Ducassé JL. Continuous positive airway pressure for cardiogenic pulmonary edema: a randomized study. *American Journal of Emergency Medicine*. 2011; 29(7):775-781
- 84 Gargani L, Frassi F, Soldati G, Tesorio P, Gheorghiade M, Picano E. Ultrasound lung comets for the differential diagnosis of acute cardiogenic dyspnoea: a comparison with natriuretic peptides. *European Journal of Heart Failure*. 2008; 10(1):70-77
- 85 Gattis WA, O'Connor CM, Gallup DS, Hasselblad V, Gheorghiade M, IMPACT-HF Investigators and Coordinators. PredischARGE initiation of carvedilol in patients hospitalized for

- decompensated heart failure: results of the Initiation Management Predischarge: Process for Assessment of Carvedilol Therapy in Heart Failure (IMPACT-HF) trial. *Journal of the American College of Cardiology*. 2004; 43(9):1534-1541
- 86 Gattis WA, O'Connor CM, Gheorghiade M. Influence of beta-blocker continuation or withdrawal on outcomes in patients hospitalized with heart failure: findings from the OPTIMIZE-HF program. Use of beta-blockers and reduction in all-cause mortality in hospitalized Medicare beneficiaries with acute diastolic heart failure: A propensity-matched study of the OPTIMIZE-HF. *Reviews in Cardiovascular Medicine*. 2002; 3(Suppl. 3):S48-S54
- 87 Gegenhuber A, Struck J, Poelz W, Pacher R, Morgenthaler NG, Bergmann A et al. Midregional pro-A-type natriuretic peptide measurements for diagnosis of acute destabilized heart failure in short-of-breath patients: comparison with B-type natriuretic peptide (BNP) and amino-terminal proBNP. *Clinical Chemistry*. 2006; 52(5):827-831
- 88 Giamouzis G, Economou D, Karayannis G, Rovithis D, Nastas I, Kyrilidis T et al. The combination of low dose furosemide and low dose dopamine is effective and prevents worsening of renal function and hypokalemia during hospitalization for acute decompensated heart failure. *European Heart Journal*. 2009; 30(Suppl. 1):431
- 89 Giamouzis G, Butler J, Starling RC, Karayannis G, Nastas J, Parisis C et al. Impact of dopamine infusion on renal function in hospitalized heart failure patients: results of the Dopamine in Acute Decompensated Heart Failure (DAD-HF) Trial. *Journal of Cardiac Failure*. 2010; 16(12):922-930
- 90 Giglioli C, Landi D, Cecchi E, Chiostrì M, Gensini GF, Valente S et al. Effects of ULTRAFiltration vs. DiureticS on clinical, biohumoral and haemodynamic variables in patients with deCOMpensated heart failure: the ULTRADISCO study. *European Journal of Heart Failure*. 2011; 13(3):337-346
- 91 Goodacre S, Gray A, Newby D, Dixon S, Masson M, Sampson F et al. Health utility and survival after hospital admission with acute cardiogenic pulmonary oedema. *Emergency Medicine Journal*. 2011; 28(6):477-482
- 92 Gorissen C, Baumgarten R, de Groot M, van Haren E, Kragten H, Leers M. Analytical and clinical performance of three natriuretic peptide tests in the emergency room. *Clinical Chemistry and Laboratory Medicine*. 2007; 45(5):678-684
- 93 Gray A, Goodacre S, Newby DE, Masson M, Sampson F, Nicholl J. Noninvasive ventilation in acute cardiogenic pulmonary edema. *New England Journal of Medicine*. 2008; 359(2):142-151
- 94 Gray A, Goodacre S, Seah M, Tilley S. Diuretic, opiate and nitrate use in severe acidotic acute cardiogenic pulmonary oedema: analysis from the 3CPO trial. *QJM*. 2010; 103(8):573-581
- 95 Gray AJ, Goodacre S, Newby DE, Masson MA, Sampson F, Dixon S et al. A multicentre randomised controlled trial of the use of continuous positive airway pressure and non-invasive positive pressure ventilation in the early treatment of patients presenting to the emergency department with severe acute cardiogenic pulmonary oedema: the 3CPO trial. *Health Technology Assessment*. 2009; 13(33):1-106
- 96 Gruson D, Ketelslegers JM, Verschuren F, Thys F. Head-to-head comparison of the prohormone proBNP1-108 with BNP and Nt-proBNP in patients admitted to emergency department. *Clinical Biochemistry*. 2012; 45(3):249-252

- 97 Gruson D, Rousseau MF, Ahn S, Van Linden F, Thys F, Ketelslegers JM et al. Accuracy of N-terminal-pro-atrial natriuretic peptide in patients admitted to emergency department. *Scandinavian Journal of Clinical and Laboratory Investigation*. 2008; 68(5):410-414
- 98 Hancock-Howard RL, Feindel CM, Rodes-Cabau J, Webb JG, Thompson AK, Banz K. Cost effectiveness of transcatheter aortic valve replacement compared to medical management in inoperable patients with severe aortic stenosis: Canadian analysis based on the PARTNER Trial Cohort B findings. *Journal of Medical Economics*. 2013; 16(4):566-574
- 99 Hanna MA, Tang WHW, Teo BW, O'Neill JO, Weinstein DM, Lau SM et al. Extracorporeal ultrafiltration vs. conventional diuretic therapy in advanced decompensated heart failure. *Congestive Heart Failure*. 2012; 18(1):54-63
- 100 Harvey S, Harrison DA, Singer M, Ashcroft J, Jones CM, Elbourne D et al. Assessment of the clinical effectiveness of pulmonary artery catheters in management of patients in intensive care (PAC-Man): a randomised controlled trial. *Lancet*. 2005; 366(9484):472-477
- 101 Havelka EG, Rzechula KH, Bryant TO, Anneken SM, Kulstad EB. Correlation between impedance cardiography and B-type natriuretic peptide levels in dyspneic patients. *Journal of Emergency Medicine*. 2011; 40(2):146-150
- 102 Hoffman JR, Reynolds S. Comparison of nitroglycerin, morphine and furosemide in treatment of presumed pre-hospital pulmonary edema. *Chest*. 1987; 92(4):586-593
- 103 Howlett JG, Cox JL, Haddad H, Stanley J, McDonald M, Johnstone DE. Physician specialty and quality of care for CHF: different patients or different patterns of practice? *Canadian Journal of Cardiology*. 2003; 19(4):371-377
- 104 Iakobishvili Z, Cohen E, Garty M, Behar S, Shotan A, Sandach A et al. Use of intravenous morphine for acute decompensated heart failure in patients with and without acute coronary syndromes. *Acute Cardiac Care*. 2011; 13(2):76-80
- 105 Januzzi JL, van Kimmenade R, Lainchbury J, Bayes-Genis A, Ordonez-Llanos J, Santalo-Bel M et al. NT-proBNP testing for diagnosis and short-term prognosis in acute destabilized heart failure: an international pooled analysis of 1256 patients: the International Collaborative of NT-proBNP Study. *European Heart Journal*. 2006; 27(3):330-337
- 106 Januzzi JLI, Camargo CA, Anwaruddin S, Baggish AL, Chen AA, Krauser DG et al. The N-terminal Pro-BNP investigation of dyspnea in the emergency department (PRIDE) study. *American Journal of Cardiology*. 2005; 95(8):948-954
- 107 Jondeau G, Neuder Y, Eicher JC, Jourdain P, Fauveau E, Galinier M et al. B-CONVINCED: Beta-blocker CONTinuation Vs. INTerruption in patients with Congestive heart failure hospitalizED for a decompensation episode. *European Heart Journal*. 2009; 30(18):2186-2192
- 108 Joynt KE, Orav EJ, Jha AK. Physician volume, specialty, and outcomes of care for patients with heart failure. *Circulation Heart Failure*. 2013; 6(5):890-897
- 109 Karpaliotis D, Kirtane AJ, Ruisi CP, Polonsky T, Malhotra A, Talmor D et al. Diagnostic and prognostic utility of brain natriuretic Peptide in subjects admitted to the ICU with hypoxic respiratory failure due to noncardiogenic and cardiogenic pulmonary edema. *Chest*. 2007; 131(4):964-971

- 110 Kelly C, Newby DE, Boon NA, Douglas NJ. Support ventilation versus conventional oxygen. *Lancet*. 2001; 357(9262):1126
- 111 Kelly CA, Newby DE, McDonagh TA, Mackay TW, Barr J, Boon NA et al. Randomised controlled trial of continuous positive airway pressure and standard oxygen therapy in acute pulmonary oedema; effects on plasma brain natriuretic peptide concentrations. *European Heart Journal*. 2002; 23(17):1379-1386
- 112 Kevin Rogers R, Stehlik J, Stoddard GJ, Greene T, Collins SP, Peacock WF et al. Adjusting for clinical covariates improves the ability of B-type natriuretic peptide to distinguish cardiac from non-cardiac dyspnoea: a sub-study of HEARD-IT. *European Journal of Heart Failure*. 2009; 11(11):1043-1049
- 113 Klemen P, Golub M, Grmec S. Combination of quantitative capnometry, N-terminal pro-brain natriuretic peptide, and clinical assessment in differentiating acute heart failure from pulmonary disease as cause of acute dyspnea in pre-hospital emergency setting: study of diagnostic accuracy. *Croatian Medical Journal*. 2009; 50(2):133-142
- 114 Kodali SK, Williams MR, Smith CR, Svensson LG, Webb JG, Makkar RR et al. Two-year outcomes after transcatheter or surgical aortic-valve replacement. *New England Journal of Medicine*. 2012; 366(18):1686-1695
- 115 Kwong JSW, Yu CM. Ultrafiltration for acute decompensated heart failure: A systematic review and meta-analysis of randomized controlled trials [Epublication]. *International Journal of Cardiology*. 2014;
- 116 L'Her E, Duquesne F, Girou E, Rosiere XD, Conte P, Renault S et al. Noninvasive continuous positive airway pressure in elderly cardiogenic pulmonary edema patients. *Intensive Care Medicine*. 2004; 30(5):882-888
- 117 Lahav M, Regev A, Ra'anani P, Theodor E. Intermittent administration of furosemide vs continuous infusion preceded by a loading dose for congestive heart failure. *Chest*. 1992; 102(3):725-731
- 118 Lainchbury JG, Campbell E, Frampton CM, Yandle TG, Nicholls MG, Richards AM. Brain natriuretic peptide and n-terminal brain natriuretic peptide in the diagnosis of heart failure in patients with acute shortness of breath. *Journal of the American College of Cardiology*. 2003; 42(4):728-735
- 119 Leon MB, Smith CR, Mack M, Miller DC, Moses JW, Svensson LG et al. Placement of AoRTic TraNscathetER Valves. *Journal of the American College of Cardiology*. 2010; 56(13 Suppl.):xii
- 120 Leon MB, Smith CR, Mack M, Miller DC, Moses JW, Svensson LG et al. Transcatheter aortic-valve implantation for aortic stenosis in patients who cannot undergo surgery. *New England Journal of Medicine*. 2010; 363(17):1597-1607
- 121 Levitt MA. A prospective, randomized trial of BiPAP in severe acute congestive heart failure. *Journal of Emergency Medicine*. 2001; 21(4):363-369
- 122 Licata G, Di Pasquale P, Parrinello G, Cardinale A, Scandurra A, Follone G et al. Effects of high-dose furosemide and small-volume hypertonic saline solution infusion in comparison with a high dose of furosemide as bolus in refractory congestive heart failure: long-term effects. *American Heart Journal*. 2003; 145(3):459-466

- 123 Lin M, Chiang HT. The efficacy of early continuous positive airway pressure therapy in patients with acute cardiogenic pulmonary edema. *Journal of the Formosan Medical Association*. 1991; 90(8):736-743
- 124 Lin M, Yang YF, Chiang HT, Chang MS, Chiang BN, Cheitlin MD. Reappraisal of continuous positive airway pressure therapy in acute cardiogenic pulmonary edema. Short-term results and long-term follow-up. *Chest*. 1995; 107(5):1379-1386
- 125 Logeart D, Saudubray C, Beyne P, Thabut G, Ennezat PV, Chavelas C et al. Comparative value of Doppler echocardiography and B-type natriuretic peptide assay in the etiologic diagnosis of acute dyspnea. *Journal of the American College of Cardiology*. 2002; 40(10):1794-1800
- 126 Lokuge A, Lam L, Cameron P, Krum H, de Villiers S, Bystrzycki A et al. B-type natriuretic peptide testing and the accuracy of heart failure diagnosis in the emergency department. *Circulation Heart Failure*. 2010; 3(1):104-110
- 127 Lowe J. Specialist or generalist care? : a study of the impact of a selective admitting policy for patients with cardiac failure. *International Journal for Quality in Health Care*. 2000; 12(4):339-345
- 128 Maisel A. B-type natriuretic peptide in the diagnosis and management of congestive heart failure. *Cardiology Clinics*. 2001; 19(4):557-571
- 129 Maisel A, Mueller C, Nowak R, Peacock WF, Landsberg JW, Ponikowski P et al. Mid-region pro-hormone markers for diagnosis and prognosis in acute dyspnea: results from the BACH (Biomarkers in Acute Heart Failure) trial. *Journal of the American College of Cardiology*. 2010; 55(19):2062-2076
- 130 Maisel AS, Krishnaswamy P, Nowak RM, McCord J, Hollander JE, Duc P et al. Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. *New England Journal of Medicine*. 2002; 347(3):161-167
- 131 Makkar RR, Fontana GP, Jilaihawi H, Kapadia S, Pichard AD, Douglas PS et al. Transcatheter aortic-valve replacement for inoperable severe aortic stenosis. *New England Journal of Medicine*. 2012; 366(18):1696-1704
- 132 Makki N, Maliske S, Blevins A, Girotra S, Cram P. The impact of ultrafiltration in acute decompensated heart failure: A systematic review and meta-analysis [Accepted for publication]. *IJC Metabolic & Endocrine*. 2013;
- 133 Marenzi G, Muratori M, Cosentino ER, Rinaldi ER, Donghi V, Milazzo V et al. Continuous Ultrafiltration for Congestive Heart Failure: The CUORE Trial. *Journal of Cardiac Failure*. 2014; 20(1):9-17
- 134 Masip J, Betbese AJ, Paez J, Vecilla F, Canizares R, Padro J et al. Non-invasive pressure support ventilation versus conventional oxygen therapy in acute cardiogenic pulmonary oedema: a randomized trial. *Lancet*. 2000; 356(9248):2126-2132
- 135 Mauri L, Foster E, Glower DD, Apruzzese P, Massaro JM, Herrmann HC et al. 4-year results of a randomized controlled trial of percutaneous repair versus surgery for mitral regurgitation. *Journal of the American College of Cardiology*. 2013; 62(4):317-328

- 136 Mauri L, Garg P, Massaro JM, Foster E, Glower D, Mehoular P et al. The EVEREST II Trial: design and rationale for a randomized study of the evalve mitralclip system compared with mitral valve surgery for mitral regurgitation. *American Heart Journal*. 2010; 160(1):23-29
- 137 McCullough PA, Nowak RM, McCord J, Hollander JE, Herrmann HC, Steg PG et al. B-type natriuretic peptide and clinical judgment in emergency diagnosis of heart failure: analysis from Breathing Not Properly (BNP) Multinational Study. *Circulation*. 2002; 106(4):416-422
- 138 McMurray J, Hart W, Rhodes G. An evaluation of the cost of heart failure to the National Health Service in the UK. *British Journal of Medical Economics*. 1993; 6:99-110
- 139 McMurray JJ, Adamopoulos S, Anker SD, Auricchio A, Bohm M, Dickstein K et al. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. *European Heart Journal*. 2012; 33(14):1787-1847
- 140 Mealing S, Eaton JN. MitraClip for patients with mitral regurgitation who are ineligible for surgical repair or replacement: A UK based cost-utility analysis. *Journal of the American College of Cardiology*. 2011; 58(20 Suppl.):B212
- 141 Mealing S, Watt M, Eaton J, Sculpher M, Brasseur P, Busca R et al. A United Kingdom-based cost utility analysis of TAVI for inoperable patients with severe Aortic Stenosis treated by medical management. *EuroIntervention*. 2011; 7(Suppl. M):M225
- 142 Mikulic E, Franciosa JA, Cohn JN. Comparative hemodynamic effects of chewable isosorbide dinitrate and nitroglycerin in patients with congestive heart failure. *Circulation*. 1975; 52(3):477-482
- 143 Miller AH, Nazeer S, Pepe P, Estes B, Gorman A, Yancy CW. Acutely decompensated heart failure in a county emergency department: a double-blind randomized controlled comparison of nesiritide versus placebo treatment. *Annals of Emergency Medicine*. 2008; 51(5):571-578
- 144 Miller DC, Blackstone EH, Mack MJ, Svensson LG, Kodali SK, Kapadia S et al. Transcatheter (TAVR) versus surgical (AVR) aortic valve replacement: occurrence, hazard, risk factors, and consequences of neurologic events in the PARTNER trial. *Journal of Thoracic and Cardiovascular Surgery*. 2012; 143(4):832-843
- 145 Moe GW, Howlett J, Januzzi JL, Zowall H. N-terminal pro-B-type natriuretic peptide testing improves the management of patients with suspected acute heart failure: primary results of the Canadian prospective randomized multicenter IMPROVE-CHF study. *Circulation*. 2007; 115(24):3103-3110
- 146 Mueller C, Laule-Kilian K, Schindler C, Klima T, Frana B, Rodriguez D et al. Cost-effectiveness of B-type natriuretic peptide testing in patients with acute dyspnea. *Archives of Internal Medicine*. 2006; 166(10):1081-1087
- 147 Mueller C, Scholer A, Laule-Kilian K, Martina B, Schindler C, Buser P et al. Use of B-type natriuretic peptide in the evaluation and management of acute dyspnea. *New England Journal of Medicine*. 2004; 350(7):647-654

- 148 Mueller T, Gegenhuber A, Poelz W, Haltmayer M. Diagnostic accuracy of B type natriuretic peptide and amino terminal proBNP in the emergency diagnosis of heart failure. *Heart*. 2005; 91(5):606-612
- 149 Mujib M, Patel K, Fonarow GC, Kitzman DW, Zhang Y, Aban IB et al. Angiotensin-converting enzyme inhibitors and outcomes in heart failure and preserved ejection fraction. *American Journal of Medicine*. 2013; 126(5):401-410
- 150 Murphy A, Fenwick E, Toff WD, Neilson MP, Berry C, Uren N et al. Transcatheter aortic valve implantation for severe aortic stenosis: the cost-effectiveness case for inoperable patients in the United Kingdom. *International Journal of Technology Assessment in Health Care*. 2013; 29(1):12-19
- 151 Murphy G and Cunningham J. Percutaneous heart valve replacement for valvular heart disease: a review of the clinical effectiveness, cost-effectiveness, and guidelines. Canadian Agency for Drugs and Technologies in Health (CADTH), 2010
- 152 National Clinical Guideline Centre. Chronic heart failure: the management of chronic heart failure in adults in primary and secondary care. NICE clinical guideline 108. London. National Clinical Guideline Centre, 2010. Available from: <http://guidance.nice.org.uk/CG108/>
- 153 National Institute for Cardiology Outcomes Research (NICOR). Secondary analysis of the National Heart Failure Audits of England and Wales 2009 - 2013 (unpublished data), 2014
- 154 National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. London: National Institute for Health and Clinical Excellence; 2008. Available from: <http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf>
- 155 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/>
- 156 National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal 2013. 2nd edition. London: National Institute for Health and Clinical Excellence; 2013. Available from: <http://publications.nice.org.uk/pmg9>
- 157 Nava S, Carbone G, DiBattista N, Bellone A, Baiardi P, Cosentini R et al. Noninvasive ventilation in cardiogenic pulmonary edema: a multicenter randomized trial. *American Journal of Respiratory and Critical Care Medicine*. 2003; 168(12):1432-1437
- 158 Nazerian P, Vanni S, Zanobetti M, Polidori G, Pepe G, Federico R et al. Diagnostic accuracy of emergency Doppler echocardiography for identification of acute left ventricular heart failure in patients with acute dyspnea: comparison with Boston criteria and N-terminal prohormone brain natriuretic peptide. *Academic Emergency Medicine*. 2010; 17(1):18-26
- 159 Nielsen HHM, Klaaborg KE, Nissen H, Terp K, Mortensen PE, Kjeldsen BJ et al. A prospective, randomised trial of transapical transcatheter aortic valve implantation vs. surgical aortic valve replacement in operable elderly patients with aortic stenosis: the STACCATO trial. *EuroIntervention*. 2012; 8(3):383-389
- 160 Novitsky JA. Controversy and conflict in the treatment of acute decompensated heart failure: limited role for nesiritide. *Pharmacotherapy*. 2007; 27(5):626-632

- 161 O'Rourke MF, Norris RM, Campbell TJ, Chang VP, Sammel NL. Randomized controlled trial of intraaortic balloon counterpulsation in early myocardial infarction with acute heart failure. *American Journal of Cardiology*. 1981; 47(4):815-820
- 162 Ohman EM, Nanas J, Stomel RJ, Leeser MA, Nielsen DWT, O'Dea D et al. Thrombolysis and counterpulsation to improve survival in myocardial infarction complicated by hypotension and suspected cardiogenic shock or heart failure: results of the TACTICS Trial. *Journal of Thrombosis and Thrombolysis*. 2005; 19(1):33-39
- 163 Organisation for Economic Co-operation and Development (OECD). Purchasing power parities (PPP). 2011. Available from: <http://www.oecd.org/std/ppp>
- 164 Orlando R, Pennant M, Rooney S, Khogali S, Bayliss S, Hassan A et al. Cost-effectiveness of transcatheter aortic valve implantation (TAVI) for aortic stenosis in patients who are high risk or contraindicated for surgery: a model-based economic evaluation. *Health Technology Assessment*. 2013; 17(33):1-86
- 165 Papaioannou V, Stakos D, Kouliatsis G, Pneumatikos I. Relation of tricuspid annular displacement and tissue doppler imaging velocities with pulmonary hypertension and length of the weaning process in mechanically ventilated patients with acute heart failure. *Intensive Care Medicine*. 2010; 36(2 Suppl.):S359
- 166 Papaioannou VE, Stakos DA, Dragoumanis CK, Pneumatikos IA. Relation of tricuspid annular displacement and tissue Doppler imaging velocities with duration of weaning in mechanically ventilated patients with acute pulmonary edema. *BMC Cardiovascular Disorders*. 2010; 10:20
- 167 Parab R, Vasudevan A, Brensilver J, Gitler B. Utility of brain natriuretic peptide as a diagnostic tool for congestive heart failure in the elderly. *Critical Pathways in Cardiology*. 2005; 4(3):140-144
- 168 Park M, Lorenzi-Filho G, Feltrim MI, Viecili PR, Sangean MC, Volpe M et al. Oxygen therapy, continuous positive airway pressure, or noninvasive bilevel positive pressure ventilation in the treatment of acute cardiogenic pulmonary edema. *Arquivos Brasileiros De Cardiologia*. 2001; 76(3):221-230
- 169 Park M, Sangean MC, Volpe MS, Feltrim MI, Nozawa E, Leite PF et al. Randomized, prospective trial of oxygen, continuous positive airway pressure, and bilevel positive airway pressure by face mask in acute cardiogenic pulmonary edema. *Critical Care Medicine*. 2004; 32(12):2407-2415
- 170 Parrinello G, Paterna S, Di Pasquale P, Torres D, Mezzero M, Cardillo M et al. Changes in estimating echocardiography pulmonary capillary wedge pressure after hypersaline plus furosemide versus furosemide alone in decompensated heart failure. *Journal of Cardiac Failure*. 2011; 17(4):331-339
- 171 Paterna S, Di Pasquale P, Parrinello G, Amato P, Cardinale A, Follone G et al. Effects of high-dose furosemide and small-volume hypertonic saline solution infusion in comparison with a high dose of furosemide as a bolus, in refractory congestive heart failure. *European Journal of Heart Failure*. 2000; 2(3):305-313
- 172 Paterna S, Pasquale P, Parrinello G, Fornaciari E, Gaudio F, Fasullo S et al. Changes in brain natriuretic peptide levels and bioelectrical impedance measurements after treatment with high-dose furosemide and hypertonic saline solution versus high-dose furosemide alone in

- refractory congestive heart failure: a double-blind study. *Journal of the American College of Cardiology*. 2005; 45(12):1997
- 173 Paterna S, Fasullo S, Di Pasquale P. High-dose torasemide is equivalent to high-dose furosemide with hypertonic saline in the treatment of refractory congestive heart failure. *Clinical Drug Investigation*. 2005; 25(3):165-173
- 174 Peacock WF, Hollander JE, Diercks DB, Lopatin M, Fonarow G, Emerman CL. Morphine and outcomes in acute decompensated heart failure: an ADHERE analysis. *Emergency Medicine Journal*. 2008; 25(4):205-209
- 175 Pitt B, Williams G, Remme W, Martinez F, Lopez-Sendon J, Zannad F et al. The EPHESUS trial: eplerenone in patients with heart failure due to systolic dysfunction complicating acute myocardial infarction. Eplerenone Post-AMI Heart Failure Efficacy and Survival Study. *Cardiovascular Drugs and Therapy*. 2001; 15(1):79-87
- 176 Pitt B, Zannad F, Remme WJ, Cody R, Castaigne A, Perez A et al. The effect of spironolactone on morbidity and mortality in patients with severe heart failure. Randomized Aldactone Evaluation Study Investigators. *New England Journal of Medicine*. 1999; 341(10):709-717
- 177 Pivac N, Rumboldt Z, Sardelic S, Bagatin J, Polic S, Ljutic D et al. Diuretic effects of furosemide infusion versus bolus injection in congestive heart failure. *International Journal of Clinical Pharmacology Research*. 1998; 18(3):121-128
- 178 Plaisance P, Pirracchio R, Berton C, Vicaut E, Payen D. A randomized study of out-of-hospital continuous positive airway pressure for acute cardiogenic pulmonary oedema: physiological and clinical effects. *European Heart Journal*. 2007; 28(23):2895-2901
- 179 Potocki M, Breidthardt T, Reichlin T, Hartwiger S, Morgenthaler NG, Bergmann A et al. Comparison of midregional pro-atrial natriuretic peptide with N-terminal pro-B-type natriuretic peptide in the diagnosis of heart failure. *Journal of Internal Medicine*. 2010; 267(1):119-129
- 180 Prondzinsky R, Lemm H, Swyter M, Wegener N, Unverzagt S, Carter JM et al. Intra-aortic balloon counterpulsation in patients with acute myocardial infarction complicated by cardiogenic shock: the prospective, randomized IABP SHOCK Trial for attenuation of multiorgan dysfunction syndrome. *Critical Care Medicine*. 2010; 38(1):152-160
- 181 Prosen G, Klemen P, Strnad M, Grmec S. Combination of lung ultrasound (a comet-tail sign) and N-terminal pro-brain natriuretic peptide in differentiating acute heart failure from chronic obstructive pulmonary disease and asthma as cause of acute dyspnea in prehospital emergency setting. *Critical Care*. 2011; 15(2):R114
- 182 Purchasing and Supply Agency. Centre for Evidence-based Purchasing. Evidence Review: Ultrafiltration therapy for fluid overload in heart failure. CEP 7016. London. Department of Health., 2007
- 183 Rasanen J, Heikkila J, Downs J. Continuous positive airway pressure by face mask in acute cardiogenic pulmonary edema. *American Journal of Cardiology*. 1985; 55(4):296-300
- 184 Ray P, Arthaud M, Lefort Y, Birolleau S, Beigelman C, Riou B. Usefulness of B-type natriuretic peptide in elderly patients with acute dyspnea. *Intensive Care Medicine*. 2004; 30(12):2230-2236

- 185 Ray P, Arthaud M, Birolleau S, Isnard R, Lefort Y, Boddaert J et al. Comparison of brain natriuretic peptide and probrain natriuretic peptide in the diagnosis of cardiogenic pulmonary edema in patients aged 65 and older. *Journal of the American Geriatrics Society*. 2005; 53(4):643-648
- 186 Reynolds MR, Magnuson EA, Wang K, Thourani VH, Williams M, Zajarias A et al. Health-related quality of life after transcatheter or surgical aortic valve replacement in high-risk patients with severe aortic stenosis: results from the PARTNER (Placement of AoRTic TraNscathetER Valve) Trial (Cohort A). *Journal of the American College of Cardiology*. 2012; 60(6):548-558
- 187 Reynolds MR, Magnuson EA, Lei Y, Leon MB, Smith CR, Svensson LG et al. Health-related quality of life after transcatheter aortic valve replacement in inoperable patients with severe aortic stenosis. *Circulation*. 2011; 124(18):1964-1972
- 188 Reynolds MR, Magnuson EA, Lei Y, Wang K, Vilain K, Li H et al. Cost-effectiveness of transcatheter aortic valve replacement compared with surgical aortic valve replacement in high-risk patients with severe aortic stenosis: results of the PARTNER (Placement of Aortic Transcatheter Valves) trial (Cohort A). *Journal of the American College of Cardiology*. 2012; 60(25):2683-2692
- 189 Reynolds MR, Magnuson EA, Wang K, Lei Y, Vilain K, Walczak J et al. Cost-effectiveness of transcatheter aortic valve replacement compared with standard care among inoperable patients with severe aortic stenosis: results from the placement of aortic transcatheter valves (PARTNER) trial (Cohort B). *Circulation*. 2012; 125(9):1102-1109
- 190 Rose EA, Gelijns AC, Moskowitz AJ, Heitjan DF, Stevenson LW, Dembitsky W et al. Long-term use of a left ventricular assist device for end-stage heart failure. *New England Journal of Medicine*. 2001; 345(20):1435-1443
- 191 Rose EA, Moskowitz AJ, Packer M, Sollano JA, Williams DL, Tierney AR et al. The REMATCH trial: rationale, design, and end points. *Randomized Evaluation of Mechanical Assistance for the Treatment of Congestive Heart Failure*. *Annals of Thoracic Surgery*. 1999; 67(3):723-730
- 192 Rutten JHW, Steyerberg EW, Boomsma F, van Saase JLCM, Deckers JW, Hoogsteden HC et al. N-terminal pro-brain natriuretic peptide testing in the emergency department: beneficial effects on hospitalization, costs, and outcome. *American Heart Journal*. 2008; 156(7):1-7
- 193 Sacchetti A, Ramoska E, Moakes ME, McDermott P, Moyer V. Effect of ED management on ICU use in acute pulmonary edema. *American Journal of Emergency Medicine*. 1999; 17(6):571-574
- 194 Salvador DRK, Rey NR, Ramos GC, Punzalan FER. Continuous infusion versus bolus injection of loop diuretics in congestive heart failure. *Cochrane Database of Systematic Reviews*. 2005;(Issue 3):CD003178
- 195 Sanz MP, Borque L, Rus A, Vicente B, Ramirez Y, Lasa L. Comparison of BNP and NT-proBNP assays in the approach to the emergency diagnosis of acute dyspnea. *Journal of Clinical Laboratory Analysis*. 2006; 20(6):227-232
- 196 Sato N, Kajimoto K, Asai K, Mizuno M, Minami Y, Nagashima M et al. Acute decompensated heart failure syndromes (ATTEND) registry. A prospective observational multicenter cohort study: rationale, design, and preliminary data. *American Heart Journal*. 2010; 159(6):949-955

- 197 Seino Y, Momomura S, Takano T, Hayakawa H, Katoh K. Multicenter, double-blind study of intravenous milrinone for patients with acute heart failure in Japan. Japan Intravenous Milrinone Investigators. *Critical Care Medicine*. 1996; 24(9):1490-1497
- 198 Seronde MF, Gayat E, Logeart D, Lassus J, Laribi S, Boukef R et al. Comparison of the diagnostic and prognostic values of B-type and atrial-type natriuretic peptides in acute heart failure. *International Journal of Cardiology*. 2013; 168(4):3404-3411
- 199 Seyfarth M, Sibbing D, Bauer I, Frohlich G, Bott-Flugel L, Byrne R et al. A randomized clinical trial to evaluate the safety and efficacy of a percutaneous left ventricular assist device versus intra-aortic balloon pumping for treatment of cardiogenic shock caused by myocardial infarction. *Journal of the American College of Cardiology*. 2008; 52(19):1584-1588
- 200 Shah MR, O'Connor CM, Sopko G, Hasselblad V, Califf RM, Stevenson LW. Evaluation Study of Congestive Heart Failure and Pulmonary Artery Catheterization Effectiveness (ESCAPE): design and rationale. *American Heart Journal*. 2001; 141(4):528-535
- 201 Shah RV, Truong QA, Gaggin HK, Pfannkuche J, Hartmann O, Januzzi JLI. Mid-regional pro-atrial natriuretic peptide and pro-adrenomedullin testing for the diagnostic and prognostic evaluation of patients with acute dyspnoea. *European Heart Journal*. 2012; 33(17):2197-2205
- 202 Shaikh K, Ahmad M. Diagnostic significance of NT-proBNP estimation in patients with acute dyspnea. *Journal of the College of Physicians and Surgeons--Pakistan*. 2011; 21(10):584-588
- 203 Sharon A, Shpirer I, Kaluski E, Moshkovitz Y, Milovanov O, Polak R et al. High-dose intravenous isosorbide-dinitrate is safer and better than Bi-PAP ventilation combined with conventional treatment for severe pulmonary edema. *Journal of the American College of Cardiology*. 2000; 36(3):832-837
- 204 Siebert U, Januzzi JL, Beinfeld MT, Cameron R, Gazelle GS. Cost-effectiveness of using N-terminal pro-brain natriuretic peptide to guide the diagnostic assessment and management of dyspneic patients in the emergency department. *American Journal of Cardiology*. 2006; 98(6):800-805
- 205 Smith CR, Leon MB, Mack MJ, Miller DC, Moses JW, Svensson LG et al. Transcatheter versus surgical aortic-valve replacement in high-risk patients. *New England Journal of Medicine*. 2011; 364(23):2187-2198
- 206 Sotomi Y, Iwakura K, Higuchi Y, Fujii K, Sakata Y, Kajimoto K et al. Right heart catheterization in patients with acute decompensated heart failure syndromes: Short-term analysis of attend registry from Japan. *Circulation*. 2012; 126(21 Suppl. 1):A12738
- 207 Sotomi Y, Sato N, Kajimoto K, Sakata S, Mizuno K, Minami Y et al. Impact of pulmonary artery catheter on outcome in patients with acute heart failure syndromes with hypotension or receiving inotropes: from the ATTEND registry [In Press]. *International Journal of Cardiology*. 2014;
- 208 Takeda S, Nejima J, Takano T, Nakanishi K, Takayama M, Sakamoto A et al. Effect of nasal continuous positive airway pressure on pulmonary edema complicating acute myocardial infarction. *Japanese Circulation Journal*. 1998; 62(8):553-558

- 209 Takeda S, Takano T, Ogawa R. The effect of nasal continuous positive airway pressure on plasma endothelin-1 concentrations in patients with severe cardiogenic pulmonary edema. *Anesthesia and Analgesia*. 1997; 84(5):1091-1096
- 210 Thiele H, Sick P, Boudriot E, Diederich KW, Hambrecht R, Niebauer J et al. Randomized comparison of intra-aortic balloon support with a percutaneous left ventricular assist device in patients with revascularized acute myocardial infarction complicated by cardiogenic shock. *European Heart Journal*. 2005; 26(13):1276-1283
- 211 Thiele H, Zeymer U, Neumann FJ, Ferenc M, Olbrich HG, Hausleiter J et al. Intra-aortic balloon counterpulsation in acute myocardial infarction complicated by cardiogenic shock (IABP-SHOCK II): final 12 month results of a randomised, open-label trial. *Lancet*. 2013; 382(9905):1638-1645
- 212 Thiele H, Zeymer U, Neumann FJ, Ferenc M, Olbrich HG, Hausleiter J et al. Intraaortic balloon support for myocardial infarction with cardiogenic shock. *New England Journal of Medicine*. 2012; 367(14):1287-1296
- 213 Thomson MR, Nappi JM, Dunn SP, Hollis IB, Rodgers JE, Van Bakel AB. Continuous versus intermittent infusion of furosemide in acute decompensated heart failure. *Journal of Cardiac Failure*. 2010; 16(3):188-193
- 214 Triposkiadis FK, Butler J, Karayannis G, Starling RC, Filippatos G, Wolski K et al. Efficacy and safety of high dose versus low dose furosemide with or without dopamine infusion: The Dopamine in Acute Decompensated Heart Failure II (DAD-HF II) Trial. *International Journal of Cardiology*. 2014; 172(1):115-121
- 215 Triposkiadis F, Parissis JT, Starling RC, Skoularigis J, Louridas G. Current drugs and medical treatment algorithms in the management of acute decompensated heart failure. *Expert Opinion on Investigational Drugs*. 2009; 18(6):695-707
- 216 Tu JV, Donovan LR, Lee DS, Wang JT, Austin PC, Alter DA et al. Effectiveness of public report cards for improving the quality of cardiac care: the EFFECT study: a randomized trial. *JAMA*. 2009; 302(21):2330-2337
- 217 University of Bristol. QUADAS-2. 2011. Available from: <http://www.bris.ac.uk/quadas/quadas-2/> [Last accessed: 18 February 2014]
- 218 Unverzagt S, Machemer MT, Solms A, Thiele H, Burkhoff D, Seyfarth M et al. Intra-aortic balloon pump counterpulsation (IABP) for myocardial infarction complicated by cardiogenic shock. *Cochrane Database of Systematic Reviews*. 2011;(Issue 7):CD007398
- 219 Villacorta H, Duarte A, Duarte NM, Carrano A, Mesquita ET, Dohmann HJF et al. The role of B-type natriuretic peptide in the diagnosis of congestive heart failure in patients presenting to an emergency department with dyspnea. *Arquivos Brasileiros De Cardiologia*. 2002; 79(6):569-8
- 220 Vital FM, Ladeira MT, Atallah AN. Non-invasive positive pressure ventilation (CPAP or bilevel NPPV) for cardiogenic pulmonary oedema. *Cochrane Database of Systematic Reviews*. 2013;(Issue 3):CD005351
- 221 Vital FM, Saconato H, Ladeira MT, Sen A, Hawkes CA, Soares B et al. Non-invasive positive pressure ventilation (CPAP or bilevel NPPV) for cardiogenic pulmonary edema. *Cochrane Database of Systematic Reviews*. 2008;(Issue 3):CD005351

- 222 Wang CS, FitzGerald JM, Schulzer M, Mak E, Ayas NT. Does this dyspneic patient in the emergency department have congestive heart failure? *JAMA*. 2005; 294(15):1944-1956
- 223 Wang HK, Tsai MS, Chang JH, Wang TD, Chen WJ, Huang CH. Cardiac ultrasound helps for differentiating the causes of acute dyspnea with available B-type natriuretic peptide tests. *American Journal of Emergency Medicine*. 2010; 28(9):987-993
- 224 Watt M, Mealing S, Eaton J, Piazza N, Moat N, Brasseur P et al. Cost-effectiveness of transcatheter aortic valve replacement in patients ineligible for conventional aortic valve replacement. *Heart*. 2012; 98(5):370-376
- 225 Watt M, Mealing S, Sculpher M, Eaton JN. Estimating the financial implications to the UK NHS of introducing corevalve alongside medical management for the treatment of severe aortic stenosis. *Value in Health*. 2010; 13(7):A346
- 226 Weitz G, Struck J, Zonak A, Balnus S, Perras B, Dodt C. Prehospital noninvasive pressure support ventilation for acute cardiogenic pulmonary edema. *European Journal of Emergency Medicine*. 2007; 14(5):276-279
- 227 Wen H, Zhang Y, Zhu J, Lan Y, Yang H. Ultrafiltration versus Intravenous Diuretic Therapy to Treat Acute Heart Failure: A Systematic Review. *American Journal of Cardiovascular Drugs*. 2013; 13(5):365-373
- 228 Weng CL, Zhao YT, Liu QH, Fu CJ, Sun F, Ma YL et al. Meta-analysis: Noninvasive ventilation in acute cardiogenic pulmonary edema. *Annals of Internal Medicine*. 2010; 152(9):590-600
- 229 Wu MY, Chang NC, Su CL, Hsu YH, Chen TW, Lin YF et al. Loop diuretic strategies in patients with acute decompensated heart failure: A meta-analysis of randomized controlled trials. *Journal of Critical Care*. 2014; 29(1):2-9
- 230 Wu YC, Zhang JF, Shen WF, Zhao Q. Transcatheter aortic valve implantation versus surgical aortic valve replacement for severe aortic stenosis: a meta analysis. *Chinese Medical Journal*. 2013; 126(6):1171-1177
- 231 Young JB. Intravenous nesiritide vs nitroglycerin for treatment of decompensated congestive heart failure: A randomized controlled trial. *JAMA*. 2002; 287(12):1531-1540
- 232 Zaninotto M, Mion M, Altinier S, Pastorello M, Rocco S, Tosato F et al. NT-proBNP in the differential diagnosis of acute dyspnea in the emergency department. *Clinical Biochemistry*. 2005; 38(11):1041-1044
- 233 Zhi Q, Liang JC. Diuretics and ultrafiltration in acute heart failure syndrome. *International Heart Journal*. 2013; 54(6):390-394
- 234 Zion MM, Balkin J, Rosenmann D, Goldbourt U, Reicher-Reiss H, Kaplinsky E et al. Use of pulmonary artery catheters in patients with acute myocardial infarction. Analysis of experience in 5,841 patients in the SPRINT Registry. *SPRINT Study Group. Chest*. 1990; 98(6):1331-1335

13 Glossary

Acronym/Term	Description
6MWT	6 minute walk test – an evaluation of exercise capacity
Abstract	Summary of a study, which may be published alone or as an introduction to a full scientific paper.
Acid-base balance	Serum bicarbonate, serum lactate and arterial blood gases.
Acute heart failure (AHF)	Rapid onset of a clinical syndrome where the heart is unable to pump adequate blood to provide for the needs of the body
Aldosterone antagonist (AA)	Also known as Mineralocorticosteroid antagonist or Mineralocorticosteroid receptor blocker. A diuretic drug that antagonises the action of aldosterone at mineralocorticoid receptors.
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.
AMI	Acute myocardial infarction (damage to the heart muscle usually due to blockage of a blood vessel supplying it)
Angiotensin receptor blocker (ARB)	Treatment for high blood pressure and heart failure.
Angiotensin-converting enzyme inhibitors (ACEI)	Treatment for high blood pressure and heart failure.
Aortic stenosis	Restriction in the opening of the aortic valve
Applicability	The degree to which the results of an observation, study or review are likely to hold true in a particular clinical practice setting.
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
Arterial blood gases	Measurement of oxygen, carbon dioxide and other parameters from an arterial blood sample
Association	Statistical relationship between 2 or more events, characteristics or other variables. The relationship may or may not be causal.
Atrial fibrillation (AF)	Irregularly irregular rhythm of the heart.
Atrial natriuretic peptide (ANP)	A hormone released by heart muscle cells
Base case analysis	In a modelling, the base case is the primary analysis based on the best estimates of each model input. (c.f. 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.

Baseline risk	The probability of an event (e.g. death) occurring in the comparator arm. This is a term used in modelling, where the baseline risk from one data source might be combined with a risk ratio from another source to estimate the probability of an event occurring for patients receiving a different intervention.
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 blocker (BB)	Treatment for heart rhythm, angina and heart attacks, high blood pressure and heart failure.
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.
Bilevel ventilation (BiPAP)	A form of non-invasive ventilation used to support people with respiratory failure
Bivariate method	Any statistical method that analyses two different variables. In the context of this guideline it refers to the method employed that was used in diagnostic meta-analysis. The two variables it is referring to are sensitivity and specificity which are modelled together.
Blinding	A way to prevent researchers, doctors and patients in a clinical trial from knowing which study group each patient is in so they cannot influence the results. The best way to do this is by sorting patients into study groups randomly. The purpose of 'blinding' or 'masking' is to protect against bias. A single-blinded study is one in which patients do not know which study group they are in (for example whether they are taking the experimental drug or a placebo). A double-blinded study is one in which neither patients nor the researchers/doctors know which study group the patients are in. A triple blind study is one in which neither the patients, clinicians or the people carrying out the statistical analysis know which treatment patients received.
BNF	British national formulary.
BNP	B-type natriuretic peptide (a protein substance secreted from the heart wall especially when stretched or when the pressure within it has risen).
BP	Blood pressure.
B-type natriuretic peptide	A hormone released by heart muscle cells
CABG	Coronary artery bypass grafting.
Cardiac resynchronisation therapy (CRT)	A form of pacing of the heart, whereby both pumping chambers as well as the right filling chamber are paced. This improves the timing and efficiency of the pumping by the heart.
Cardiogenic shock	A severe form of heart failure characterised by sustained inadequate blood

	flow to the whole body
Cardiomyopathy (CM)	A condition that has several forms. They are characterised by disease processes that primarily affect the heart muscle.
Carer (caregiver)	Someone who looks after family, partners or friends in need of help because they are ill, frail or have a disability.
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.
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.
Chronic heart failure (CHF)	A long term condition where the heart is unable to pump adequate blood to meet the demands of the body
Clinical efficacy	The extent to which an intervention is active when studied under controlled research conditions.
Clinician	A healthcare professional who provides patient care. For example, a doctor, nurse or physiotherapist.
Cochrane Review	The Cochrane Library consists of a regularly updated collection of evidence-based medicine databases including the Cochrane Database of Systematic Reviews (reviews of randomised controlled trials prepared by the Cochrane Collaboration).
Cohort study	A study with 2 or more groups of people – cohorts – with similar characteristics. One group receives a treatment, is exposed to a risk factor or has a particular symptom and the other group does not. The study follows their progress over time and records what happens. See also observational study.
Comorbidity	A disease or condition that someone has in addition to the health problem being studied or treated.
Comparability	Similarity of the groups in characteristics likely to affect the study results (such as health status or age).
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

	<p>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.</p> <p>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.</p> <p>A wide confidence interval indicates a lack of certainty about the true effect of the test or treatment - often because a small group of patients has been studied. A narrow confidence interval indicates a more precise estimate (for example, if a large number of patients have been studied).</p>
Confounding factor	<p>Something that influences a study and can result in misleading findings if it is not understood or appropriately dealt with.</p> <p>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.</p>
Consensus methods	<p>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.</p>
Conservative assumption	<p>Where there is uncertainty modellers may have a choice of which value to give to a model input. A conservative assumption is where the modeller chooses the parameter in such a way that it cannot bias in favour of the new treatment (and is likely to be biasing in favour of the standard treatment).</p>
Continuous positive airway pressure (CPAP)	<p>A form of non-invasive ventilation used to support people with respiratory failure</p>
Control group	<p>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.</p> <p>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.</p>
COPD	<p>Chronic obstructive pulmonary disease (A condition that affects the lungs and the airways, characterised by breathlessness, wheeze and cough)</p>
Coronary heart disease (CHD)	<p>Narrowing of the arteries that supply blood to the heart muscle by the build-up of fatty deposits</p>
Cost of illness analysis	<p>A non-comparative study which estimates the cost per year associated with a particular disease. Such an analysis might include the cost of time off work as well as direct medical costs.</p>
Cost-benefit analysis (CBA)	<p>Cost-benefit analysis is one of the tools used to carry out an economic</p>

	evaluation. The costs and benefits are measured using the same monetary units (for example, pounds sterling) to see whether the benefits exceed the costs.
Cost–consequences analysis (CCA)	Cost-consequence analysis is one of the tools used to carry out an economic evaluation. This compares the costs (such as treatment and hospital care) and the consequences (such as health outcomes) of a test or treatment with a suitable alternative. Unlike cost-benefit analysis or cost-effectiveness analysis, it does not attempt to summarise outcomes in a single measure (like the quality-adjusted life year) or in financial terms. Instead, outcomes are shown in their natural units (some of which may be monetary) and it is left to decision-makers to determine whether, overall, the treatment is worth carrying out.
Cost-effective	Good value for money - that is sufficient additional (health) gains achieved relative to the additional cost incurred
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-effectiveness plane	A graph used to present results of cost-effectiveness analyses where incremental costs are plotted against incremental health effects (e.g. QALYs gained).
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.
CV mortality	Cardiovascular mortality (Death caused by disease of the heart and the blood vessels)
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.
Discount rate	The rate per year at which future costs and outcomes are discounted – see discounting. This has been set by the Treasury at 3.5% for economic evaluations, reflecting long-term interest rates. So a cost of £103.50 next year is valued today at £100.
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.
Distress	A state of extreme anxiety, pain or sorrow
Disutility	The reduction in utility attributed to experiencing a clinical event or health

	state.
Diuretics	Medications which act to increase urine output
Diuretic resistance	Dose escalation beyond a patient's previously recognised dose ceiling or a dose approaching the maximum recommended daily dose without incremental improvement in diuresis
DM	Diabetes mellitus
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.
Dyspnoea	Shortness of breath
ECG	Electrocardiogram (Recording of the electrical activity of the heart)
Echocardiography	An ultrasound examination of the heart
Economic evaluation	<p>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.</p> <p>There are several types of economic evaluation: cost-benefit analysis, cost-consequence analysis, cost-effectiveness analysis, cost-minimisation analysis and cost-utility analysis. They use similar methods to define and evaluate costs, but differ in the way they estimate the benefits of a particular drug, programme or intervention.</p>
Effect (as in effect measure, treatment effect, estimate of effect, effect size)	<p>A measure that shows the magnitude of the outcome in one group compared with that in a control group.</p> <p>For example, if the absolute risk reduction is shown to be 5% and it is the outcome of interest, the effect size is 5%.</p> <p>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).</p>
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-5D)	A standardised instrument used to measure a health outcome. It provides a single index value for health status.
ER	Emergency room
ESC	European society of cardiology

Estimated glomerular filtration rate (eGFR)	A measure of the function of the kidneys, reflecting the volume of blood that is liable to be cleared by the kidney per minute. The lower the number the worse is the function of the kidneys.
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 (clinical study)	Criteria that define who is not eligible to participate in a clinical study.
Exclusion criteria (literature review)	Explicit standards used to decide which studies should be excluded from consideration as potential sources of evidence.
Extended dominance	If Option A is both more clinically effective than Option B and has a lower cost per unit of effect, when both are compared with a do-nothing alternative then Option A is said to have extended dominance over Option B. Option A is therefore more efficient 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.
GDG	Guideline development group. Multiprofessional group responsible for developing this guideline
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.
GP	General Practitioner.
GPP	Good practice point.
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.
Haemodynamic	Relating to the circulation of the blood, usually describes the mechanical effects of the circulatory system such as the pressure in a chamber or vessel.
Harms	Adverse effects of an intervention.
Health economics	Study or analysis of the cost of using and distributing healthcare resources.
Health Technology Assessment (HTA)	An evaluation exploring clinical and cost effectiveness and other related issues, for example organisational implications, of a health technology (e.g., drug, medical device, clinical or surgical procedure).
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.
Heart failure with preserved ejection fraction (HFPEF)	A form of heart failure associated with preserved [good] contraction of the heart muscle.

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.
HF	Heart failure.
HRQoL	Health-related quality of life
Hypercapnic	Levels of carbon dioxide in the blood above the normal range
Hypertrophic cardiomyopathy	A form of heart muscle abnormality, frequently characterised by an unexplained increase in the thickness of the heart muscle due to a genetic abnormality.
Implantable cardioverter defibrillator (ICD)	A type of pacemaker capable of delivering an electrical shock inside the heart, to stop a lethal rhythm abnormality.
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 monetary benefit INMB)	The value, in monetary terms, of an intervention net of its cost compared with a comparator intervention. The INMB can be calculated for a given cost-effectiveness (willingness to pay) threshold. If the threshold is £20,000 per QALY gained then the INMB is calculated as: (£20,000 x QALYs gained) – Incremental cost.
Indirectness	The available evidence is different to the review question being addressed, in terms of PICO (population, intervention, comparison and outcome).
(Positive) Inotrope	A medication which increases the force of heart muscle contraction
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.
International normalised ratio (INR)	A measure of how thinned the blood is, in comparison to normal, as a result of blood thinning medication
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.
Intra-aortic balloon counterpulsation (IABP)	A device placed via the femoral artery with a balloon that sits within the descending aorta that inflates and deflates at intervals timed with the cardiac cycle.
Intraoperative	The period of time during a surgical procedure.
Intravenous	The administration of medication in to a vein
Intubation rate	The number of people requiring intubation of the airway and mechanical invasive ventilation
Invasive Mechanical ventilation	The use of a breathing machine and tube in to the persons wind pipe in order to treat respiratory failure
IQR	Inter-quartile range
Ischaemic heart disease (IHD)	Disease of the heart caused by insufficient blood supply to the heart.
ISDN+Hyd	Isosorbide dinitrate and hydralazine.
- isosorbide dinitrate	A vasodilator medication
Isovolumic relaxation time (IVRT)	A short period in the cycle of the heart where the heart muscle is relaxing, but the amount of blood in the pumping chamber is not changing.
ISWT	Incremental Shuttle Walk Test. A field test of functional capacity or exercise tolerance
Jugular venous pressure (JVP)	A measure of the pressure in the neck veins, assessed by the height of distended vein in the neck of the patient who is propped up at 45 degrees.
K+	Potassium (One of the essential salts for the function of the body)
Kappa statistic	A statistical measure of inter-rater agreement that takes into account the agreement occurring by chance.
Left ventricular (LV)	Refers to the left pumping chamber of the heart.
Left ventricular assist devices (LVADs)	Sophisticated device, implanted surgically to help a badly failing heart, to pump blood into the circulation.
Left ventricular assist devices (LVADs)	A mechanical device aimed at assisting the function of the left ventricle
Left ventricular ejection fraction (LVEF)	The percentage of the volume of the blood that leaves the heart with each beat, this is a measure of the pumping function of the left pumping chamber of the heart.
Left ventricular systolic dysfunction (LVSD)	The condition where the left pumping chamber's ability to pump is impaired. This is characterised by low left ventricular ejection fraction, and leads to heart failure.
Length of stay	The total number of days a participant stays in hospital.
Level 2 care	Patients requiring more detailed observation or intervention, including support for a single failing organ system or post-operative care and those stepping down from higher levels of care
Licence	See 'Product licence'.
Life-years	The average years of remaining life expectancy. The life-years gained are the extra years of life attributable to one treatment compared with an alternative.

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.
Loss to follow-up	People in a clinical trial who do not finish the complete observation period
LYG	Life year gained.
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.
MICE	Male, history of myocardial infarction, crepitations, ankle oedema
Mid-regional Pro-atrial Natriuretic Peptide (MR-proANP)	A hormone released by the heart muscle cells
Minimal important difference (MID)	The smallest difference in score in the outcome of interest that informed patients or informed proxies perceive as important, either beneficial or harmful, and that would lead the patient or clinician to consider a change in the management.
Minnesota living with heart failure questionnaire (MLHF/MLWHFQ)	Measures the effect of heart failure and treatment for heart failure on an individual's quality of life.
Mitral regurgitation	Failure of function of the mitral valve resulting in blood leaking back towards the left atrium during the pumping phase of the left ventricle.
Model	A model represents the essential aspects of a complex system in a usable form. Modelling is usually conducted when simply observing the outcomes in a controlled setting is not feasible. A decision model uses data often from different sources to quantify specific outcomes with one course of action compared with another.
Multivariate model	A statistical model for analysis of the relationship between 2 or more predictor (independent) variables and the outcome (dependent) variable.
Myocardial infarction (MI)	Heart attack
Myocardial ischaemia	Insufficient blood supply to the muscle of the heart.
Myocardium	Heart muscle
National Service Framework (NSF)	Policies set out by the National Health Service to clearly define standards of care for major medical issues.
Natriuretic peptide (NP)	A protein substance secreted by the wall of the heart when it is stretched or under increased pressure. It has several forms.
NCC	National Collaborating Centre.
NCGC	National Clinical Guideline Centre.
Negative predictive value	In screening or diagnostic tests: A measure of the usefulness of a screening

Non-invasive positive pressure ventilation	A form of non-invasive respiratory support
(NPV)	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>
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence.
Nitrates	A type of medication use to dilate blood vessels
Nitroprusside	A medication used to dilate blood vessels
NR	Not reported
N-terminal atrial natriuretic peptide (Nt-proANP)	A hormone released by heart muscle cells
N-terminal pro-B-type natriuretic peptide (NT-proBNP)	One of the natriuretic peptides, protein substances secreted by the wall of the heart when it is stretched or under increased pressure. It has several forms.
Number needed to treat (NNT)	<p>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.</p> <p>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.</p>
NYHA	New York Heart Association (functional classification): (These allow an assessment of the patient's ability to carry out exercise before they develop their symptoms).
Observational study	<p>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.</p> <p>There is a greater risk of selection bias than in experimental studies.</p>
Odds ratio	<p>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.</p> <p>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.</p> <p>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</p>

	reference category. Odds ratios would be worked out for occasional smokers compared with non-smokers and for regular smokers compared with non-smokers. See also confidence interval, relative risk, risk ratio.
Opportunity cost	The loss of other health care programmes displaced by investment in or introduction of another intervention. This may be best measured by the health benefits that could have been achieved had the money been spent on the next best alternative healthcare intervention.
Outcome	The impact that a test, treatment, policy, programme or other intervention has on a person, group or population. Outcomes from interventions to improve the public's health could include changes in knowledge and behaviour related to health, societal changes (for example, a reduction in crime rates) and a change in people's health and wellbeing or health status. In clinical terms, outcomes could include the number of patients who fully recover from an illness or the number of hospital admissions, and an improvement or deterioration in someone's health, functional ability, symptoms or situation. Researchers should decide what outcomes to measure before a study begins.
P value	<p>The p value is a statistical measure that indicates whether or not an effect is statistically significant.</p> <p>For example, if a study comparing 2 treatments found that one seems more effective than the other, the p value is the probability of obtaining these results by chance. By convention, if the p value is below 0.05 (that is, there is less than a 5% probability that the results occurred by chance) it is considered that there probably is a real difference between treatments. If the p value is 0.001 or less (less than a 1% probability that the results occurred by chance), the result is seen as highly significant.</p> <p>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.</p>
Paroxysmal nocturnal dyspnoea (PND)	Episodes of waking up suddenly with breathlessness.
PCT	Primary Care Trust.
Percutaneous	A procedure performed via the skin in a minimally invasive way
Perioperative	The period from admission through surgery until discharge, encompassing the pre-operative and post-operative periods.
Perspective	<p>In economic evaluation the perspective is the body, whose costs and outcomes are accounted for in the model. In NICE guidelines, costs are measured from an NHS and personal social services perspective.</p> <p>Alternatively, some studies take a broader societal perspective, taking all costs into account.</p>
PICO	Population, Intervention, Comparison and Outcome.
Placebo	A fake (or dummy) treatment given to participants in the control group of a clinical trial. It is indistinguishable from the actual treatment (which is given to participants in the experimental group). The aim is to determine what effect the experimental treatment has had - over and above any placebo effect caused because someone has received (or thinks they have received)

	care or attention.
Polypharmacy	The use or prescription of multiple medications.
Positive predictive value (PPV)	In screening or diagnostic tests: A measure of the usefulness of a screening or diagnostic test. It is the proportion of those with a positive test result who have the disease, and can be interpreted as the probability that a positive test result is correct. It is calculated as follows: <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.
PPIP	Patient and Public Involvement Programme.
PPP	Purchasing Power Parity.
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.
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 modelling, this is where distributions are applied to each model parameter instead of point estimates. This allows us to consider the uncertainty around the model results. This is also known as probabilistic sensitivity analysis.
Probabilistic sensitivity analysis	See probabilistic 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 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.

Pulmonary oedema	The accumulation of fluid in the lungs
Purchasing Power Parity	Rate of currency conversion that reflects the prices of the same good or service in different countries
P-value	The probability that an observed difference could have occurred by chance, assuming that there is in fact no underlying difference between the means of the observations. If the probability is less than 1 in 20, the P value is less than 0.05; a result with a P value of less than 0.05 is conventionally considered to be 'statistically significant'.
PWD	Pulsed wave Doppler (one of the tools to assess the speed of movement by ultrasound. It has important applications in the assessment of the heart valves and heart muscle function)
QUADAS	Quality Assessment of Diagnostic Studies. A 14-item tool used to assess the quality of diagnostic accuracy studies.
Quality of life	See 'Health-related quality of life'.
Quality-adjusted life year (QALY)	<p>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.</p> <p>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.</p>
Randomisation	Assigning participants in a research study to different groups without taking any similarities or differences between them into account. For example, it could involve using a random numbers table or a computer-generated random sequence. It means that each individual (or each group in the case of cluster randomisation) has the same chance of receiving each intervention.
Randomised controlled trial (RCT)	A study in which a number of similar people are randomly assigned to 2 (or more) groups to test a specific drug or treatment. One group (the experimental group) receives the treatment being tested, the other (the comparison or control group) receives an alternative treatment, a dummy treatment (placebo) or no treatment at all. The groups are followed up to see how effective the experimental treatment was. Outcomes are measured at specific times and any difference in response between the groups is assessed statistically. This method is also used to reduce bias.
RCT	See 'Randomised controlled trial'.
Receiver operated characteristic (ROC) curve	A graphical method of assessing the accuracy of a diagnostic test. Sensitivity is plotted against 1 minus specificity. A perfect test will have a positive, vertical linear slope starting at the origin. A good test will be somewhere close to this ideal.
Reference standard	The test that is considered to be the best available method to establish the presence or absence of the outcome – this may not be the one that is routinely used in practice.
Rehabilitation	Process to assist patients to achieve optimal function. May include a period of exercise training.

Rehospitalisation	Requiring to go back to hospital once initially discharged
Relative risk (RR)	<p>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).</p> <p>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.</p>
Relative risk reduction	The proportional reduction in risk in one treatment group compared to another. It is one minus the risk ratio.
Renal function	The function of the kidneys
Reporting bias	See 'Publication bias'.
Resource implication	The likely impact in terms of finance, workforce or other NHS resources.
Respiratory distress	Extreme anxiety due to failure of the lungs
Retrospective study	A research study that focuses on the past and present. The study examines past exposure to suspected risk factors for the disease or condition. Unlike prospective studies, it does not cover events that occur after the study group is selected.
Review question	In guideline development, this term refers to the questions about treatment and care that are formulated to guide the development of evidence-based recommendations.
Risk ratio	See Relative risk.
SD	Standard deviation.
SE	Standard error.
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	<p>Selection bias occurs if:</p> <ul style="list-style-type: none"> 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	<p>How well a test detects the thing it is testing for.</p> <p>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').</p> <p>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.</p> <p>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</p>

	<p>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').</p> <p>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.</p>
Sensitivity analysis	<p>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.</p> <p>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.</p> <p>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.</p> <p>Threshold sensitivity analysis: the critical value of parameters above or below which the conclusions of the study will change are identified.</p> <p>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).</p>
Serious adverse events	Any undesirable experience associated with an intervention being investigated in a clinical trial
Serum bicarbonate	The level of bicarbonate within the blood
Serum lactate	The level of lactate within the blood
Significance (statistical)	A result is deemed statistically significant if the probability of the result occurring by chance is less than 1 in 20 ($p < 0.05$).
Specialist	<p>The term "specialist" is applicable to a wide range of healthcare professionals; however within the context of this guideline, the term specialist is used in relation to establishing the diagnosis of heart failure through non-invasive procedures and to taking the decisions on the management of the heart failure syndrome and its multiple causes.</p> <p>Throughout this guideline the term "specialist" denotes a physician with sub-specialty interest in heart failure (often a consultant cardiologist) who leads a specialist multidisciplinary heart failure team of professionals with appropriate competencies from primary and secondary care. The team will involve, where necessary, other services (such as rehabilitation, tertiary care and palliative care) in the care of individual patients.</p> <p>Unless otherwise specified, within this guideline specialist assessment or management refers to assessment or management by this specialist multidisciplinary heart failure team. The team will decide who is the most appropriate team member to address a particular clinical problem.</p>
Specialist heart failure service	The provision of care for people with heart failure by a specific and expert multidisciplinary team

Specialist management unit	The area where specialist care is provided
Specificity	<p>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.</p> <p>See related term 'Sensitivity'.</p> <p>In terms of literature searching a highly specific search is generally narrow and aimed at picking up the key papers in a field and avoiding a wide range of papers.</p>
Stable	Where the medical condition of the person is not deteriorating
Stakeholder	<p>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:</p> <ul style="list-style-type: none"> • manufacturers of drugs or equipment • national patient and carer organisations • NHS organisations • organisations representing healthcare professionals.
Subcutaneous	Under the skin
Surgical aortic valve replacement	Replacement of the aortic valve via a surgical 'open heart' approach.
Symptomatic hypotension	Low blood pressure which causes symptoms
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.
Systemic hypertension	High blood pressure
Transcatheter aortic valve implantation (TAVI)	The insertion of an artificial valve in place of the aortic valve using a minimally invasive approach
Transoesophageal	Via the food pipe
UK	United Kingdom
Ultrafiltration	A technique to remove excess water and salt from the body by filtering the blood
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).
Vasodilators	Medication to dilate blood vessels
Vasopressor	Medication to constrict blood vessels
Venodilatation	The enlargement or dilation of veins
Venous return	Blood flow returning to the heart via the veins
Venous tone	The balance of dilation and constriction of the veins

Ventricular fibrillation	A type of serious heart rhythm characterized by very rapid, irregular, uncoordinated electrical activity of the pumping chambers with no pumping effect, it is fatal if not corrected immediately
Ventricular tachycardia (VT)	Ventricular tachycardia - A type of serious heart rhythm problem arising in the ventricles resulting in (usually) very rapid contraction of the ventricles.
Willingness to pay (WTP)	How much a group of people or institution would be prepared to pay to receive a certain outcome. For example, we sometimes consider the theoretical willingness to pay for a QALY to be between £20,000

