NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

DIAGNOSTICS ASSESSMENT PROGRAMME

Diagnostics consultation document

Tests to help assess risk of acute kidney injury for people being considered for critical care admission (ARCHITECT and Alinity i Urine NGAL assays, BioPorto NGAL test and NephroCheck test)

The National Institute for Health and Care Excellence (NICE) is producing guidance on using the ARCHITECT and Alinity i Urine NGAL assays, BioPorto NGAL test and NephroCheck test to help assess the risk of acute kidney injury in the NHS in England. The diagnostics advisory committee has considered the evidence and the views of clinical and patient experts.

This document has been prepared for public consultation. It summarises the evidence and views that have been considered, and sets out the recommendations made by the committee. NICE invites comments from registered stakeholders, healthcare professionals and the public. This document should be read along with the evidence (the diagnostics assessment report and the diagnostics assessment report addendum).

The advisory committee is interested in receiving comments on the following:

- Has all of the relevant evidence been taken into account?
- Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
- Are the recommendations sound, and a suitable basis for guidance to the NHS?

Equality issues

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that the

recommendations may need changing to meet these aims. In particular, please tell us if the recommendations:

- could have a different effect on people protected by the equality legislation than on the wider population, for example by making it more difficult in practice for a specific group to access the technology
- could have any adverse effect on people with a particular disability or disabilities.

Please provide any relevant information or data you have about such effects and how they could be avoided or reduced.

Note that this document is not NICE's final guidance on the ARCHITECT and Alinity i Urine NGAL assays, BioPorto NGAL test and NephroCheck test to help assess the risk of acute kidney injury. The recommendations in section 1 may change after consultation.

After consultation, the committee will meet again to consider the evidence, this document and comments from the consultation. After considering the comments, the committee will prepare its final recommendations, which will be the basis for NICE's guidance on the use of the technology in the NHS in England.

For further details, see the diagnostics assessment programme manual.

Key dates:

Closing date for comments: 21 January 2020

Second diagnostics advisory committee meeting: 29 January 2020

1 Recommendations

1.1 There is not enough evidence to recommend the routine use of the ARCHITECT and Alinity i Urine neutrophil gelatinase-associated lipocalin (NGAL) assays, NephroCheck test or BioPorto NGAL test to help assess the risk of acute kidney injury for people being considered for critical care admission.

1.2 Further research is recommended to assess:

- the clinical effectiveness of defined care bundles to prevent or reduce the effect of acute kidney injury in defined NHS patient populations who could benefit from preventive care for acute kidney injury (see section 5.1)
- the effect on clinical outcomes of having the tests to guide care to prevent acute kidney injury (see section 5.2).

Why the committee made these recommendations

Using the tests may help to identify people with acute kidney injury earlier than monitoring serum creatinine and urine levels alone. But it is not clear how much this will benefit people being considered for admission to critical care in the NHS, for example, by reducing their hospital stay or likelihood of needing renal replacement therapy in hospital.

The cost-effectiveness estimates for the tests are very uncertain. But they are likely to be much higher than what NICE normally considers a cost-effective use of NHS resources. Therefore, these tests are not recommended for use in the NHS.

There is considerable uncertainty about which patients in the NHS could benefit from the tests. This is because preventive care for acute kidney injury may already be done (in full or in part) as standard practice, which limits the effect that the test results can have on guiding care. Further research may identify specific populations in the NHS who could benefit from the tests, and by how much.

2 The diagnostic tests

Clinical need and practice

Acute kidney injury

2.1 Acute kidney injury ranges from minor loss of kidney function to complete kidney failure. In current practice, reduced kidney function is identified, and staged, by elevated serum creatinine levels or reduced urine output, or both. There are no direct therapies for treating most types of acute

kidney injury. Care focuses on optimising haemodynamics and fluid status, avoiding nephrotoxic treatments, and identifying and resolving the underlying cause as quickly as possible. A goal of care is to prevent any further kidney injury and to stop acute kidney injury progressing; in particular, to try and prevent progression to a stage when renal replacement therapy is needed.

2.2 The NephroCheck and neutrophil gelatinase-associated lipocalin (NGAL) tests could potentially detect kidney injury earlier than current methods for monitoring kidney function; serum creatinine and urine levels. Serum creatinine levels are slow to rise after kidney injury. Also, using intravenous fluids and diuretics can cause issues when detecting kidney injury by measuring urine levels. Earlier identification of acute kidney injury could allow earlier adoption of measures, such as care bundles (a group of interventions, or processes, which when implemented together can help to reduce the severity of acute kidney injury). These could prevent the condition progressing to more severe injury and reduce the risk of adverse outcomes for patients.

2.3 The NephroCheck test is indicated for use in people who are critically ill, but the NGAL tests potentially have a broader indication. At the scoping workshop and assessment subgroup meeting, clinical experts considered the most relevant population for this assessment. They considered the different types of care for people who are critically ill to determine who could benefit from use of the tests in the NHS. People who are admitted to NHS critical care should already have a range of interventions designed to prevent acute kidney injury because they are extremely unwell. Therefore, the potential for the tests to improve outcomes in this population is limited in the NHS because the results of the tests are unlikely to change management decisions. Clinical experts highlighted that the tests could be useful for people who are being considered for admission to critical care; that is, when a decision about admission has not been made and the test

results could guide the use of preventive care for acute kidney injury. The decision question for this assessment therefore focuses on this population.

The interventions

The NephroCheck test

2.4 The NephroCheck test (Astute Medical) measures the level of 2 biomarkers (tissue inhibitor of metalloproteinase 2 [TIMP-2] and insulin-like growth factor binding protein 7 [IGFBP-7]) in urine and uses the concentrations to help assess risk of moderate to severe acute kidney injury (defined as per the Kidney Disease Improving Global Outcomes [KDIGO] guidelines) in the subsequent 12 hours. The company states that the test result is intended to be used in conjunction with clinical evaluation as an aid in the risk assessment of acute kidney injury in the critically ill.

2.5 The concentrations of TIMP-2 and IGFBP-7 are used to calculate an AKIRisk score (the concentrations of each [nanograms/millilitre; ng/ml] multiplied together and divided by 1,000). A score of over 0.3 indicates a higher risk of developing moderate to severe acute kidney injury within 12 hours of assessment. The test can be run on the Astute 140 meter, the VITROS 3600 immunodiagnostic system and the VITROS 5600 and VITROS 7600 integrated system clinical chemistry analysers. The company states that the test is marketed in the UK for people over 21 years.

ARCHITECT and Alinity i Urine NGAL assays

2.6 The ARCHITECT and Alinity i Urine NGAL assays (Abbott) are chemiluminescent microparticle immunoassays for the quantitative determination of NGAL in human urine. The company states that 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.

2.7 The company has no set threshold for a positive result. The ARCHITECT and Alinity i Urine NGAL assays are run on different analysers but use the same reagents. The ARCHITECT assay is run on the ARCHITECT system (i1000SR, i2000, i2000SR, ci4100, ci8200 or ci16200). The test has no age restrictions on use.

The BioPorto NGAL test

2.8 The BioPorto NGAL test (BioPorto Diagnostics) is a particle-enhanced turbidimetric immunoassay for the quantitative determination of NGAL in human urine, ethylenediaminetetraacetic acid (EDTA) plasma and heparin plasma. The company states that this is not a standalone test and clinicians should interpret the significance of any raised NGAL level alongside a person's clinical features.

2.9 The company advises that the NGAL concentration in an isolated sample of urine or EDTA plasma should exceed 250 ng/ml to indicate the presence of renal disorder, including acute kidney injury. The assay can be run on various clinical chemistry analyser systems in a laboratory. The test has no age restriction on use.

The comparator

2.10 No additional testing to identify people at high risk of developing acute kidney injury (other than standard serum creatinine and urine output monitoring).

3 Evidence

The diagnostics advisory committee (section 8) considered evidence on the ARCHITECT and Alinity i Urine neutrophil gelatinase-associated lipocalin (NGAL) assays, BioPorto NGAL test and NephroCheck test for detecting emerging acute kidney injury from several sources. Full details of all the evidence are in the committee papers.

Clinical effectiveness

3.1 The external assessment group (EAG) did a systematic review to identify evidence on the diagnostic accuracy and clinical effectiveness of the ARCHITECT and Alinity i Urine NGAL assays, BioPorto NGAL test and NephroCheck test to help assess, and reduce, the risk of acute kidney injury for critically ill patients who are being considered for critical care admission. Although the population in the scope is people being considered for critical care admission, to maximise the available data the EAG included data from studies that enrolled patients already admitted to critical care.

3.2 In total, 56 studies (reported in 71 articles) were included. Of these, 46 enrolled adults only, 8 enrolled children only and 2 enrolled both adults and children. Twenty-eight studies were done in Europe (4 in the UK), 15 in North America, 9 in Asia, 2 in North America and Europe, 1 in Australia and 1 study did not provide details of location. In most studies data were collected prospectively.

3.3 The studies either reported data on using the biomarkers to detect or predict acute kidney injury or to predict clinical outcomes (mortality or need for renal replacement therapy [RRT]) in critically ill patients admitted to hospital. No randomised controlled trials or controlled clinical trials were identified. No studies compared using the biomarkers with standard clinical care for clinical effectiveness outcomes.

3.4 The studies assessed using the tests in various clinical settings. The EAG divided the studies in adults and children into 3 groups based on clinical setting: people who had cardiac surgery, people who had major non-cardiac surgery and people admitted to critical care (including critically ill patients presenting to the emergency department, patients admitted to intensive care or patients considered for critical care for various medical conditions).

Evidence on accuracy to detect emerging acute kidney injury

3.5 Test accuracy was determined by the ability of the tests to identify the presence of acute kidney injury according to current clinical criteria (that is, using serum creatinine and urine output). A rise in serum creatinine levels or fall in urine output, or both, occurring within a certain time after the NephroCheck or NGAL test was done (this varied between studies, from within 12 hours to within 7 days) were used to indicate if acute kidney injury occurred (reference standard). The EAG could extract or derive the necessary data for calculating sensitivity and specificity estimates from 33 of the included studies.

3.6 The QUADAS-2 tool was used for quality assessment of the studies. The EAG commented that it was not clear in most studies if the tests were interpreted without knowledge of the reference standard (unclear risk of bias). Studies that used NephroCheck were judged at low risk of bias for interpretation of the test because they used a common threshold for a positive result. However, for the NGAL studies a common threshold was not used. The EAG also commented that in the NGAL studies the threshold was not pre-specified before data were collected. Two studies were assessed as being at high risk of bias on the patient flow domain because more than 50% of the participants were excluded from the analysis (Jaques et al. 2019) or because of poor reporting (Asada et al. 2016). The EAG considered that the applicability of the index test results to the NHS was unclear in many studies because there was wide variation in the NGAL threshold used to define a positive test result and in the timing of the test sample collection. The EAG commented that it had no major concerns that the patient population, index text and reference standard were not applicable to the review question. However, in some of the included studies people were already admitted to critical care.

3.7 Because the threshold used for a positive test result varied in the identified studies, the EAG ran meta-analyses using the hierarchical

summary ROC (HSROC) model to estimate summary values for sensitivity and specificity. If multiple thresholds were used in a study, the EAG selected 1 threshold to use in its analysis. Meta-analysis was only done if data from 4 or more studies were available.

NephroCheck test (adults)

3.8 All studies assessed used the NephroCheck test on urine samples. No studies were done in the UK. Two studies assessed using NephroCheck to detect acute kidney injury after cardiac surgery and 5 studies assessed its use in hospitalised patients admitted to intensive or critical care for various clinical reasons. No studies were identified in people who had major non-cardiac surgery. The summary estimate for sensitivity was 0.75 (95% confidence interval [CI] 0.58 to 0.87) and for specificity was 0.61 (95% CI 0.49 to 0.72). The EAG commented that there was heterogeneity across studies and noted that estimates of specificity were generally low.

ARCHITECT Urine NGAL assay (adults)

3.9 Two studies provided test accuracy data on using the ARCHITECT NGAL assay for detecting acute kidney injury in patients who had cardiac surgery and 4 studies assessed its use in hospitalised patients admitted to intensive or critical care for various clinical reasons. No studies were done in the UK or were identified in people who had major non-cardiac surgery. The summary estimate for sensitivity was 0.67 (95% CI 0.58 to 0.76) and for specificity was 0.72 (95% CI 0.64 to 0.79). The EAG commented that there was heterogeneity across studies.

The BioPorto NGAL test – urine (adults)

3.10 Eight studies assessed using the BioPorto NGAL test with urine for detecting acute kidney injury: 1 study in people who had cardiac surgery, 1 study in people who had major non-cardiac surgery and 6 studies in hospitalised patients admitted to intensive or critical care for various clinical reasons. One study was done in the UK (Matsa et al. 2014). The

summary estimate for sensitivity was 0.73 (95% CI 0.65 to 0.80) and for specificity was 0.83 (95% CI 0.64 to 0.93). The EAG commented that there was heterogeneity across studies.

The BioPorto NGAL test – plasma (adults)

3.11 The EAG only identified studies in the critical care setting for the BioPorto NGAL test used with blood plasma (4 studies). One study was done in the UK (Matsa et al. 2014). The summary estimate for sensitivity was 0.76 (95% CI 0.56 to 0.89) and for specificity was 0.67 (95% CI 0.40 to 0.86). The EAG commented that there was heterogeneity across studies.

Children

3.12 Seven studies assessed using the NGAL assays with urine samples to detect acute kidney injury in children. No studies were done in the UK. No studies assessing the use of NephroCheck in children were identified.

ARCHITECT Urine NGAL assay (children)

3.13 Five studies assessed using the ARCHITECT Urine NGAL assay for detecting acute kidney injury in children who had cardiac surgery. The summary estimate for sensitivity was 0.68 (95% CI 0.53 to 0.80) and for specificity was 0.79 (95% CI 0.63 to 0.89). The EAG commented that there was considerable heterogeneity across studies. No studies were identified in a population who had major non-cardiac surgery. One study assessed using the ARCHITECT Urine NGAL assay for detecting acute kidney injury in children admitted to intensive or critical care for various clinical reasons. The sensitivity and specificity were 0.77 (95% CI 0.60 to 0.90) and 0.85 (95% CI 0.74 to 0.92), respectively.

The BioPorto NGAL test – urine (children)

3.14 One study assessed using the BioPorto NGAL test with urine for detecting acute kidney injury in children who had cardiac surgery. NGAL was measured using a concentration normalised by units of creatinine. The

sensitivity and specificity were 0.77 (95% CI 0.69 to 0.84) and 0.47 (95% CI 0.40 to 0.54), respectively.

Evidence on ability to predict intermediate outcomes

3.15 The EAG identified 11 studies with data on the ability of the tests to predict mortality, 4 studies with data on predicting the need for RRT and 3 studies that assessed the ability of the tests to predict worsening of acute kidney injury. All studies were in critically ill patients at risk of acute kidney injury. For predicting mortality, area under the curve (AUC) values varied from 0.55 to 0.91. For predicting the need for RRT, AUC values varied from 0.68 to 0.86. For predicting worsening of acute kidney injury, AUC values varied from 0.66 to 0.71.

3.16 The EAG commented that adding the tests to existing clinical models generally improved risk prediction of newly developed acute kidney injury, or worsening of acute kidney injury, and mortality. However, it cautioned that there were limited data available and the statistical models used varied between studies. Also, information on potential candidate variables considered in studies was often not provided.

3.17 No studies were identified that reported the effect of using the tests on clinical or patient-reported outcomes.

Cost effectiveness

Systematic review of cost-effectiveness evidence

3.18 The EAG did a systematic review to identify any published economic evaluations of the ARCHITECT and Alinity i Urine NGAL assays, the BioPorto NGAL test (plasma and urine) and the NephroCheck test for evaluating people at risk of developing acute kidney injury. Two of the studies identified used modelling strategies that were similar, and that the EAG considered appropriate for the current decision problem. One of these (Hall et al. 2018) was done in the UK, and the EAG considered it a

comprehensive and high-quality assessment. But because the setting was outside the scope of this assessment (people already admitted to intensive care units), the EAG adapted the model for critically ill patients who are at risk of acute kidney injury and being considered for admission to critical care.

Model structure

3.19 The EAG developed a de novo economic model designed to assess the cost effectiveness of using the tests (in addition to standard clinical monitoring) to help detect the risk of developing acute kidney injury and to help start early preventive care.

3.20 This was a 2-stage model using TreeAge Pro software. Limited direct evidence was identified that showed the effect of using the tests (compared with standard monitoring alone) on health outcomes (such as acute kidney injury status, mortality, development of chronic kidney disease). So the EAG used observational associations to infer how preventing or reducing the severity of acute kidney injury may affect changes in health outcomes (a linked-evidence approach). An initial decision-tree phase modelled:

- The accuracy of the tests to identify people with emerging acute kidney injury.
- For people with a positive biomarker test result, the effect of preventive measures (a Kidney Disease Improving Global Outcomes [KDIGO] care bundle) on reducing the probability that they develop acute kidney injury or reducing the severity of the condition if they develop it.
- The effect of developing acute kidney injury, and its severity, on short-term outcomes (within 90 days): whether a person is admitted to intensive care, length of stay in intensive care or hospital, development of chronic kidney disease and 90-day mortality.

After this initial 90-day period, a longer-term Markov model was used to model the effect of developing acute kidney injury while in hospital on the risk of developing chronic kidney disease, and the effect of this condition on the rest of a person's life.

Population

3.21 The modelled population was people in hospital at risk of developing acute kidney injury, having their serum creatinine and urine output monitored. The EAG used the Grampian population register of hospitalisations to characterise this population. This dataset included 17,630 adults admitted to hospital in Grampian in 2003. It is the complete population of all patients who had an abnormal kidney function blood test on hospital admission and had at least an overnight stay in hospital, including all patients who developed acute kidney injury. The model starting base-case cohort was 63 years old, 54.3% women, with about 11% having chronic kidney disease (in the model, more could develop this condition over time). The base-case prevalence of acute kidney injury (that is, people who will develop the condition while in hospital under standard monitoring) was assumed to be 9.2%.

Model inputs

3.22 The sensitivity and specificity of the tests to identify people who will develop acute kidney injury (as shown by a later increase in serum creatinine or drop in urine output, or both) was taken from the systematic review and meta-analysis done in the clinical effectiveness section. The EAG used values pooled from all studies identified for each of the tests across all clinical settings. The incidence of acute kidney injury and the effect of developing the condition on clinical outcomes (admission to intensive care, 90-day mortality) was estimated by the EAG largely using data from the Grampian observational dataset. The model could vary which clinical outcomes were affected by acute kidney injury status, and the size of this effect.

3.23 The EAG assumed that a KDIGO care bundle would be the preventive care used if the tests were positive. It did a literature search to identify studies to estimate the effectiveness of this intervention for the model. The EAG did not include the identified studies in its clinical effectiveness review because the studies did not report the direct effect of using the tests on clinical outcomes. Instead the EAG included the studies in its cost-effectiveness section (as part of the rationale for parameter values used in the model). The EAG used data from Meersch et al. (2017) for the effect of the KDIGO care bundle in the model. This was a single-centre randomised controlled trial done in Germany in people who had cardiac surgery (n=276). People who had a positive NephroCheck test (using a score of over 0.3) were randomised to either standard care (less intensive care than in the KDIGO care bundle) or standard care plus a KDIGO care bundle. People having standard care followed the recommendations of the American College of Cardiology Foundation (2011), which included keeping mean arterial pressure over 65 mmHg and central venous pressure between 8 and 10 mmHg. The KDIGO care bundles included avoiding nephrotoxic agents, discontinuing angiotensin-converting enzyme inhibitors and angiotensin receptor blockers, close monitoring of urine output, serum creatinine, avoiding hyperglycaemia (for 72 hours), considering alternatives to radiocontrast agents, and optimising fluids. Although there was a significant reduction in occurrence of acute kidney injury by 72 hours for the KDIGO arm compared with standard care (odds ratio 0.48 [95% CI 0.29 to 0.80]), the EAG commented that this did not appear to translate to other clinical outcomes (need for RRT in hospital, 90-day all-cause mortality and length of stay in intensive care or hospital).

3.24 The EAG found 2 other studies reporting the effects of KDIGO care bundles; Gocze et al. (2018) and Schanz et al. (2018). Both were done in Germany and assessed the effect of NephroCheck-guided application of a KDIGO care bundle compared with standard care (no use of a care bundle). Gocze et al. was a smaller study (n=121) than Meersch et al. and

reported that NephroCheck-guided care (after major non-cardiac surgery) showed a trend towards a lower probability of acute kidney injury. But the results were not statistically significant; the odds ratio for standard care compared with NephroCheck was 1.96 (95% CI 0.93 to 4.10). There was, however, a statistically significant increase in the odds of stage 2 or 3 acute kidney injury in the standard care group compared with NephroCheck; 3.43 (95% CI 1.04 to 11.32). Schanz et al. (n=100) compared the effect of a NephroCheck-triggered implementation of KDIGO recommendations for acute kidney injury with standard care alone in an emergency department in Germany. Acute kidney injury outcomes were similar in both groups. The probability of acute kidney injury stage 2 or 3 was 32.1% for the intervention group and 33.3% for the control group after 1 day. After 3 days this was 38.9% for intervention group and 39.1% for the control group. The effect size from Gocze et al. was used in a scenario analysis. Data from intensive care registers, reports and studies were used for parameters in the longer-term Markov model.

Costs

3.25 In its base-case analysis, the EAG assumed that an Astute 140 meter would need to be purchased to use NephroCheck, so included the cost of this. The EAG assumed that the NGAL tests are run on platforms that are already available in hospital laboratories, so the cost of these analysers was assumed to be negligible and was not included in the analysis. A scenario analysis was done in which no capital costs (including an analyser) or training costs were included for the tests.

Table 1 Test-related costs

Cost per test	NephroCheck	BioPorto NGAL test [a]	Abbott NGAL assays	
			ARCHITECT	Alinity [b]
Platform cost	£0.53	–	–	–
Equipment cost	£49.80	£20.00	£25.71	£28.29
Maintenance/ consumables	£4.23	£1.90	£3.51	£3.51
Staff costs	£37.62	£37.62	£37.62	£37.62
Staff training costs	£0.08	£0.03	£0.03	£0.03
Total cost	**£92.26**	**£59.55**	**£66.87**	**£69.44**

[a] Costs assumed to be the same for plasma and urine samples.

[b] The Alinity NGAL assay was not included in the base-case analysis because of a lack of data for this assay.

Abbreviation: NGAL, neutrophil gelatinase-associated lipocalin

3.26 The EAG assumed that the KDIGO care bundle would be applied for an additional 3 days over and above standard care for people who tested positive on the NephroCheck or NGAL tests (based on clinical opinion and consistent with the primary outcome measure from Meersch et al. 2017). Resources included in the care bundle costs included intravenous fluids (including nurse time), nephrologist and pharmacist review time and stopping blood pressure medication. The total additional cost of applying the KDIGO bundle was assumed to be £106.36 per person.

Health-related quality of life

3.27 The EAG updated the searches run in Hall et al. (2018) to identify any additional source of utility data for its model for both the initial decision-tree phase and longer-term Markov model. The age- and sex-matched EQ-5D UK population norms were calculated using an equation published by Ara and Brazier (2010). These were used to derive age- and sex-adjusted utility multipliers from the raw pooled estimates from studies, based on the age and sex distribution of the source studies.

Base-case assumptions

3.28 The following assumptions (in addition to those described in previous sections) were applied in the base-case analyses:

- Acute kidney injury, and more severe acute kidney injury, can be prevented by earlier NephroCheck or NGAL-guided use of a KDIGO care bundle (for people who would otherwise develop it with standard monitoring alone) in base case 1. In base case 2, NGAL-guided care cannot prevent acute kidney injury (but can reduce the severity of the condition).
- In base case 1, the NephroCheck biomarkers and NGAL rise at similar times and the earlier identification of emerging kidney injury (relative to serum creatinine and urine output changes) is the same for both tests.
- There are no adverse effects on health caused by a false-positive NephroCheck or NGAL test result.
- No adaptions to standard monitoring were made for people testing negative on NephroCheck or NGAL tests (although standard monitoring done alongside would detect acute kidney injury for false-negative tests, just at a later time). This was because the EAG assumed that de-escalation of care would not occur solely because of a negative test result.
- Everyone with a positive NephroCheck or NGAL test immediately had a KDIGO care bundle.
- After 5 years post-transplant, mortality reverted to the general population all-cause mortality probability. The annual probability of transplant failure remained as that reported from years 3 to 5 in the UK renal registry.
- The proportion of the cohort whose transplant failed returned to dialysis. Their probability of progressing from end-stage renal disease on dialysis to a second transplant was the same as for progressing to the first transplant.

Base-case results

3.29 No evidence for NGAL test-guided implementation of preventive care for acute kidney injury on clinical outcomes was identified. Therefore, the EAG did 2 base cases:

- Base case 1: Using the NGAL test had the same effect as the NephroCheck test to prevent acute kidney injury and reduce severity of the condition if it occurred (based on Meersch et al. 2017).
- Base case 2: Using the NGAL test could only reduce the severity of acute kidney injury (as for base case 1), not prevent it from occurring (NephroCheck effects are unchanged).

3.30 The results of base case 1 (probabilistic) are shown in table 2. Because of uncertainty about the extent of any effect of acute kidney injury on other clinical outcomes, the EAG did several scenario analyses (B, C and D). This was in addition to the base case varying which outcomes acute kidney injury occurrence (and severity) affected, and the size of this effect. Scenario C was the most pessimistic (no effect of preventing acute kidney injury, or reducing severity, on clinical outcomes) and scenario D was the most optimistic (full effect of preventing acute kidney injury, or reducing severity, on clinical outcomes).

Table 2 Cost-effectiveness results (probabilistic) for base case 1

Test	Total cost	Total QALYs	ICER (probability cost effective at £20,000 per QALY gained)	
			Fully incremental	Compared with standard monitoring
BioPorto NGAL test (urine)	£22,887	6.07332	– (43.5%)	Dominant (54.6%)
BioPorto NGAL test (plasma)	£22,900	6.07332	£2,694,918 (11.1%)	Dominant (47.6%)
Standard monitoring only	£22,901	6.07296	Dominated (45.1%)	–
ARCHITECT NGAL	£22,912	6.07328	Dominated (0.1%)	£32,131 (41.4%)
NephroCheck	£22,938	6.07332	Dominated (0.2%)	£101,456 (31.9%)

Abbreviations: NGAL, neutrophil gelatinase-associated lipocalin; QALYs, quality-adjusted life years; ICER, incremental cost-effectiveness ratio

3.31 Scenario C assumed no benefit of reducing acute kidney injury occurrence or severity on clinical outcomes (most pessimistic). Standard care dominated all the tests in this scenario, with all tests having 0% probability of being cost effective at a maximum acceptable incremental cost-effectiveness ratio (ICER) of £20,000 per quality-adjusted life year (QALY) gained. Scenario D assumed a full benefit of reducing acute kidney injury occurrence or severity on clinical outcomes (most optimistic; see table 3).

Table 3 Cost-effectiveness results (probabilistic) for scenario D (in base case 1)

Test	Total cost	Total QALYs	ICER (probability cost effective at £20,000 per QALY gained)	
			Fully incremental	Compared with standard monitoring
Standard monitoring only	£22,959	6.08383	– (0.7%)	–
BioPorto NGAL test (urine)	£23,013	6.11006	£2,052 (40.7%)	£2,052 (99.3%)
BioPorto NGAL test (plasma)	£23,028	6.11091	£17,702 (47.5%)	£2,538 (99.1%)
ARCHITECT NGAL	£23,031	6.10799	Dominated (1.1%)	£2,981 (98.8%)
NephroCheck	£23,065	6.11064	Dominated (10.0%)	£3,955 (97.7%)

Abbreviations: NGAL, neutrophil gelatinase-associated lipocalin; QALYs, quality-adjusted life years; ICER, incremental cost-effectiveness ratio

3.32 The EAG also did several further scenario analyses. Changes made to several parameters improved the cost effectiveness of the tests, in that they all dominated standard care (in a pairwise comparison):

- Increasing long-term costs and risk of mortality in the Markov model (scenario G) for people who were admitted to intensive care while in hospital (in the decision-tree phase).
- For people having acute kidney injury while in hospital, extending the time of increased risk of developing chronic kidney disease from 1 year to the rest of a person's life (scenario H).
- Increasing the prevalence of acute kidney injury to 23% (from 9.2% in base case; scenario K).

Assuming false-positive tests increased mortality (scenario M), which worsened the cost effectiveness of the tests.

3.33 In scenario Q, the EAG used alternative accuracy estimates from studies that enrolled children only. Data were only available for the ARCHITECT NGAL and the BioPorto NGAL (urine) tests. The EAG cautioned that the model was not configured for children but used parameters from an adult population. Because limited accuracy data for the tests in children were available and there was a lack of data for other parameters, the EAG considered the analysis to be exploratory only.

3.34 In base case 2 (probabilistic analysis), NephroCheck dominated all other tests, with an ICER of about £106,000 per QALY gained compared with standard monitoring. The probability of NephroCheck being the most cost-effective test across scenario analyses increased considerably.

3.35 In scenario T (provided in an addendum to the diagnostics assessment report), the EAG used Gocze et al. (rather than Meersch et al.) to inform estimates of the effect of a KDIGO care bundle on reducing the risk of developing acute kidney injury, or the severity of the condition if it did develop. This improved the cost-effectiveness estimates of the tests. In base case 1, all tests dominated standard monitoring. In base case 2, NephroCheck dominated all other tests and standard monitoring.

4 Committee discussion

Preventing or reducing the severity of acute kidney injury could benefit patients

4.1 The patient expert explained that a diagnosis of acute kidney injury can be very unexpected and can have a substantial effect on people and their families. Acute kidney injury can mean prolonged stays in hospital, which are distressing for patients and cost family members time and money. The patient expert also explained that earlier detection of acute kidney injury might make temporary renal replacement therapy (RRT) less likely. This would benefit people by reducing the need for this invasive treatment and

would release resources. Also, developing acute kidney injury increases the risk of chronic kidney disease. The patient expert emphasised that end-stage renal disease changes people's lives (and that of their families), because it affects their lifestyle and ability to work. If the tests helped detect acute kidney injury earlier and allowed interventions to prevent or reduce the severity of the condition, this could benefit patients by improving clinical outcomes.

There is considerable uncertainty about which patients in the NHS could benefit from the tests

4.2 Clinical experts explained that the definition of critical care varied across the world. People tend to be more unwell before they are admitted to critical care in the NHS than in the US or the rest of Europe. So in the NHS they should already be having all available interventions to prevent acute kidney injury. Therefore, the potential of the tests to change care and improve outcomes in NHS critical care is very limited. Clinical experts also commented that it was uncertain which patients in the NHS could benefit from targeted use of preventive care bundles for acute kidney injury. They commented that care bundles (in addition to standard care) were the only option currently available to try and prevent acute kidney injury or reduce its severity. They also explained that a care bundle is a very complex intervention. It involves implementing measures (such as avoiding nephrotoxic agents, avoiding hyperglycaemia and optimising fluids) that can protect the kidneys from further damage. Many of these will already be done as part of standard care, depending on the clinical setting and the person's condition (that is, they are more likely to have been done already in more intensive care). Care bundles can also be tailored to a person's condition. Not all measures may be done if this is not clinically appropriate. Therefore, the effect of the care bundles could vary between different populations. Clinical experts suggested that critical care outreach teams could potentially use the tests to guide preventive

care. The committee concluded that there was considerable uncertainty about who in the NHS could benefit from the tests.

Clinical effectiveness

The accuracy of the tests to detect emerging acute kidney injury, and the clinical significance of their results, is highly uncertain

4.3 Most of the available data for the tests were sensitivity and specificity estimates. These measured the tests' ability to identify people who will be diagnosed with acute kidney injury using current clinical criteria (serum creatinine or urine levels). The time of acute kidney injury diagnosis varied from within 12 hours to within 7 days. The EAG commented that there was considerable clinical and statistical heterogeneity seen across the studies, which included very different populations, and therefore the results should be interpreted with caution. The committee also noted that even the best estimates of sensitivity and specificity showed that using the tests could result in large proportions of falsely positive or negative results. Clinical experts commented that the staging of the condition in classification systems (such as Kidney Disease Improving Global Outcomes [KDIGO]) was developed by clinical consensus and there was uncertainty about the clinical significance of the different stages of acute kidney injury. The committee concluded that there was uncertainty about how well the tests could detect emerging acute kidney injury, and the clinical significance of what they detect in studies of test accuracy.

Cost effectiveness

There is considerable uncertainty about the effect of care bundles on developing acute kidney injury, and whether this would be seen in the NHS

4.4 Clinical experts commented that there was considerable uncertainty about how much benefit the care bundles used in the NHS would provide to prevent, or reduce the severity of, acute kidney injury if used earlier (when

NephroCheck or NGAL tests indicate risk of acute kidney injury; see section 4.2). In its model, the EAG used data from Meersch et al. (2017; see section 3.23) for the effect of test-guided preventive care (a KDIGO care bundle; see section 3.23) on reducing the chance of developing acute kidney injury or reducing the severity of the condition if it developed. The committee noted that people in the control arm did not have the KDIGO care bundle. This was unlikely to reflect NHS practice because although using the tests could allow earlier use of the care bundle, everyone at risk would eventually have an acute kidney injury care bundle at a later time, once serum creatine or urine levels showed acute kidney injury. The absence of the KDIGO care bundle in the control group could therefore have overestimated the treatment effect from Meersch et al. compared with NHS practice. Using the treatment effect size from Gocze et al. rather than Meersch et al. improved the cost effectiveness of the tests (see section 3.35). Clinical experts commented that standard care in Germany (the control arms of the 3 identified studies on the effectiveness of the KDIGO care bundle) may differ from standard care in the UK. Therefore, the generalisability of the results of these studies to the NHS was potentially limited. The committee concluded that there was substantial uncertainty about how much effect a KDIGO care bundle had on developing, or reducing the severity of, acute kidney injury. It also concluded that it was uncertain whether a treatment effect size determined in German studies would be seen in the NHS, and therefore if the modelled effect of the KDIGO care bundle on acute kidney injury would be seen in the NHS.

It is not appropriate to assume that the results of the NephroCheck and NGAL tests are equivalent in the economic model

4.5 No studies were identified that showed the effect of NGAL-guided use of the KDIGO care bundle. In base case 1, the EAG assumed that the effect of NephroCheck and NGAL-guided preventive care on acute kidney injury incidence was the same. It used data from Meersch et al. (2017), a study

done in people who had a positive NephroCheck test, to estimate the effect of test-guided preventive care on acute kidney injury incidence. Clinical experts commented that the biomarkers used in the NephroCheck test (tissue inhibitor of metalloproteinase 2 [TIMP-2] and insulin-like growth factor binding protein 7 [IGFBP-7]) may perform very differently to NGAL as indicators of acute kidney injury because they are released during different physiological processes. The committee concluded that it was not appropriate to assume that the results of the NephroCheck and NGAL tests were equivalent. It also concluded that data from Meersch et al. should not be used to inform estimates of how well NGAL-guided use of the KDIGO bundle affects acute kidney injury incidence in the economic model.

It is uncertain how much the incidence, and severity, of acute kidney injury affects clinical outcomes

4.6 In its model, the EAG used observational data to link incidence and severity of acute kidney injury to the probability of further outcomes, such as length of stay in hospital, 90-day mortality and need for RRT. However, the committee noted that in Meersch et al. use of the KDIGO bundle reduced acute kidney injury incidence, but not length of stay in hospital or intensive care, need for RRT in hospital or 90-day all-cause mortality. In Gocze et al. length of hospital and intensive care stay was significantly shorter in the KDIGO bundle study arm, but there was no significant difference in need for RRT or mortality in hospital. Clinical experts explained that how each stage of acute kidney injury affected shorter- and longer-term clinical outcomes was not clearly understood (see section 4.3). The EAG investigated how much varying the effect of having acute kidney injury, and severity, had on further clinical outcomes in scenario analyses. This led to large variation in cost effectiveness (see section 4.9). The committee concluded that it was uncertain how much the incidence and severity of acute kidney injury affected clinical outcomes.

The economic model should include the cost of analysers for the NGAL assays

4.7 The EAG did not include the cost of analysers needed to run the NGAL assays in their estimates of cost per NGAL test. This was because it assumed that the NGAL tests are run on platforms already available in hospital laboratories, so the cost of these analysers was negligible. Clinical experts commented that the analysers needed to run the different NGAL assays would not be in every hospital. The committee concluded that it would have been reasonable to include the cost of analysers needed to run the NGAL assays in the cost per test used in the model, as had been done for the NephroCheck test.

The tests may be used very differently for children and the cost-effectiveness estimates for this group are highly uncertain

4.8 The committee discussed the lack of data available for children. It noted that the EAG did a scenario analysis that used alternative accuracy estimates from studies that enrolled children only (scenario Q; see section 3.33). Because of a lack of data for other parameters, the EAG had to use values derived from adult populations. The EAG cautioned that this analysis should be considered as exploratory. In addition, clinical experts commented that the potential use for children could be very different to that for adults in the NHS. The committee concluded that, because of a lack of data to inform model parameters and uncertainty in their intended use, the cost-effectiveness estimates of the tests for children was highly uncertain. The committee considered that future studies should consider the utility of the tests for children (see section 4.11).

The cost-effectiveness estimates are highly uncertain and potentially much higher than the committee normally consider cost effective

4.9 The EAG did multiple scenario analyses to reflect the uncertainty about which clinical outcomes would be affected by both the incidence and

severity of acute kidney injury. It cautioned that the results of the cost-effectiveness modelling were largely speculative and should be interpreted with caution. Also, it considered it impossible to determine the best incremental cost-effectiveness ratio (ICER) given the available evidence. Incremental quality-adjusted life years (QALYs) were very low across the scenarios, with tests often having ICERs over £50,000 per QALY gained, or being dominated (that is, they had higher costs and lower QALYs) compared with standard monitoring. Varying the parameter values used in scenario analyses substantially affected the cost-effectiveness estimates for the tests. Changes to some parameters improved the cost effectiveness of the tests, to the extent that they dominated standard care (in base case 1) when compared in a pairwise manner (see sections 3.30 and 3.33). The committee further recalled that it did not consider it appropriate to use data from NephroCheck-guided use of the KDIGO care bundle to estimate the effect of NGAL-guided use of the KDIGO care bundle (see section 4.5). The committee concluded that there was substantial uncertainty about the best cost-effectiveness estimate for the tests in the defined clinical population. However, the estimate could potentially be much higher than what NICE normally considers cost effective.

There is too much uncertainty about the cost effectiveness of the tests to recommend adoption

4.10 The committee agreed that there was substantial uncertainty about how the tests could be used in the NHS (see section 4.2) and their likely cost effectiveness. This was mainly because there was uncertainty about the effect that test-guided care could have on the incidence of acute kidney injury (see section 4.4) and on other clinical outcomes (see section 4.6) in the defined NHS clinical population. Also, how clinicians would react to the test results in the NHS was unclear (that is, the changes to care they would make in response to a positive or negative result). The cost-effectiveness estimates for the tests are very uncertain and are potentially

much higher than what NICE normally considers a cost-effective use of NHS resources (see section 4.9). The committee concluded that there was too much uncertainty about the cost effectiveness of the tests to recommend their adoption in the NHS. It also noted that in most scenarios, estimates for the tests were far higher than what NICE usually considers a cost-effective use of NHS resources. Further research could provide clarity on how the tests would affect care and outcomes in the NHS and allow their cost effectiveness to be estimated.

Research considerations

Consideration should be given to defining populations in the NHS who would benefit from test-guided preventive care

4.11 The committee recalled that there was uncertainty about which patient populations in the NHS could benefit from test-guided use of preventive care for acute kidney injury (see section 4.2). If care bundles were already being used, in full or in part, in a patient population this would limit the effect that the test results can have on guiding care. Clinical experts commented that the potential use for children and young people can also be very different to adults, so specific consideration is needed for this group. The cost of the NephroCheck (about £90) and of providing the KDIGO care bundle earlier (about £105) were similar. Therefore, the committee questioned whether providing the care bundle earlier to everyone (that is, without testing) could be the most cost-effective strategy for some patient populations in the NHS. The committee concluded that, before further studies are done, it was important that companies define the patient populations in the NHS who could benefit from test-guided preventive care. It noted that people who are critically unwell in the NHS would likely already be having all available care to prevent acute kidney injury.

5 Recommendations for further research

5.1 Companies should specify patient populations in the NHS who could benefit from test-guided preventive care. Further research is then recommended to assess the clinical effectiveness of defined care bundles designed to prevent or reduce the effect of acute kidney injury in the NHS. Studies should be done in a defined patient population who could benefit from preventive care for acute kidney injury. Research should be done in children, young people and adults, but specific considerations may need to be given to children and young people when care differs from that for an adult population (see section 4.11).

5.2 Further research is recommended to assess the effect of test-guided preventive care (see section 5.1) on clinical outcomes (such as length of stay in hospital, mortality and need for renal replacement therapy and progression to chronic kidney disease). Research should be done in children, young people and adults, but specific considerations may need to be given to children and young people when care differs from that for an adult population. Studies should investigate the effects of both positive and negative test results on clinical decisions and subsequent care.

6 Implementation

NICE intends to develop tools, in association with relevant stakeholders, to help organisations put this guidance into practice.

In addition NICE will support this guidance through a range of activities to promote the recommendations for further research. The research proposed will be considered by the NICE Medical Technologies Evaluation Programme research facilitation team for developing specific research study protocols as appropriate. NICE will also incorporate the research recommendations in section 5 into its guidance research recommendations database and highlight these recommendations to public research bodies.

7 Review

NICE reviews the evidence 3 years after publication to ensure that any relevant new evidence is identified. However, NICE may review and update the guidance at any time if significant new evidence becomes available.

Mark Kroese

Chair, diagnostics advisory committee

December 2019

8 Diagnostics advisory committee members and NICE project team

Committee members

This topic was considered by the diagnostics advisory committee, which is a standing advisory committee of NICE.

Committee members are asked to declare any interests in the test to be assessed. If it is considered there is a conflict of interest, the member is excluded from participating further in that assessment.

The minutes of each committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Additional specialist committee members took part in the discussions for this topic:

Specialist committee members

Dr Banwari Agarwal

Consultant in critical care medicine, Royal Free Hospital

Dr Sally Brady

Consultant clinical scientist, Viapath

Dr Mark Devonald

Consultant nephrologist, Nottingham University Hospitals

Mr Guy Hill

Lay member

Dr Christopher Kirwan

Consultant in critical care and renal Medicine, Barts Health NHS Trust

Dr Mark Thomas

Consultant physician and nephrologist, Birmingham Heartlands Hospital

Dr Kay Tyerman

Consultant paediatric nephrologist, Leeds Children's Hospital

NICE project team

Each diagnostics assessment is assigned to a team consisting of a technical analyst (who acts as the topic lead), a technical adviser and a project manager.

Thomas Walker

Topic lead

Rebecca Albrow

Technical adviser

Donna Barnes

Project manager

ISBN: