Final Protocol 03 May 2019

HTA Reference Number: 128897

1. Title of the project

The NephroCheck test and NGAL assays for the assessment of acute kidney injury in critically ill patients.

2. Name of External Assessment Group (EAG) and project lead

Aberdeen Health Technology Assessment Group
Miriam Brazzelli
Senior Research Fellow
Health Services Research Unit
2nd Floor
University of Aberdeen
Health Sciences Building
Foresterhill
Aberdeen AB25 2ZD
Tel: 01224 438082
Email: m.brazzelli@abdn.ac.uk

Reserve contact
Mari Imamura
Research Fellow
Health Services Research Unit
2nd Floor
University of Aberdeen
Health Sciences Building
Foresterhill
Aberdeen AB25 2ZD
Tel: 01224 438121
Email: m.imamura@abdn.ac.uk

3. Plain English summary

The kidneys are filters in the human body. They remove waste products (toxins) and excess fluid from the blood to form urine. Acute kidney injury (AKI) is a common and serious complication that typically occurs in the context of an acute severe illness. AKI often occurs after surgery or among people who are very ill and are considered for critical care or admitted to critical care in the hospital. AKI happens when kidney function suddenly declines. The consequent build-up of toxins sometimes also prevents other organs from working as well as they should. When the kidneys do not work properly, doctors may consider a blood purification technique called renal replacement therapy, or 'dialysis'. Without timely treatment, it can lead to long-term kidney disease, or death.

AKI can occur for many different reasons. It is usually caused by reduced blood flow to the kidney because of another illness, damage to the kidney itself potentially because of a reaction to some drugs, the contrast dye used in some procedures, severe infections, or a blockage preventing drainage from the kidneys. To pre-empt or avoid lasting consequences for people with AKI, early detection may be beneficial. The diagnosis of AKI is a clinical one made by a healthcare provider, but is usually made based on evidence of a build-up of toxins in the blood or a reduction in urine output. One such toxin is 'creatinine', a waste product produced by the muscles and filtered by the kidneys to eliminate it from the body. An elevated level of creatinine usually means impaired kidney function. A rise in creatinine thought to have occurred suddenly within the past week is the main element for defining AKI.

Despite its widespread use in the monitoring of kidney health and disease, creatinine is not a perfect indicator of kidney function. This is because when kidney function suddenly falls (AKI), creatinine levels in the blood may take hours or even days to rise to the level specified by current international rules. Rises in creatinine levels in the blood may also occur even when no damage to the kidney has occurred. These problems related to the use of creatinine levels have led to the search for novel biomarkers that may detect kidney damage earlier or more reliably.

Biomarkers are specific cells, genes, molecules and other characteristics that appear in a particular health condition and can therefore be used as an indicator of the condition. Emerging biomarker tests for AKI include the NGAL test (neutrophil gelatinase-associated lipocalin), which can be measured using a sample of urine or blood, and the NephroCheck test, which measures a combination of two protein biomarkers (TIMP-2 and IGFBP7) in the urine. These novel biomarkers have been developed to identify early damage or stress in the kidneys. If the reliable use of these biomarkers is demonstrated, they may enable earlier identification of AKI and, therefore, early treatment of patients with a modifiable disease course.

AKI is serious condition and the care of people with AKI is costly. Novel biomarkers may have the potential to generate health benefits for people with kidney disease. They may be a cost-effective use of NHS resources, if their use demonstrated to improve the disease course for people with AKI by identifying their condition early in a reliable way.

The purpose of this assessment is to bring together existing evidence to evaluate whether novel biomarkers for early detection of AKI in critically ill people who are admitted to critical care or considered for critical care are effective and whether they represent a good value for money and efficient use of NHS resources (cost-effectiveness). We will assess cost-effectiveness by comparing costs (e.g. test cost, treatment cost) and benefits (e.g. patient survival and quality of life) of the novel biomarkers to determine the best use of NHS resources and inform clinical practice and policy.

4. Decision problem

4.1 Purpose of the decision to be made

Acute kidney injury (AKI) is a common and serious complication that typically occurs in the context of an acute critical illness, especially in intensive and post-surgical care. It is associated with prolonged hospital stay, severe morbidity and increased mortality.[1, 2] Delayed identification of AKI contributes to worse outcomes.[3]

To pre-empt or avoid lasting consequences of AKI, early detection may be beneficial. Traditionally, AKI diagnosis relies on a rise in serum creatinine levels and fall in urine output. Despite its widespread use in the monitoring of kidney health and disease, creatinine is an imperfect marker of kidney function.[4] In a steady state (normal circumstances), the level of creatinine in the blood also depends on the total body muscle mass, which varies between individual people. Some creatinine is also eliminated from the body by mechanisms other than filtering by the kidneys, which can be influenced by a variety of medications, including some commonly used antibiotics. In an illness where kidney function suddenly falls (AKI), creatinine levels in the blood start to rise. While this may start to happen straight away, it can take hours or sometimes days for the level to rise to be sufficient for AKI to be evidence based on current international rules. In addition, in response to stress or even kidney damage, the kidneys have reserve capacity and can compensate so that kidney function is maintained. For this reason, in some clinical settings significant kidney damage can occur without AKI being apparent from changes in blood creatinine. In other settings, such as during a reduction in blood flow to kidneys, rises in creatinine and a reduction in urine can occur, even when no damage has occurred. These drawbacks of our current use of creatinine have led to the search for novel biomarkers that may detect kidney damage or kidney stress earlier or more reliably.

Emerging biomarker tests for AKI include the NGAL test (neutrophil gelatinase-associated lipocalin), which can be measured using a sample of urine or blood.[5] NGAL is released from neutrophils and is induced by inflammation, indicating tubular injury.[4] One limitation of NGAL is that it is produced throughout the body making it difficult to distinguish systemic inflammation from localised renal inflammation.[4] Novel NGAL tests include the ARCHITECT Urine NGAL test (Abbott), the NGAL plasma test (BioPorto Diagnostics) and the NGAL urine test (BioPorto Diagnostics).

Another emerging test for AKI is 'NephroCheck', a combination of two urinary biomarkers, the tissue inhibitor of metalloproteinase-2 (TIMP-2) and the insulin-like growth factor-binding protein 7 (IGFBP-7). Both TIMP-2 and IGFBP7 are cell-cycle arrest proteins that are released into urine as markers of cellular stress in the early phase of tubular cell injury due to a variety of insults (e.g., toxins, drugs, oxidative stress and inflammation), which lead to AKI.[6] The US Food and Drug Administration has approved these combined biomarkers to assess the risk of AKI in critically ill patients.[4]

These novel biomarkers have been developed to detect early damage or stress in the kidneys. If reliable use of these biomarkers can be demonstrated, they may enable earlier identification of AKI, and, therefore, early management of those with a modifiable disease course - with potential for downstream benefits in patients' clinical outcomes.

The purpose of this assessment is to review current evidence on the diagnostic accuracy, predictive validation, impact on clinical outcomes and cost-effectiveness of novel biomarkers (NephroCheck and NGAL assays) for the assessment of AKI in critically ill patients.

If demonstrated, the predictive ability of these novel biomarkers for early detection of AKI would have the potential of improving current AKI management by adopting measures that could prevent progression to more severe kidney injury as well as by informing decisions about the 'step down' of low risk patients to a lower level of hospital care, reducing the use of hospital resources. Additionally, the potential prognostic value of urinary TIMP-2 and IGFBP7 in predicting major adverse kidney events in patients at high risk for AKI, including the need for renal replacement therapy (RRT), in-hospital mortality and the risk of chronic kidney disease, is important for clinical decision-making. The purpose of this assessment is to review the diagnostic accuracy, prognostic accuracy, clinical impact and cost-effectiveness of novel biomarkers (TIMP-2, GFBP7 and NGAL assays), for the assessment of AKI in critically ill patients.

4.2 Clear definition of the intervention

The NephroCheck test and the NGAL assays may help to assess AKI in critically ill people who are admitted to critical care or considered for admission to critical care in hospital. These tests can potentially detect kidney injury earlier than methods currently used for monitoring kidney function; serum creatinine and urine levels. Serum creatinine levels are imprecise following kidney injury. In addition, the use of intravenous fluids and diuretics in critical care settings can affect the assessment of urine levels for detection of kidney injury.

4.2.1 The NeproCheck test

The NephroCheck test (Astute Medical, Inc., USA) measures the level of 2 biomarkers in urine, the TIMP-2 (tissue inhibitor of metalloproteinase 2) and IGFBP-7 (insulin-like growth factor binding protein 7), to assess risk of moderate to severe acute kidney injury (defined as per KDIGO guidelines) in the subsequent 12 hours. The test result must be used in conjunction with clinical evaluation and results of other tests such as serum creatinine and urine output.

The concentrations of TIMP-2 and IGFBP-7 are used to calculate an AKIRisk score (the concentrations of each [ng/ml] multiplied together and divided by 1,000). A score of 0.3 or less indicates a low risk of developing *moderate to severe* AKI within 12 hours of assessment, while a score of greater than 0.3 indicates a high risk of developing *moderate to severe* AKI within 12 hours of assessment.[5]

When used with the Astute 140 Meter,NephroCheck test system consists of the following components:

- Astute140 Meter Kit (a benchtop analyser)
- Astute140 Electronic Quality Control device
- NephroCheck Test Kit (includes a single-use NephroCheck test cartridge and reagents)
- NephroCheck Liquid Control kit
- NephroCheck Calibration Verification kit

A fresh or thawed urine sample (mixed with reagent) is added to a single-use test cartridge, which is then inserted into an Astute140 Meter for incubation and result calculation. Preparation takes 3 to 5 minutes and results of NephroCheck are available in about 20 minutes. In the NHS, the Astute 140 Meter would be used in a laboratory and not at the point of care.

The test can also be run on the VITROS 3600 immunodiagnostic System and on the VITROS 5600 Integrated System clinical chemistry analysers. All systems generate a single numerical result (the AKIRisk score).

For surgical patients the NephroCheck test should be carried out 2 to 4 hours after surgery. A second administration of the test within the first 24 hours may be considered.

In the UK, Nephrocheck test is marketed for people aged over 21 years old.

4.2.2 Neutrophil gelatinase-associated lipocalin (NGAL) assays

4.2.2.1 ARCHITECT Urine NGAL

The ARCHITECT Urine NGAL test (Abbot, Germany) is a chemiluminescent micro particle immunoassay for the quantitative determination of NGAL in human urine. NGAL can be used as a marker of kidney injury.

ARCHITECT Urine NGAL test might be used as follows:

- Early detection of acute kidney injury;
- Provides a measure of the severity of acute kidney injury;
- Predicts the requirement for renal replacement therapy;
- Helps differentiate acute kidney injury from chronic kidney disease and dehydration.

For diagnostic purposes, the test results should be used in conjunction with clinical assessment and the results of any other testing that has been done (including serum creatinine and urine output). In addition, if the NGAL results are inconsistent with clinical assessment and other test results, additional testing can be done to confirm the NGAL results.

The test could be used daily until a diagnosis is made or treatment for acute kidney injury is initiated.

The expected range for the assay (for people without kidney injury) is less than or equal to 131.7 ng/ml, based on the 95th percentile from specimens from non-hospitalised donors, but results from individual laboratories may vary. The test has no age restrictions in use.

The assay is run on the ARCHITECT system (i1000SR, i2000, i2000SR, ci4100, ci8200 or ci16200) in a laboratory. The throughput of the system is up to 200 tests per hour, and the time to first result is 36 minutes.

In addition to the ARCHITECT Urine NGAL Reagent Kit, the following materials are also needed:

- ARCHITECT Urine NGAL Calibrators
- ARCHITECT Urine NGAL Controls or other control material
- ARCHITECT i pre-trigger solution
- ARCHITECT i trigger solution
- ARCHITECT i wash buffer
- ARCHITECT i reaction vessels
- ARCHITECT i sample cups
- ARCHITECT i septum
- ARCHITECT i replacement caps

4.2.2.2 The NGAL Test (using urine or plasma)

The NGAL Test (BioPorto Diagnostics, Denmark) is particle-enhanced turbidimetric immunoassay for the quantitative determination of NGAL in human urine, EDTA plasma and heparin plasma on automated clinical chemistry analysers. NGAL measurements may be useful in pre-empting the diagnosis of acute kidney injury, which may lead to acute renal failure. Urinary NGAL can serve as an early marker of acute kidney injury after cardiopulmonary bypass surgery and both urinary and plasma levels of NGAL provide an early indication of acute renal injury in unselected patients in intensive care.

The NGAL test is intended to be used alongside monitoring of serum creatinine and urine output (not as a stand-alone test) and the significance of any raised NGAL level should be interpreted in the light of a patient's clinical features.

The NGAL test can be administered as a single measurement but also as a serial measurement to detect any further development of acute kidney injury during hospitalisation, or any improvement in the clinical condition. In patients admitted to intensive care the test can be used to predict stage 2/3 AKI or as a negative predictive marker to rule out the presence of acute kidney injury.

To indicate the presence of renal disorder, including acute kidney injury, NGAL concentration in an isolated sample of urine and/or EDTA plasma should exceed 250 ng/mL. This threshold has been chosen to minimise the risk of an unacceptably high proportion of false positive results.

The assay can be run on clinical chemistry analyser systems from Roche (Cobas, Modular P), Siemens (ADVIA), Abbott (AEROSET, ARCHITECT) and Beckman Coulter (Olympus AU) in a laboratoty. The assay time is 10 minutes.

In addition to the NGAL Test Reagent Kit, the following materials are also needed:

- The NGAL Test Calibrator Kit
- The NGAL Test Control Kit
- 0.9% w/v aqueous sodium chloride solution as zero calibrator
- Analyzer-specific reagent containers

The test has no age restrictions on use.

4.3 Population and relevant subgroups

The population under consideration is critically ill people at risk of AKI who are admitted to critical care or considered for admission to critical care.

Relevant subgroups may include:

- Type of surgery (e.g., major vascular/cardiac surgery, major non-vascular surgery, trauma, solid organ transplant)
- Type of setting (i.e., intensive or critical care, cardiac care, emergency department)
- Type of sample media (i.e., urine, blood serum, blood plasma)
- People with a different underlying risk of AKI (e.g., people with chronic kidney disease, sepsis, hip fracture, major trauma, chronic liver disease)
- People with or without urinary infection and other inflammatory conditions (tests may perform differently in these populations)

4.4 Target condition: acute kidney disease

Acute kidney injury (AKI) ranges from minor loss of kidney function to complete kidney failure. In current practice, reduced kidney function is identified by elevated serum creatinine levels and/or reduced urine output.

There are many causes of acute kidney injury (NHS Choices Acute Kidney Injury)[7], including:
Pre-renal: Reduced blood flow to the kidneys, caused by:

- low blood volume (after bleeding, excessive vomiting or diarrhoea and severe dehydration),
- reduced blood flow from the heart (potentially caused by sepsis or heart/liver failure)
- damage to blood vessels which can be caused by inflammation or blockages within the kidneys
- medications that affect blood flow to the kidneys

Intrinsic: Damage to the kidney potentially caused by drugs, infections or contrast agents

Post-renal: A blockage preventing drainage from the kidneys (potentially caused by an enlarged prostate, a tumour in the pelvis or kidney stones).

People often develop acute kidney injury after major surgery.[4] In general, incidence of post-operative AKI depends on the type surgery. Rates of AKI after cardiac surgery have been reported to range from 8% to 40% according to the patient populations.[4] Recent meta-analyses have reported a pooled incidence of AKI in patients admitted to intensive care after abdominal surgery of 13.4% (95% CI 10.9% to 16.4%)[8] and after major trauma of 24% (95% CI 20% to 29%)[7] and 21% (95% CI 16.5% to 24.9%).[9]

Several tools are available for determining the stage of AKI. The NICE Clinical Knowledge Summary[10] on acute kidney injury outlines a summarised staging system for acute kidney injury in adults based on the RIFLE (Risk, Failure, Loss of kidney function, End-stage disease), AKIN (Acute Kidney Injury Network) and KDIGO (Kidney Disease: Improving Global Outcome) systems (see Table 1 below). A person's acute kidney injury should be staged by the criterion, which gives the highest stage. A classification of stage 1 or above indicates acute kidney injury.

Table 1 Summary of the staging system for acute kidney injury in adults (based on the RIFLE, AKIN, and KDIGO systems)

Stage	Criteria
1	Creatinine rise of 26 micromol or more within 48 hours OR Creatinine rise of 50–99% from baseline within 7 days* (1.50–1.99 x baseline) OR Urine output** < 0.5 mL/kg/h for more than 6 hours
2	100–199% creatinine rise from baseline within 7 days* (2.00–2.99 x baseline) OR Urine output** < 0.5 mL/kg/hour for more than 12 hours
3	200% or more creatinine rise from baseline within 7 days* (3.00 or more x baseline) OR Creatinine rise to 354 micromol/L or more with acute rise of 26 micromol/L or more within 48 hours or 50% or more rise within 7 days OR Urine output** < 0.3 mL/kg/hour for 24 hours or anuria for 12 hours
* The rise is known (based on previous blood tests) or presumed (based on history) to have occurred within 7 days. ** Measurement of urine output may not be practical in a primary care population, but can be considered in a person with a catheter.	

Source: NICE Clinical knowledge summaries on acute kidney injury (2018)[10]

People with AKI have a higher risk of mortality and spend longer in hospital.[1, 2] In addition, acute kidney injury is associated with a higher risk of developing chronic kidney disease (CKD) and end-stage renal disease. The risk of CKD increases with severity of acute kidney injury. More severe acute kidney injury has also been associated with increased mortality, length of hospital stay and use of intensive care services, in addition to a reduced chance of renal recovery.[1, 2] People with more severe acute kidney injury (and a greater loss of renal function) are more likely to need temporary renal replacement therapy.

4.5 Clinical pathway

The NICE clinical guideline on acute kidney injury[11] recommends measuring serum creatinine and comparing with baseline for adults, children and young people with acute illness if risk factors for the condition are likely or present. Risk factors include sepsis, hypovolemia and deteriorating early warning scores (using a paediatric version for children and young people). NHS England and NHS Improvement have endorsed the National Early Warning Score (NEWS) for use in acute and ambulance settings. An updated version of the score (NEWS2)[12] was published in December 2017. The score should not be used in children (under 16 years) or pregnant women.

The NICE guideline further recommends monitoring serum creatinine regularly in all adults, children and young people with or at risk of acute kidney injury. The guideline development group did not wish to define 'regularly' because this would vary according to clinical need, but recognised that daily measurement was typical while in hospital.

An **AKI algorithm** to help with detection and diagnosis of the condition has been endorsed by NHS England.[13] In some hospitals the algorithm has been integrated into Laboratory Information Management Systems (LIMS) to help identify potential cases of acute kidney injury from laboratory data in real time.

The **KDIGO Clinical Practice Guideline for Acute Kidney Injury**[14] highlights the importance of screening patients who have had an exposure that may cause acute kidney injury (such as sepsis or trauma) and that high-risk patients should continue to be monitored until risk subsides. The guideline states that intervals of checking serum creatinine is a matter of clinical judgement, but suggest as a general rule that high risk in-patients should have serum creatinine measured at least daily and more frequently after an exposure. Critically ill patients should also have urine output monitoring.

For adults who are at risk of acute kidney injury, the NICE AKI guideline[11] also recommends that systems are in place to recognise and respond to oliguria (urine output less than 0.5 ml/kg/hour). For children and young people who are at risk of acute kidney injury, the guideline recommends:

- measure urine output
- record weight twice daily to determine fluid balance
- measure urea, creatinine and electrolytes
- think about measuring lactate, blood glucose and blood gases.

Further detail on these recommendations, and further recommendations on the ongoing assessment of the condition of patients in hospital, can be found in section 1.2 of the NICE clinical guideline on AKI.[11]

The NICE guideline recommends diagnosing acute kidney injury in line with the RIFLE (or paediatric-modified RIFLE - pRIFLE), AKIN or KDIGO definitions, by using any of the following criteria:

- a rise in serum creatinine of 26 micromol/litre or greater within 48 hours
- a 50% or greater rise in serum creatinine known or presumed to have occurred within the past 7 days
- a fall in urine output to less than 0.5 ml/kg/hour for more than 6 hours in adults and more than 8 hours in children and young people
- a 25% or greater fall in eGFR in children and young people within the past 7 days.

There are no direct therapies for treating acute kidney injury. Care focuses on optimising hemodynamics and fluid status, avoiding nephrotoxic treatments, and carrying out investigations to identity and resolve the underlying cause as quickly as possible. In general, the goal of care is to prevent any further kidney injury and stop progression of the disease; in particular, to prevent progression to a stage where renal replacement therapy is needed.

The **NICE clinical guideline on AKI**[11] highlights the importance of identifying the cause, or causes, of acute kidney injury and has recommendations on the use of urinalysis and ultrasound for this purpose.

The **KDIGO Clinical Practice Guideline for Acute Kidney Injury**[14] also recommends prompt evaluation of people with acute kidney injury to determine the cause. Identifying possible reversible causes of the condition is highlighted as important to reduce severity of the condition.

The NICE clinical guideline on AKI[11] has recommendations on managing acute kidney injury (section 1.5); covering removing urological obstruction, pharmacological management, renal replacement therapy and referral to nephrology services. **The KDIGO Clinical Practice Guideline for Acute Kidney Injury**[14] recommends staging severity of acute kidney injury with serum creatinine and urine

output, and to manage the condition according to stage and cause. General management principles for people at high risk of acute kidney injury (or with the condition) are to:

- discontinue nephrotoxic agents if possible,
- monitor volume status and perfusion pressure,
- consider functional haemodynamic monitoring,
- monitor serum creatinine and urine output,
- avoid hyperglycaemia,
- consider alternatives to radiocontrast procedures.

Further actions should only be considered at higher stages of acute kidney injury, such as renal replacement therapy. Dosages of drugs may also need to be adapted because of reduced kidney function. The KDIGO guideline also has more detailed guidance on the prevention and treatment of acute kidney injury (section 3). This includes haemodynamic monitoring and support, glycemic control and nutritional support, the use of diuretics and vasodilator therapy.

In UK clinical practice the NephroCheck test and NGAL assays are likely to be used for the assessment of AKI in people who are considered for admission to critical care rather than in patients already admitted to critical care. It is worth pointing out that the NephroCheck test, the ARCHITECT Urine NGAL test, the NGAL plasma and the NGAL urine test would not replace serum creatinine and urine output monitoring but they would be used alongside current monitoring to facilitate earlier detection of kidney injury and prompt adoption of strategies to prevent further progression of kidney disease.

4.6 Key factors to be addressed

There are several components to this work that fall within the scope of the following research questions:

1. Do novel biomarkers (NephroCheck test, ARCHITECT Urine NGAL test, NGAL plasma test and NGAL urine test) accurately detect emerging AKI in critically ill people who are admitted to critical care or considered for critical care?
2. Do the novel biomarkers (NephroCheck test, ARCHITECT Urine NGAL test, NGAL plasma test and NGAL urine test) predict the development of future events (e.g., AKI, mortality, need for long-term renal replacement therapy) in critically ill people at risk of AKI who are admitted to critical care or considered for critical care?
3. Does the use of novel biomarkers (NephroCheck test, ARCHITECT Urine NGAL test, NGAL plasma test and NGAL urine test) lead to improvements in clinical outcomes of critically ill people who are admitted to critical care or considered for critical care? (i.e.,

reduction in events rates – such as mortality and long-term renal replacement therapy - among patients whose management is guided by the novel biomarkers)

4. Does routine use of novel biomarkers (NephroCheck test, ARCHITECT Urine NGAL test, NGAL plasma test and NGAL urine test) affect costs to the NHS, length or quality of life (i.e. Quality Adjusted Life Years, QALYs), or cost-effectiveness (incremental cost per QALY gained)?

In brief, the main objectives of this assessment are the following:

- To determine the diagnostic accuracy, prognostic accuracy/predictive validation and clinical impact of the use of novel biomarkers (NephroCheck test, ARCHITECT Urine NGAL test, NGAL plasma test and NGAL urine test) for the assessment of acute kidney injury in critically ill patients (adults and children) admitted to critical care or considered for critical care;
- To develop an economic model to assess the cost-effectiveness of the use of novel biomarkers (NephroCheck test, ARCHITECT Urine NGAL test, NGAL plasma test and NGAL urine test) for the assessment of acute kidney injury in critically ill patients (adults and children) admitted to critical care or considered for critical care.

5. Evidence synthesis methods

This section describes methods for research questions 1 to 3. Methods for research question 4 will be described in Section 6.

5.1 Inclusion and exclusion criteria

5.1.1 Population

Adults and children with critical illness at risk of developing AKI who are admitted to critical care or considered for admission to critical care in hospital. People at risk of developing AKI are broadly defined as those '*who are having their serum creatinine and urine output monitored'* but do not have established AKI. Even though both adults and children will be considered suitable for inclusion, we anticipate that only a minority of studies will focus on children

5.1.2 Intervention

The interventions under investigation are the following fluid biomarkers:

- the NephoCheck test (Asture Medical)
- the ARCHITECT Urine NGAL test (Abbott)
- the NGAL plasma test (BioPorto Diagnostics)
- the NGAL urine test (BioPorto Diagnostics)

All tests must be used alongside existing care including monitoring of serum creatinine and urine output and clinical judgement.

At present, there is no universally accepted reference standard for the diagnosis of acute kidney injury. The current methods for detecting or predicting AKI are in line with the RIFLE (or paediatric-modified RIFLE), AKIN and KDIGO classification systems, which are based on the assessment of serum creatinine levels and urine output alongside clinical judgement. A classification of 1 or above indicates a diagnosis of AKI, whilst a classification of 2 or 3 indicates moderate to severe AKI (see Table 1 above and the NICE guideline on AKI). In the absence of a universally accepted reference standard, the current methods for detecting or predicting AKI will be used as the reference standard.

5.1.3 Study design and Outcomes

Type of suitable studies and relevant clinical outcome for research questions 1-3 are illustrated below.

Table 2 Eligibility criteria for research question 1 (diagnostic accuracy of novel biomarkers)

Population	Critically ill patients at risk of AKI
Biomarkers under investigation	• the NephoCheck test • the ARCHITECT Urine NGAL test • the NGAL plasma test • the NGAL urine test All used in conjunction with existing care
Reference standard	Existing clinical criteria for the monitoring of serum creatinine and urine output used in conjunction with clinical judgement
Measures of accuracy	Sensitivity and specificity
Study design	Include: • Any cross-sectional study which investigates the diagnostic accuracy of a single biomarker (NephroCheck test or NGAL assays) against the reference standard in the same study population • Any fully paired direct comparison (observational or randomised direct comparison) in which one novel biomarker and a comparator biomarker (or assay) are evaluated in the same study population against the reference standard Exclude: • Studies with insufficient information to complete a two-by-two contingency table

Table 3 Eligibility criteria for research question 2 (prognostic accuracy/prospective validation)

Population	**Critically ill patients at risk of AKI**
Biomarkers under investigation	• the NephoCheck test • the ARCHITECT Urine NGAL test • the NGAL plasma test • the NGAL urine test All used in conjunction with existing care
Comparator	Existing clinical criteria for the monitoring of serum creatinine and urine output used in conjunction with clinical judgement
Outcome for prediction	• Mortality • Need for long-term renal replacement therapy • Chronic kidney disease >90 days post AKI
Study design	• Randomised controlled trials • Cohort studies, preferably with prospective enrolment • Case-control studies

Table 4 Eligibility criteria for research question 3 (impact on clinical outcomes)

Population	**Critically ill patients at risk of AKI**
Biomarkers under investigation	• the NephoCheck test • the ARCHITECT Urine NGAL test • the NGAL plasma test • the NGAL urine test All used in conjunction with existing care
Comparator	Existing care, including monitoring of serum creatinine and urine output alongside clinical judgement
Outcome	Clinical outcomes • Mortality • AKI-associated morbidity (e.g., chronic kidney disease/end stage renal disease, other organ failure) Patient-reported outcome • Health-related quality of life Intermediate outcomes • Incidence/duration of acute renal replacement therapy within 7 days • Incidence of chronic kidney disease-related renal replacement therapy post acute kidney injury • Length of stay in critical/intensive care • Length of stay in hospital

	• Incidence of hospital readmission post-discharge • Impact of test result on clinical decision making • Impact on steady state estimated glomerular filtration rate at 90 days • Equivalence of biomarkers (e.g., the NGAL assays)
Study design	• Randomised controlled trials • Prospective cohort studies with a concurrent comparison group

For addressing the above research questions we will include the following types of study design:

- Performed in secondary and tertiary settings
- Published in English

We will exclude the following types of study reports:

- Narrative or systematic reviews
- Case reports or case series
- Studies published as abstracts or conference proceedings for which the full text is not available
- Studies that assess people immediately after a kidney transplant
- Studies that use solid tissue (not fluid) biomarkers or imaging modalities for the diagnosis of AKI

5.4 Search methods for identification of studies

5.4.1 Electronic searches

The following databases will be searched, with no date, language, or publication type restriction:

- Ovid MEDLINE
- Ovid EMBASE
- Science Citation Index Expanded
- CINAHL
- CENTRAL

A highly sensitive search strategy will be developed, to include index terms, free-text words, abbreviations and synonyms, to combine biomarkers and AKI. The preliminary MEDLINE and EMBASE search strategies are presented in Appendix 1. These searches will be adapted for other electronic databases. All search strategies will be reported in full in the final version of this assessment.

5.4.2 Searching other sources

We will search the following sources for additional ongoing or unpublished studies:

- ClinicalTrials.gov (www.clinicaltrials.gov/)
- WHO International Clinical Trials Registry Platform (ICTRP) Search Portal (apps.who.int/trialsearch)
- WHO Global Index Medicus (www.who.int/library/about/The Global Index Medicus/en/).

Furthermore, websites of relevant professional organisations and health technology agencies, as well as appropriate clinical experts, will be consulted for additional reports. The reference lists of all papers retrieved in full will be perused to identify further relevant studies.

5.5 Study selection and data extraction strategies

One reviewer will screen the citations identified by the search strategies. A second reviewer will independently screen a random sample of citations (20%). Full text versions of potentially relevant articles will be retrieved. Two reviewers will independently assess each article for eligibility based on the pre-specified inclusion criteria. We will resolve any disagreement by discussion or consultation with a third reviewer. Multiple publications of the same studies will be linked and considered together. For excluded studies, we will document reasons for exclusion. We will illustrate the study selection process by means of a PRISMA flow diagram.

Two reviewers will independently extract data from each eligible study using a form developed *ad hoc* for the purpose of this assessment. Any disagreements will be resolved by discussion or consultation with a third reviewer.

The following information will be recorded from each study:

1. Characteristics of studies: first author, year of publication, country, language, setting, objectives, inclusion and exclusion criteria, type of enrolment
2. Characteristics of study participants: age, target condition, clinical history, previous tests, number of participants enrolled and included in the analysis, reasons for withdrawal
3. Characteristics of the technologies under investigation (NephroCheck test, ARCHITECT Urine NGAL test, NGAL plasma test and NGAL urine test).
4. The reported number of true positives, false positives, false negatives and true negatives and area under the receiver-operating characteristic curve (AUC) for each test for each relevant outcome.
5. For studies using the tests to predict clinical outcomes such as renal replacement therapy and in-hospital mortality we will also extract information on the duration of follow-up, number of

patients followed up and number of patients who experienced the outcome of interest during the follow-up period.

5.6 Quality assessment strategy

We will use QUADAS-2 criteria to assess the quality of included diagnostic studies.[15] QUADAS-2 consists of four domains: patient selection, index test, reference standard and flow and timing. Each domain is assessed in terms of 'low', 'high' or 'unclear' risk of bias, and the first three in terms of concerns regarding 'low', 'high' or 'unclear' applicability.

For prognostic and prediction model studies, we will assess risk of bias by using an approach based on the Prediction model Risk Of Bias ASsessment Tool (PROBAST), which is structured into four domains: participants, predictors, outcome and analysis.[16] These domains contain signaling questions to facilitate the judgement of risk of bias. We will also consider the use of PROBAST for studies of diagnostic test accuracy that include a validation or prediction model.

We will use the Cochrane risk of bias tool[17] for the assessment of randomised trials evaluating the clinical utility of the novel biomarkers under (NephroCheck test, ARCHITECT Urine NGAL test, the NGAL plasma test and NGAL urine test). For assessing the quality of non-randomised evidence reporting quantitative data on the clinical utility of the tests we will use the checklist developed by the HSRU, University of Aberdeen, in partnership with the NICE Review Body for Interventional Procedures (ReBIP). The ReBIP checklist was adapted from several sources[18-21] and comprises 17 items, which assess the following aspects: generalisability, sample definition and selection, description of the intervention, outcome assessment, adequacy of follow-up, and performance of the analysis.

One reviewer will extract the data and a second reviewer will check the data extracted. Any disagreements will be resolved by consensus or consultation with a third reviewer.

5.7 Methods of analysis/synthesis

We will analyse TP, FP, FN and TN for each test in every study. We will enter diagnostic data in the two-by-two tables into Review Manager software (Review Manager 5.0), which will allow the sensitivity and specificity estimates together with their 95% confidence intervals (CIs) to be presented in forest plots and plotted in the receiver operating characteristic (ROC) space for each diagnostic test under investigation.

Where appropriate we will perform meta-analysis of each pair of sensitivity and specificity estimates from each included study for each relevant test. We intent to use the Hierarchical Summary ROC

(HSROC) model implemented in SAS (using Proc NLMIXED). Although the bivariate method may be appropriate for studies that use the same threshold to define a positive test result, it may not be appropriate for studies that use different positivity thresholds.[22] The HSROC will also allow exploration of heterogeneity by incorporating co-variates into the model when there is a sufficient number of included studies. We will perform separate meta-analyses for each diagnostic test.

Heterogeneity will be assessed initially by visual inspection of the forest plots of sensitivity and specificity and of the prediction region in the SROC plots. If there are sufficient data, we will investigate sources of heterogeneity in estimates of test accuracy by adding covariates to the statistical model. We will consider the following potential sources of heterogeneity: characteristics of the population (stage of AKI, baseline serum creatinine level), aspects of the diagnostic/prognostic tests (positivity threshold) and of the reference standard (criteria for assessing the presence of AKI), characteristic of the clinical setting (e.g., post-surgery care/cardiac care, critical care), and type of sample mode (e.g., urine versus blood serum or plasma).

If sufficient data are available, we will use sensitivity analyses to assess the impact of studies methodological quality on the results of our analyses. In particular, we will restrict analysis to studies judged at low risk of bias.

We will consider calculating the AUC with 95% CI and cut-off values to measure the performance of each test for the prediction of AKI. We will perform meta-analyses for the AUC for each test for each relevant outcome. An AUC $\geq$0.70 will indicate a useful risk predictor. We will use the I^2 statistic to determine the proportion of between-study variation due to heterogeneity rather than sample error in the AUC-ROC. We will use the following thresholds for the interpretation of I^2: <30% will indicate low heterogeneity, 30–60% moderate heterogeneity and >60% high heterogeneity. Where possible for the analysis of prediction models we will consider recently published guidance and recommendations.[23]

In addition, we will assess current evidence about the prognostic value of these biomarkers using risk estimates such as relative risks, hazard ratios and odd ratios (provided their outcome incidence rate is low) along with their relative 95% CIs. Summary statistics for any continuous outcomes for example length of stay will be compared between groups (participants who are assessed using the biomarkers in addition to standard care versus those who receive standard care only). We will take into account suitable transformations and subsequent normalising since urinary biomarker levels are often skewed and reporting on a log-scale offers better linearity thereby reducing heterogeneity among studies.[24] Again the Q- and I^2-statistics will be used to identify the presence of heterogeneity between studies.

For studies with confirmed homogeneity, fixed-effects model will be adopted, otherwise, random-effects models will be used.

We will not assess reporting bias as, at present, there is no consensus on recommended methods for evaluating publication bias in the diagnostic field.[22]

When appropriate, we intend to summarise the results of the studies (RCTs and observational studies) evaluating the clinical impact of the NephroCheck test and NGAL assays using standard meta-analysis methods.[17] We will consider a narrative synthesis of results if considerable clinical and methodological heterogeneity is observed between studies.

6. Report methods for synthesising evidence of cost-effectiveness

The specific objectives for the assessment of cost-effectiveness are to:

- Review and critically appraise existing economic evaluations of the NephroCheck test, the ARCHITECT Urine NGAL test, the NGAL plasma test and the NGAL urine test for evaluating critically ill patients (adults and children) at risk of AKI.
- Develop a *de novo* economic model to assess the cost-effectiveness of the NephroCheck test, the ARCHITECT Urine NGAL test, the NGAL plasma test and the NGAL urine test in combination with standard clinical assessment, compared with standard clinical assessment alone (i.e., serum creatinine and urine output) for evaluating critically ill patients (adults and children) at risk of AKI from a UK NHS and personal social services perspective.

6.1 Identifying and systematically reviewing published cost-effectiveness studies

Comprehensive search strategies will be developed to identify economic evaluations of the NephroCheck test, the ARCHITECT Urine NGAL test, the NGAL plasma test and the NGAL urine test. The following databases will be searched, with no date, language, or publication type restriction:

- Ovid MEDLINE
- Ovid EMBASE
- NHS Economic Evaluations Database
- HTA Database
- Research Papers in Economics
- ISPOR Scientific Presentations Database

Websites of relevant professional organisations and health technology agencies will be consulted for additional reports. Reference lists of all included studies will be perused, and appropriate experts contacted for details of additional reports of economic evaluations.

Any identified full economic evaluations matching the NICE final scope will be included. Full economic evaluations are defined as comparative analyses of costs and outcomes in the framework of cost-utility, cost-effectiveness, cost-benefit or cost-minimisation analyses. Economic evaluations conducted alongside single effectiveness studies (e.g., RCTs) or decision analysis models will be deemed eligible for inclusion. Included evaluations will be appraised against the NICE reference case for the assessment of the cost-effectiveness of diagnostic tests.[25] The main findings will be summarised in a narrative review, and results across studies will be tabulated for comparison.

6.2 Evaluation of costs and cost effectiveness

Following the review of cost-effectiveness evidence, an economic model will be developed to assess the cost-effectiveness of the NephroCheck test, the ARCHITECT Urine NGAL test, the NGAL plasma test and the NGAL urine test in combination with standard clinical assessment, compared with standard clinical assessment alone (including serum creatinine and urine output) for evaluating critically ill patients (adults and children) at risk of AKI. The evidence on costs and cost-effectiveness will be evaluated according to the recommendations of the NICE Diagnostics Assessment Programme manual.[25]

6.3 Development of a health economic model

The model type and structure will be determined in consultation with clinical experts, including the NICE assessment subgroup. It is anticipated that the event pathways will be modelled through a number of mutually exclusive Markov health states. Transition probabilities between the health states (expressed on a constant cycle length) will govern the flow of cohorts through the model. However, we will retain the flexibility to move to an individual simulation approach if the preferred conceptual model becomes too complex to implement as a Markov cohort model.

It is likely that disease modelling will incorporate two phases, to reflect the pathways of care that a cohort of patients would follow, first through AKI and then progression through to longer-term health outcomes and survival, which will include chronic kidney disease (CKD) and end stage renal disease. The model will include available evidence regarding adverse health outcomes associated with AKI, such as higher mortality, increased hospital length of stay, reduced chance of renal recovery, and the risk of developing CKD. Where appropriate evidence to guide key model parameters does not exist, assumptions based on clinical expert opinion may be required. Where clinical expert opinion is used to populate the model, uncertainty is greater, and so these assumptions will be tested in sensitivity analyses.

The care pathways and care bundles will be modelled in line with relevant NICE guidelines[11], and supplemented with discussion with clinical experts. Where feasible, and where sufficient data exist, the model will consider the direct impact of test strategy on health outcomes. Alternatively, if there are no direct data on the effects of diagnostic strategies on important health outcomes, a linked evidence approach will be used to model the impacts of improved diagnostic accuracy and risk prediction on short and long-term health outcomes and associated resource use. This will depend on whether there is an appropriate source of evidence to demonstrate benefits of modified treatment packages for patients with AKI, whose care is modifiable, in response to a diagnostic test result.

An NHS and PSS perspective will be adopted throughout, and the model will be run over a time period that is sufficient to realise all the costs and benefits associated with early diagnosis of AKI. If evidence supports long-term differences in health outcomes between the alternative testing strategies, this will be a life-time horizon. Costs and benefits (QALYs) that occur into the future will be discounted at an annual rate of 3.5% per annum.[25]

The model will be populated using data obtained and synthesised from the systematic review of diagnostic accuracy and / or clinical outcomes studies as appropriate, as well as any relevant data obtained from the systematic review of cost-effectiveness studies. To further inform the economic model, additional broad searches will be carried out to identify existing decision models for AKI. The structure of and inputs for these models will be considered potentially informative for the purposes of developing the economic model. Previous NICE guidance relevant to the decision problem will be consulted to ensure consistency with previous NICE evaluations where it is possible and appropriate to do so.[11, 26] Additional searches will also be conducted, where appropriate, to inform population of key model parameters (e.g., resource use, probabilities, utilities). Where feasible and appropriate, model parameter searches from a recent NIHR-HTA assessing various biomarkers for the early detection of AKI in ICU patients will be updated.[27] Priority will be given to data that are consistent with the NICE reference case (e.g. descriptive health related quality of life data elicited from UK patients using the EQ-5D, and valued using general population preferences).[25] Routine sources of unit cost data from a UK NHS perspective will be used whenever possible, and where necessary will be supplemented by study specific cost estimates, based on expert opinion, or provided by manufacturers as appropriate.[28, 29] If feasible and if sufficient data exist, risks (probabilities) of the included events under standard practice will be informed by a review of published observational/registry data applicable to the UK clinical setting. Data from the control arms of identified randomised controlled trials will also be assessed for generalisability to the UK context.

Costs associated with the delivery of the comparator (standard clinical management) will be based on a review of current clinical guidelines and published data on the frequency of monitoring tests in the

UK NHS. The costing will include staff time, consumables and equipment and will be validated with clinical experts. Costs associated with the diagnostic tests will include the costs of standard clinical management in addition to costs specific to each test under evaluation. The costing will be informed by the frequency of testing and the experience/opinions of the specialist committee members for the assessment. Unit costs for the alternative devices, test cartridges, and any associated consumables will be sourced from the companies (at the price most relevant to the NHS). Tests that require the use of platforms / analysers (not included in the test price) will incur additional costs. For example, in the absence of an analyser for the NephroCheck test at a hospital, a benchtop Astute 140 meter (~£3000 ex VAT) would be required. Any required capital equipment costs such as platforms for the tests will be selected based on standard UK practice and will be amortised over the estimated useful lifespan of the device, and allocated on a per patient or per test basis using estimates of annual throughput per device. The impact of applying different assumptions with respect to testing frequency, platform types available at different hospitals, and throughput will be explored through sensitivity analyses.

The results of the model will be presented in terms of a cost-utility analysis. A multi-treatment comparison will be undertaken, with each strategy compared incrementally to its next less effective non-dominated comparator, to estimate its incremental cost per quality adjusted life year gained (QALY).[30] ICERs for pairwise comparisons against current clinical management will also be reported. The modelling exercise will use the net benefit framework to identify the optimal testing strategy at different threshold ratios of willingness to pay per QALY. To characterise the uncertainty surrounding point estimates of incremental costs and effects, probabilistic sensitivity analyses will be undertaken.[31] The results of these analyses will be presented in the form of cost-effectiveness acceptability curves (CEACs) and frontiers (CEAFs). Further deterministic sensitivity and scenario analyses will be used to address other forms of uncertainty. This will focus in particular on areas where assumptions regarding the care pathway are required and for parameters where little or no high quality evidence exists. Where evidence allows, subgroup analyses will explore the impact on cost-effectiveness of testing following different types of major surgery and different characteristics of the patient population (e.g., age group).

7. Handling information from the companies

Following a request for information, any 'commercial in confidence' data provided by a company and specified as such will be highlighted in blue and underlined in the assessment report (followed by an indication of the relevant company name e.g. in brackets). Any academic-in-confidence data provided will be highlighted in yellow and underlined.

8. Competing interests of authors

None

9. References

1. Bedford M, Stevens PE, Wheeler TW, Farmer CK. What is the real impact of acute kidney injury? *BMC Nephrology* 2014;**15:**95.

2. Selby NM, Crowley L, Fluck RJ, McIntyre CW, Monaghan J, Lawson N, et al. Use of electronic results reporting to diagnose and monitor AKI in hospitalized patients. *Clinical Journal of The American Society of Nephrology: CJASN* 2012;**7:**533-40.

3. National Confidential Enquiry into Patient Outcome and Death. Adding insult to injury: A review of the care of patients who died in hospital with a primary diagnosis of acute kidney injury (acute renal failure). A report by the National Confidential Enquiry into Patient Outcome and Death. London: NCEPOD; 2009

4. Park JT. Postoperative acute kidney injury. *Korean Journal of Anesthesiology* 2017;**70:**258-66.

5. National institute for Health and Care Excellence. *Medtech innovation briefing [MIB156] – NeproCheck test to help assess the risk of acute kidney injury in critically ill patients*. London/Manchester: 2018. https://www.nice.org.uk/advice/mib156 [Accessed 5 April 2019]

6. Gocze I, Koch M, Renner P, Zeman F, Graf BM, Dahlke MH, et al. Urinary biomarkers TIMP-2 and IGFBP7 early predict acute kidney injury after major surgery. *PLoS ONE [Electronic Resource]* 2015;**10:**e0120863.

7. NHS Digital. 2019. Acute kidney injury [Online] NHS Digital. Available from: https://www.nhs.uk/conditions/acute-kidney-injury/. [Accessed: 18 April 2019]

8. O'Connor ME, Kirwan CJ, Pearse RM, Prowle JR. Incidence and associations of acute kidney injury after major abdominal surgery. *Intensive Care Medicine* 2015521-30.

9. Haines RW, Fowler AJ, Kirwan CJ, Prowle JR. The incidence and associations of acute kidney injury in trauma patients admitted to critical care: A systematic review and meta-analysis. *The Journal of Trauma and Acute Care Surgery* 2019;**86:**141-7.

10. National Institute for Health and Care Excellence. *Clinical Knowledge Summary: Acute Kidney Injury [Online]* London/Manchester: 2018. https://cks.nice.org.uk/acute-kidney-injury. [Accessed 5 April 2019]

11. National Institute for Health and Care Excellence. *Acute kidney injury: prevention, detection and management Clinical guideline [CG169]*. London/Manchester: 2013. https://www.nice.org.uk/guidance/cg169/chapter/1-Recommendations#assessing-risk-of-acute-kidney-injury [Accessed 5 April 2019]

12. Royal College of Physicians. 2017. National Early Warning Score (NEWS)2 [Online] London: Royal College of Physicians. Available from: https://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news-2. [Accessed: 18 April 2019]

13. NHS England. n.d. Acute Kidney Injury (AKI) Algorithm [Online] Available from: https://www.england.nhs.uk/akiprogramme/aki-algorithm/. [Accessed: 18 April 2019]

14. Kidney Disease: Improving Global Outcomes. *KDIGO Clinical Practice Guideline for Acute Kidney injury (Online Appendices A-F)*: 2012. https://kdigo.org/wp-content/uploads/2016/10/KDIGO-AKI-Suppl-Appendices-A-F_March2012.pdf [Accessed 4 April 2019]

15. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155:**529-36.

16. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. *Annals of Internal Medicine* 2019;**170:**W1-W33.

17. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. London: The Cochrane Collaboration. 2011

18. Centre for Reviews and Dissemination. *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*. York: 2009. www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf [Accessed March 2013]

19. Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Bouter LM, et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *Journal of Clinical Epidemiology* 1998;**51:**1235-41.

20. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health* 1998;**52:**377-84.

21. Jackson R, Ameratunga S, Broad J, Connor J, Lethaby A, Robb G, et al. The GATE frame: critical appraisal with pictures. *Evidence Based Medicine* 2006;**11:**35-8.

22. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* The Cochrane Collaboration; 2010.

23. Debray TP, Damen JA, Snell KI, Ensor J, Hooft L, Reitsma JB, et al. A guide to systematic review and meta-analysis of prediction model performance. *BMJ* 2017;**356:**i6460.

24. Zhou LT, Lv LL, Pan MM, Cao YH, Liu H, Feng Y, et al. Are Urinary Tubular Injury Markers Useful in Chronic Kidney Disease? A Systematic Review and Meta Analysis. *PLoS ONE [Electronic Resource]* 2016;**11:**e0167334.

25. National Institute for Health and Care Excellence. *Diagnostics Assessment Programme manual*. London/Manchester: 2011. https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-diagnostics-guidance/Diagnostics-assessment-programme-manual.pdf [Accessed 5 April 2019]

26. National Institute for Health and Care Excellence. *Chronic kidney disease in adults: assessment and management Clinical guideline CG182*. London/Manchester: 2014. https://www.nice.org.uk/guidance/cg182 [Accessed 2 May 2019]

27. Hall PS, Mitchell ED, Smith AF, Cairns DA, Messenger M, Hutchinson M, et al. The future for diagnostic tests of acute kidney injury in critical care: evidence synthesis, care pathway analysis and research prioritisation. *Health Technology Assessment (Winchester, England)* 2018;**22:**1-274.

28. NHS Improvement. 2017. Reference Costs [Online] London: NHS Improvement. Available from: https://improvement.nhs.uk/resources/reference-costs/. [Accessed: April 2019]

29. Curtis L, Burns A. Unit Costs of Health and Social Care. Cantebury: Kent Uo. 2018

30. Drummond MF, O'Brien B, Sitoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 2005

31. Briggs A, Claxton K, Sculpher M. Analysing and presenting simulation output from probabilistic models. New York: Oxford Univerisity Press; 2006

Appendix 1 Literature search strategy

Database: Embase <1974 to 2019 March 20>, Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) <1946 to March 20, 2019>

1 exp acute kidney injury/ use ppezv (42642)
2 acute kidney failure/ use oemezd (71413)
3 (Acute adj3 (kidney disease* or kidney injury or kidney failure or kidney dysfunction or kidney insufficienc*)).tw,kw. (51171)
4 (Acute adj3 (renal disease* or renal injury or renal failure or renal dysfunction or renal insufficienc*)).tw,kw. (63271)
5 ((Acute adj3 (Tubular Necrosis or nephrotoxic*)) or "nephrotoxic injur*").tw,kw. (9593)
6 AKI.tw,kw. (27405)
7 "contrast induced nephropathy".tw,kw. (5148)
8 acute kidney tubule necrosis/ use oemezd (4502)
9 or/1-8 (161591)
10 nephrocheck.tw,kw. use ppezv (22)
11 nephrocheck.tw,dv,kw. use oemezd (52)
12 10 or 11 (74)
13 "Tissue Inhibitor of Metalloproteinase-2"/ use ppezv (3422)
14 "tissue inhibitor of metalloproteinase 2"/ use oemezd (6766)
15 Metalloproteinase inhibitor 2.tw,nm,kw. use ppezv (14)
16 Metalloproteinase inhibitor 2.tw,kw. use oemezd (27)
17 tissue inhibitor of metalloproteinase-2.tw,nm,kw. use ppezv (3671)
18 tissue inhibitor of metalloproteinase-2.tw,kw. use oemezd (872)
19 TIMP metallopeptidase inhibitor 2.tw,nm,kw. use ppezv (10)
20 TIMP metallopeptidase inhibitor 2.tw,kw. use oemezd (11)
21 (TIMP 2 or TIMP2 or DDC8 or CSC-21K).tw,nm,kw. use ppezv (4775)
22 (TIMP 2 or TIMP2 or DDC8 or CSC-21K).tw,kw. use oemezd (6037)
23 or/13-22 (14375)
24 (IGFBP7 or IBP-7 or IGFBP-rP1).tw,nm,kw. use ppezv (403)
25 (IGFBP7 or IBP-7 or IGFBP-rP1).tw,kw. use oemezd (599)
26 IGF-binding protein 7.tw,nm,kw. use ppezv (15)
27 IGF-binding protein 7.tw,kw. use oemezd (22)
28 Insulin-like growth factor-binding protein 7.tw,nm,kw. use ppezv (217)
29 Insulin-like growth factor-binding protein 7.tw,kw. use oemezd (316)
30 MAC25 protein.tw,nm,kw. use ppezv (5)

31 MAC25 protein.tw,kw. use oemezd (5)
32 PGI2-stimulating factor.tw,nm,kw. use ppezv (6)
33 PGI2-stimulating factor.tw,kw. use oemezd (9)
34 Prostacyclin-stimulating factor.tw,nm,kw. use ppezv (29)
35 Prostacyclin-stimulating factor.tw,kw. use oemezd (31)
36 Tumor-derived adhesion factor.tw,nm,kw. use ppezv (15)
37 Tumor-derived adhesion factor.tw,kw. use oemezd (7)
38 or/24-37 (1248)
39 23 or 38 (15354)
40 9 and 39 (437)
41 12 or 40 (446)
42 remove duplicates from 41 (321)