Biomarkers for assessing acute kidney injury for people who are being considered for admission to critical care: a systematic review and cost-effectiveness analysis

Miriam Brazzelli,[1]* Lorna Aucott,[1]
Magaly Aceves-Martins,[1] Clare Robertson,[1]
Elisabet Jacobsen,[2] Mari Imamura,[1]
Amudha Poobalan,[3] Paul Manson,[1]
Graham Scotland,[1,2] Callum Kaye,[4] Simon Sawhney[3]
and Dwayne Boyers[2]

[1]Health Services Research Unit, University of Aberdeen, Aberdeen, UK
[2]Health Economics Research Unit, University of Aberdeen, Aberdeen, UK
[3]Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK
[4]Anaesthetics and Intensive Care Medicine, NHS Grampian, Aberdeen Royal Infirmary, Aberdeen, UK

*Corresponding author m.brazzelli@abdn.ac.uk

Declared competing interests of authors: Lorna Aucott is a member of the National Institute for Health Research Public Health Research funding committee (February 2017 to present).

Published January 2022
DOI: 10.3310/UGEZ4120

Scientific summary

Biomarkers for assessing acute kidney injury

Health Technology Assessment 2022; Vol. 26: No. 7
DOI: 10.3310/UGEZ4120

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Acute kidney injury is a common and serious complication that typically occurs in the context of an acute critical illness or during a postoperative period. It is associated with prolonged hospital stay, increased morbidity and increased mortality. Earlier detection of kidney injury may facilitate the adoption of strategies to preserve renal function and prevent further progression of kidney disease.

Currently, acute kidney injury diagnosis relies on a rise in serum creatinine levels and/or fall in urine output. Despite its widespread use in the monitoring of kidney health and disease, creatinine is an imperfect marker of kidney function because its level in the blood is not solely dependent on kidney function, and changes in creatinine lag behind reductions in kidney function. The limitations have led to the search for novel biomarkers that may detect kidney stress or damage earlier and more reliably.

Biomarker tests for acute kidney injury include neutrophil gelatinase-associated lipocalin (NGAL), which can be measured in urine or blood. NGAL is released from neutrophils and is induced by inflammation, indicating tubular injury. Another recent biomarker for acute kidney injury is NephroCheck® (Astute Medical, Inc., San Diego, CA, USA), which tests for the presence of tissue inhibitor of metalloproteinase-2 (TIMP-2) and insulin-like growth factor-binding protein 7 (IGFBP7) in the urine. Both TIMP-2 and IGFBP7 are cell-cycle arrest proteins and are used as markers of cellular stress in the early phase of tubular cell injury. Both NephroCheck and NGAL immunoassays are intended to be used in conjunction with existing clinical care. This assessment focuses specifically on the ARCHITECT® and Alinity i™ urine NGAL assays (Abbott Laboratories, Abbott Park, IL, USA), the BioPorto urine and plasma NGAL tests (BioPorto Diagnostics A/S, Hellerup, Denmark) and the NephroCheck test.

If these biomarkers allow early identification of patients at risk of acute kidney injury, they could enhance current acute kidney injury management by enabling timely measures to prevent progression of kidney injury and by informing decisions about the ‘step-down’ of low-risk patients to a lower level of hospital care, thereby reducing the use of hospital resources.

Objectives

The aim of this project was to summarise the current evidence on the clinical effectiveness and cost-effectiveness of the NephroCheck test, the ARCHITECT and Alinity i urine NGAL assays, and the BioPorto urine and plasma NGAL tests to assess the risk of acute kidney injury in critically ill hospitalised patients (adults and children) who are considered for admission to critical care.

Methods

Assessment of clinical effectiveness

Comprehensive electronic searches were undertaken to identify relevant reports of published studies up to June 2019.

The population of interest was critically ill people at risk of developing acute kidney injury who are considered for admission to critical care. Studies were eligible for inclusion only if they enrolled at least 100 participants at risk of acute kidney injury. The biomarkers under investigation were the

NephroCheck test, the ARCHITECT and Alinity i urine NGAL assays, and the urine and plasma BioPorto tests, used in conjunction with existing care. At present, there is no universally accepted reference standard for diagnosing acute kidney injury. The relevant comparator for this assessment was existing clinical criteria for monitoring serum creatinine levels and urine output, in conjunction with clinical judgement and in line with current clinical classification systems [risk, injury, failure, loss of kidney function, and end-stage disease (RIFLE); paediatric-modified RIFLE; the Acute Kidney Injury Network classification; and the Kidney Disease: Improving Global Outcomes (KDIGO) classification system] (see National Institute for Health and Care Excellence. *Acute Kidney Injury: Prevention, Detection and Management. Clinical guideline [CG169]*. London: National Institute for Health and Care Excellence; 2013).

The outcomes of interest were detection of acute kidney injury, prediction of acute kidney injury, prediction of mortality, prediction of the need for long-term renal replacement therapy and prediction of developing chronic kidney disease after acute kidney injury.

The quality of included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies, version 2, tool and the Prediction model Risk Of Bias ASsessment Tool.

Assessment of cost-effectiveness

A decision tree, using a linked-evidence approach based on observational data, was used to model the impact of test accuracy on acute kidney injury status and associated 90-day costs and outcomes, including need for intensive care unit care, length of hospital stay, 90-day mortality or development of chronic kidney disease. These observational associations necessitate causal assumptions, but, although a causal link between acute kidney injury and poor outcomes is plausible, the extent of this relationship is uncertain and controversial. These hypothesised links were tested extensively in sensitivity analyses.

The surviving proportion from each decision tree pathway at 90 days entered a Markov cohort model (starting age 63 years) with six mutually exclusive health states (outpatient follow-up, chronic kidney disease stages 1–4, end-stage renal disease without dialysis, end-stage renal disease with dialysis, transplantation and death). The cohort can enter the Markov model in the outpatient or chronic kidney disease states, with the starting proportions dependent on the experience of acute kidney injury up to 90 days.

The NHS-perspective costs (2018 Great British pounds values) in the first 90 days included costs of diagnostic biomarkers; costs of 3 days of a KDIGO care bundle for test-positive patients; and hospital costs, including days in intensive care unit, days on ward and need for acute renal replacement therapy. Markov health-state costs included post-discharge follow-up and costs of chronic kidney disease, end-stage renal disease, long-term dialysis, transplantation, immunosuppression and post-transplant follow-up.

Health-state utility values were sourced from the literature, based on the EuroQol-5 Dimensions questionnaire when possible, and were combined with mortality estimates for each health state to calculate quality-adjusted life-years. Utilities were applied separately for the duration of stay in an intensive care unit or hospital ward, or discharged up to day 90 with an additional utility decrement for acute renal replacement therapy and to each Markov health state. It was assumed that, following transplant recovery, utility reverted back to the outpatient post-discharge state. All utilities were adjusted for UK age- and sex-specific general population norms.

The model captured the cumulative cost and quality-adjusted life-year implications of transitions through the health states in annual cycles over a lifetime time horizon from an NHS perspective. All future costs and quality-adjusted life-years were discounted at 3.5% per annum. All analyses were reported probabilistically.

Results

Assessment of clinical effectiveness

A total of 56 studies, mainly prospective cohort studies, were included in the systematic review of clinical effectiveness. No studies addressing the impact of routine use of biomarkers on clinical outcomes were identified. Forty-six studies enrolled adults only, eight enrolled children only and two enrolled both adults and children. The total number of participants was 17,967, of whom 16,247 were adults (average age range 49 to 77 years) and 1720 were children (average age range 1 day to 5 years). The 46 studies on adults assessed patients after cardiac surgery ($n = 12$), non-surgical cardiac care ($n = 4$), major abdominal surgery ($n = 1$) or hepatobiliary surgery ($n = 1$); patients admitted to intensive care units ($n = 16$); patients with liver disease ($n = 5$, mainly cirrhosis), sepsis ($n = 2$) or chronic kidney disease ($n = 2$); and patients presenting to the emergency department ($n = 3$). Of the eight studies assessing children only, six assessed children (including neonates) undergoing cardiac surgery and two assessed children admitted to a paediatric or neonatal intensive care unit. The two studies that included both adults and children assessed patients undergoing cardiac surgery.

For statistical and cost-effectiveness analyses, participants were grouped into three categories according to clinical setting: patients undergoing cardiac surgery, patients undergoing major non-cardiac surgery and patients admitted to critical care (mixed patient population).

Of the 56 studies, 41 studied NGAL: 37 used urine NGAL assays and four used plasma NGAL assays. NephroCheck was assessed in eight studies. Seven studies provided data on more than one assay (six studies on urine NGAL and plasma NGAL assays; and one study on NephroCheck, urine NGAL and plasma NGAL assays). Of the NGAL studies, 24 used the urine NGAL ARCHITECT platform and 20 used the urine BioPorto NGAL assay. All 11 plasma NGAL studies used the BioPorto NGAL assay. No studies used the Alinity i NGAL platform.

The included studies were considered applicable to the remit of this assessment. The main sources of bias across diagnostic studies were the lack of information on blinding and the lack of a common threshold for NGAL. The statistical prediction models differed between prediction studies and often were not sufficiently detailed.

Few studies assessed patients after cardiac surgery or major non-cardiac surgery. The results of the meta-analyses of sensitivity and specificity estimates suggest that the biomarkers under investigation may have a role in the detection of acute kidney injury in patients already admitted to critical care. The NephroCheck test at a common threshold of 0.3 ng/ml^2/1000 had the higher pooled sensitivity (0.83), but the worst pooled specificity (0.51), whereas the ARCHITECT urine NGAL and the BioPorto urine NGAL tests had slightly lower pooled sensitivity estimates (0.70 and 0.72, respectively), but better pooled specificity estimates (0.72 and 0.87, respectively). The BioPorto urine NGAL pooled sensitivity was similar to that of the BioPorto plasma NGAL (0.72 and 0.76, respectively), whereas the pooled specificity was better for the BioPorto urine NGAL than for the BioPorto plasma NGAL (0.87 and 0.67, respectively). NGAL thresholds varied across studies. The biomarkers had a similar performance across all clinical settings.

Although these findings show some diagnostic usefulness of biomarkers, this should be tempered by the considerable heterogeneity observed across studies.

Moreover, for studies with a small number of acute kidney injury events, the relationship between sensitivity and specificity estimates appeared to be quite different from that of studies for which prevalence of acute kidney injury events was higher.

There was an indication that the addition of biomarkers to existing clinical models might improve the prediction of relevant clinical outcomes; however, few studies were available for each biomarker in each clinical setting, and studies varied substantially in terms of study characteristics and statistical methods used to assess prediction, thereby hindering any reliable conclusion.

In general, studies varied considerably in terms of clinical setting, NGAL threshold levels, time of sample collection, definition of acute kidney injury, time of acute kidney injury diagnosis, number of acute kidney injury events and assay platforms. Therefore, we have limited confidence in the validity and reliability of the findings.

Results of the cost-effectiveness model (including sensitivity analyses)

Published data show that NephroCheck-guided implementation of a KDIGO care bundle may avert acute kidney injury. However, no such data exist for the NGAL tests. Therefore, two base-case analyses were considered. Base case 1 can be considered an optimistic scenario for the NGAL biomarkers: assuming that all NGAL tests are equally as effective as NephroCheck in terms of the potential to avert acute kidney injury. Base case 2 can be considered a more conservative analysis. It assumes, in the absence of evidence, that only NephroCheck can avert acute kidney injury, but that all tests have the potential to reduce acute kidney injury severity if it occurs.

Fifteen scenario analyses were conducted for each potential base case, ranging from a set of optimistic assumptions whereby biomarker-guided care bundles led to substantial improvements in health outcomes (need for intensive care unit, hospital length of stay, chronic kidney disease, mortality) to a set of more conservative assumptions whereby change in acute kidney injury status had no effect on health outcomes.

Incremental cost-effectiveness ratios were highly uncertain, and subject to wide variation depending on the set of scenarios chosen. The probability of cost-effectiveness at an incremental cost-effectiveness ratio of < £20,000 per quality-adjusted life-year gained for scenarios in which all NGAL biomarkers were assumed to be equally as effective as NephroCheck in preventing acute kidney injury ranged from 0% to 15% (NephroCheck), from 0% to 55% (BioPorto urine NGAL), from 0% to 2% (ARCHITECT urine NGAL) and from 0% to 48% (BioPorto plasma NGAL). BioPorto urine NGAL was usually the test associated with the greatest probability of cost-effectiveness, albeit with great uncertainty, when compared with standard care. This was because the BioPorto urine NGAL biomarker was estimated to have slightly better diagnostic test accuracy data from the meta-analysis and incurred slightly lower test costs than the comparators. However, there was substantial uncertainty in diagnostic accuracy information, driven by substantial study heterogeneity. The cost-effectiveness results should therefore be interpreted with caution.

When it was assumed that NGAL biomarkers could not avert acute kidney injury but could only reduce its severity, the cost-effectiveness case for NephroCheck improved substantially, while remaining highly uncertain, with a probability of cost-effectiveness ranging from 0% to 99% across the explored scenarios.

Discussion

Strengths and limitations of the analyses, and uncertainties

The methods used to conduct this assessment were detailed, thorough and in line with current methodological standards.

The main limitations of the clinical effectiveness assessment were as follows:

- considerable clinical and statistical heterogeneity in the diagnostic and prediction analyses
- use of an imperfect reference standard for detection of acute kidney injury (clinical assessment based on serum creatinine levels and urine output)
- variation in the use of the NGAL assays and lack of a common threshold for identification of acute kidney injury
- uncertainty regarding the best timing of biomarker measurements
- variation in acute kidney injury prevalence across studies, with a very small number of acute kidney injury events in some studies
- lack of data on the impact of the routine use of the biomarkers on health outcomes.

The majority of the included studies were conducted outside the UK and assessed hospitalised patients admitted to critical care, with large variation in the delivery of critical and intensive care across the world. There is great uncertainty in how well findings of studies that are predominantly conducted outside the UK, based in intensive care, and heterogeneous, could be applied to a UK clinical scenario of people at risk of acute kidney injury who do not currently receive critical care.

With regard to the economic modelling, we identified three key areas of uncertainty, which mirror those identified for the clinical effectiveness assessment and limit the robustness of the cost-effectiveness results:

1. lack of direct evidence on the impact of the use of the biomarkers on health outcomes
2. heterogeneity in the diagnostic accuracy data (including uncertainty in the prevalence of acute kidney injury in a broad, poorly defined population)
3. uncertainty around the impact of an NGAL-guided implementation of a KDIGO care bundle on the frequency and severity of acute kidney injury.

Given these uncertainties, the results of the cost-effectiveness modelling were largely speculative and should be interpreted with caution. Although we conducted extensive probabilistic analyses for all scenario analyses, these may still not capture the full magnitude of uncertainty faced in the implementation of these biomarkers in clinical practice.

Generalisability of the findings

Owing to the limitations listed previously, it is unclear how the findings of this assessment can be generalised to current UK practice.

Conclusions

Future studies should evaluate the targeted use of the biomarkers among specific clinical populations and in circumstances where there is potential for benefit with a plausible and feasible intervention. They should focus on the assessment of the impact of routine biomarker use on mortality, major clinical adverse events, modification of clinical care, and resource use.

There is also a need to harmonise the methods and platforms for collection, and handling and storage of urine and plasma biomarker samples, as well as reporting of biomarkers' concentrations (units of measurement).

Study registration

This study is registered as PROSPERO CRD42019147039.

Funding

This project was funded by the National Institute for Health Research (NIHR) Evidence Synthesis programme and will be published in full in *Health Technology Assessment*; Vol. 26, No. 7. See the NIHR Journals Library website for further project information.

Health Technology Assessment

NICE TAR and DAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.014

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This report

The research reported in this issue of the journal was commissioned and funded by the Evidence Synthesis Programme on behalf of NICE as project number NIHR128897. The protocol was agreed in May 2019. The assessment report began editorial review in November 2019 and was accepted for publication in November 2020. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).