Biomarkers for the diagnosis and risk stratification of acute kidney injury: A systematic review

SG Coca[1,2], R Yalavarthy[3], J Concato[1,2] and CR Parikh[1,2]

[1]Clinical Epidemiology Research Center, Veterans Affairs Medical Center, West Haven, Connecticut, USA; [2]Department of Medicine, Yale University School of Medicine, New Haven, Connecticut, USA and [3]Denver Health Medical Center and University of Colorado at Denver and Health Sciences Center, Denver, Colorado, USA

The diagnosis of acute kidney injury (AKI) is usually based on changes in serum creatinine, but such measurements are a poor marker of acute deterioration in kidney function. We performed a systematic review of publications that evaluated the accuracy and reliability of serum and urinary biomarkers in human subjects when used for the diagnosis of established AKI or early AKI, or to risk stratify patients with AKI. Two reviewers independently searched the MEDLINE and EMBASE databases (January 2000–March 2007) for studies pertaining to biomarkers for AKI. Studies were assessed for methodologic quality. In total, 31 studies evaluated 21 unique serum and urine biomarkers. Twenty-five of the 31 studies were scored as having 'good' quality. The results of the studies indicated that serum cystatin C, urine interleukin-18 (IL-18), and urine kidney injury molecule-1 (KIM-1) performed best for the differential diagnosis of established AKI. Serum cystatin C and urine neutrophil gelatinase-associated lipocalin, IL-18, glutathione-*S*-transferase-π, and γ-glutathione-*S*-transferase performed best for early diagnosis of AKI. Urine *N*-acetyl-β-D-glucosaminidase, KIM-1, and IL-18 performed the best for mortality risk prediction after AKI. In conclusion, published data from studies of serum and urinary biomarkers suggest that biomarkers may have great potential to advance the fields of nephrology and critical care. These biomarkers need validation in larger studies, and the generalizability of biomarkers to different types of AKI as well as the incremental prognostic value over traditional clinical variables needs to be determined.

Kidney International (2008) **73,** 1008–1016; doi:10.1038/sj.ki.5002729; published online 19 December 2007

KEYWORDS: acute renal failure; clinical epidemiology; creatinine; mortality; renal injury; diagnosis

Correspondence: *C Parikh, Section of Nephrology, VAMC and Yale University School of Medicine, Clinical Epidemiology Research Center, 950 Campbell Ave, Mail Code 151B, Building 35A, Room 219, West Haven, Connecticut 06516, USA. E-mail: chirag.parikh@yale.edu*

Received 6 July 2007; revised 11 October 2007; accepted 16 October 2007; published online 19 December 2007

The diagnosis of acute kidney injury (AKI—formally known as acute renal failure) is usually based on either an elevation of serum creatinine or the detection of oliguria.[1] Serum creatinine is a poor marker of early renal dysfunction, because serum concentration is greatly influenced by numerous non-renal factors (such as body weight, race, age, gender, total body volume, drugs, muscle metabolism, and protein intake).[2] The utility of serum creatinine is worse in AKI, because the patients are not in steady state; hence, serum creatinine lags far behind renal injury. Thus, substantial rises in serum creatinine are often not witnessed until 48–72 h after the initial insult to the kidney.[1,3] In addition, significant renal disease can exist with minimal or no change in creatinine because of renal reserve, enhanced tubular secretion of creatinine, or other factors.[4,5] A 'troponin-like' biomarker of AKI that is easily measured, unaffected by other biological variables, and capable of both early detection and risk stratification would substantially assist the diagnosis of AKI.

The American Society of Nephrology has designated the development of biomarkers for early detection of AKI as a top research priority.[6] Several biomarkers of AKI have been identified over the past few years that are elevated in ischemic renal injury in experimental animals, and also in humans with clinical AKI, in some cases prior to a 'gold standard' diagnostic threshold (for example, rise in serum creatinine by 50%).[7,8] These biomarkers include both serum tests such as cystatin C,[9] and urinary tests such as interleukin-18 (IL-18)[10] and neutrophil gelatinase-associated lipocalin (NGAL),[11] along with several others. Uncertainty still exists, however, as to whether these biomarkers possess adequate prognostic accuracy for both established AKI and for early detection of AKI. Limited data are available that directly compare the performance of these biomarkers as tests for diagnosis of AKI, and their consistency across certain subgroups of patients (for example, post-cardiac surgery, sepsis, post-kidney transplant). Finally, the differences in the reporting criteria for 'positivity' have made it difficult to compare the performance of these various biomarkers directly.

Our objectives were to evaluate, in human subjects, the accuracy and reliability of serum and/or urinary biomarkers for the diagnosis of established AKI, for the early diagnosis of AKI, and for risk stratification of AKI.

RESULTS

The combined search identified a total of 830 citations, of which 715 were judged ineligible after title and abstract review (Figure 1). The major reasons for exclusion were study populations without AKI or acute renal failure, non-serum or non-urine biomarker of diagnosis of AKI, and duplicate publications. Full text analysis of the remaining 115 articles led to 31 studies[9–36] meeting inclusion criteria (Tables 2–4). The methodological quality based on scoring of the 10 validity criteria taken from the STARD (Standards for Reporting of Diagnostic Accuracy) recommendations is listed in Table 1. Only tests with results that were scored of good or fair quality are described in the text below.

Biomarker performance

Diagnosis of established AKI. *Serum biomarkers.* Cystatin C, a marker of glomerular filtration, performed extremely well for identifying established AKI (receiver operating curve (area under the curve (AUC)) 0.88–0.97)[13,15–17] (Table 2). Serum NGAL, a marker of renal tubular inflammation rather than glomerular filtration, correlated well with acute changes in serum creatinine in two studies.[19,26] Finally, carbamylated hemoglobin performed quite well in differentiating AKI from elevated creatinine due to chronic kidney disease (CKD).[18]

Urine biomarkers. Urine IL-18 (a marker of renal tubular inflammation) concentrations were significantly greater in patients with acute tubular necrosis vs all other types of patients (healthy controls, prerenal azotemia, UTI, and CKD).[20] The AUC of IL-18 for the diagnosis of established AKI (acute tubular necrosis) was nearly flawless (0.95) in this cohort.

Kidney injury molecule-1 (KIM-1) is renal tubular protein that has been demonstrated to be elevated in experimental animals with AKI. Clinical studies reveal that KIM-1 is highly effective at distinguishing true acute tubular necrosis from other types of renal injury (including CKD) or controls.[22,37] In the first published report, a 1 U increase in KIM-1 was associated with 12-fold higher odds for ischemic acute tubular necrosis.[40] In the second report, the AUC of KIM-1 for the diagnosis of AKI was 0.90.[37] Two other biomarkers (enzymes that are normally filtered and degraded) were also tested in this cohort. *N*-acetyl-β-D-glucosaminidase (NAG)

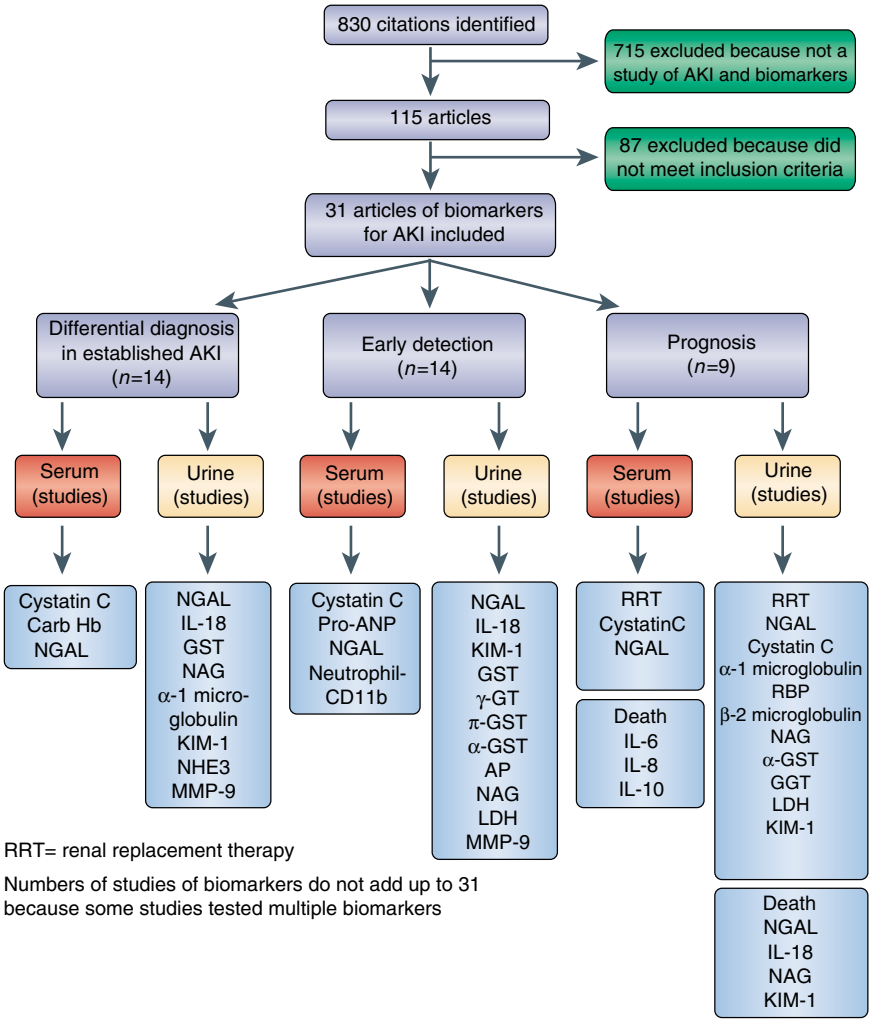

Figure 1 | Selection of Studies. RRT, renal replacement therapy.

Table 1 | Scoring system for validity used in this systematic review

Validity Criterion	Explanation	Scoring	Comments
Participant recruitment	Was recruitment based on presenting symptoms, results from previous tests, or fact that participants received index tests?	Presenting sx=1 Previous tests or index tests=0	Based on presenting symptoms in all 31 studies.
Participant sampling	Was it study population or a convenience sample or a consecutive series?	Consecutive series=1 Convenience sample=0	Based on convenience sample in 5[10,19,20,22,25,37] studies.
Data collection	Was data collection planned before the index test and reference standard were performed prospectively or retrospectively?	Prospective=1 Retrospective=0	Planned and performed prospectively in all 31 studies.
Reference standard	Was the rationale for the reference standard stated?	Stated=1 Not stated=0	Not stated for two[12,18] studies.
Materials and methods	Were technical specifications of material and methods stated including how and when measurements were taken?	Stated=1 Not stated=0	Stated in all but one[19] of the articles.
Index test	Were the definitions of and rationales for the units, cutoffs, and/or categories of the results of the index tests stated?	Stated=1 Not stated=0	Not stated in two[12,19] of the 31 articles.
Blinding	Were readers of index test and reference standard blinded?	Blinded=1 Not blinded or not stated=0	Stated that the readers of index test and reference standard were blinded in 17[10,11,19,29,30,35–39] articles.
Completion	Was the number of participants that did not undergo index tests (no. of tests vs sample size) stated?	Stated=1 Not stated=0	Stated in all but one[19] of the studies.
Time interval	Was the time interval from index test to reference standard stated?	Stated=1 Not stated=0	The time interval between index test and reference standard (clinical diagnosis of AKI or severity end point such as dialysis or death) was not stated in five[19,20,25,34,36] articles.
Distribution of severity of disease	Was there a representative distribution of severity of disease? (mild, moderate, severe AKI; non-oliguric vs oliguric)	Yes=1 No=0	A broad distribution of disease severity was found in all but four [12,25,27,37] of the studies.

Table 2 | Studies of biomarkers for diagnosis of established AKI

References	Biomarker	Clinical setting	Subjects	Sensitivity/ specificity	Area under ROC	LR	Quality score
Serum							
Abu-Omar *et al.*[12]	Cystatin C	Cardiac surgery	60	NR	NR	N/A	6
Biancofiore *et al.*[13]	Cystatin C	Post-OLT	68	0.967/0.85	NR	6.4	9
Benohr *et al.*[14]	Cystatin C	Cisplatin Rx	41	NR	NR	N/A	9
Stabuc *et al.*[15]	Cystatin C	Cisplatin Rx	72	0.87/1.0	0.967	Infinity	9
Zhu *et al.*[16]	Cystatin C	Cardiac surgery	60	NR	0.876	N/A	9
Villa *et al.*[17]	Cystatin C	ICU	50	NR	0.927	N/A	9
Wynckel *et al.*[18]	Carbamylated Hb	Hospitalized	41	0.94/0.92	NR	11.8	8
Mori *et al.*[19]	NGAL	Hospitalized	37	NR	NR	N/A	5
Bachorzewska-Gajewska *et al.*[26]	NGAL	PCI	35	NR	NR	NR	9
Urine							
Mori *et al.*[19]	NGAL	Hospitalized	46	NR	NR	N/A	5
Bachorzewska-Gajewska *et al.*[26]	NGAL	PCI	35	NR	NR	N/A	9
Parikh *et al.*[20]	IL-18	Hospitalized	72	0.85/0.88	0.95	7.1	7
Boldt *et al.*[21]	GST, NAG, α-1-M	Cardiac surgery	80	NR	NR	N/A	9
Han *et al.*[22]	KIM-1	Hospitalized	32	NR	NR	12.4†	8
Han *et al.*[37]	KIM-1	Hospitalized	74	NR	0.90	N/A	8
Han *et al.*[37]	MMP-9	Hospitalized	74	NR	0.74	N/A	8
Han *et al.*[37]	NAG	Hospitalized	74	NR	0.97	N/A	8
du Cheyron *et al.*[23]	NHE3	ICU	68	NR	NR	N/A	9

α-1-M, alpha-1 microglobulin; AKI, acute kidney injury; AP, alkaline phosphatase; CKD, chronic kidney disease; ESRD, end stage renal disease; GFR, glomerular filtration rate; GGT, γ-glutamyltransferase; GST, glutathione-*S*-transferase; GT, glutamyl transpeptidase; ICU, intensive care unit; KIM-1, kidney injury molecule-1; LDH, lactate dehydrogenase; LR, likelihood ratio; MMP-9, matrix metalloproteinase-9; N/A, not applicable; NAG, *N*-acetyl-β-D-glucosaminidase; NGAL, neutrophil gelatinase-associated lipocalin; NHE3, sodium hydrogen exchanger 1; NR, not reported; OLT, orthotopic liver transplantation; PCI, percutaneous coronary intervention; ProANP, prohormone of atrial natriuretic peptide; Rx, treatment; RIFLE, Risk Injury Failure Loss End-stage Renal Disease; ROC, receiver operating curve.

was excellent at identifying cases of AKI (AUC 0.97), yet matrix metalloproteinase-9 (MMP-9) did not fare as well to identify AKI (AUC 0.74).[37] In the same study, the combination of KIM-1, NAG, and MMP-9 yielded perfect results (AUC 1.0) for diagnosing AKI.

Sodium–hydrogen exchanger isoform-3 (NHE-3) is an abundant proximal tubular protein that is shed into the urine after tubular injury. One clinical study[23] demonstrated that urinary NHE-3 was elevated by sixfold in patients with AKI compared to controls with prerenal azotemia.

Early diagnosis of AKI. *Serum biomarkers*. The two studies[9,24] of cystatin C that were of good methodologic quality demonstrated excellent accuracy for the early diagnosis of AKI at 24 and 48 h before the clinical diagnosis of AKI,[9] or in another study performed on days 1–3 of admission to the medical intensive care unit (Table 3).[24] Serum NGAL was significantly higher in children with AKI following cardiac surgery compared to controls, and possessed excellent specificity, but its sensitivity for the early diagnosis of AKI was low (Table 3).[11] In a study of 75 adult patients undergoing cardiac surgery, peri-operative elevations of CD11b (an inflammatory marker) had high odds for the development of AKI (Table 3).[27] The peak CD11b increase also correlated with the percentage change in creatinine.

Urine biomarkers. The two biomarkers studied most often for early diagnosis of AKI have been urine NGAL and IL-18. In a study of children undergoing cardiac surgery,[11] the urinary concentration of NGAL 2 h post-surgery possessed nearly 100% accuracy for correctly identifying AKI 24–72 h after surgery (Tables 3 and 5). In contrast, subsequent studies have not yielded results as robust as those witnessed in the initial study. In a cohort of adults undergoing cardiac surgery,[28] the best AUC for urinary NGAL concentration was only 0.80 at a time point of 18 h after surgery. Other time points, both immediately after surgery (2 and 12 h) and 24 h after surgery, did not possess diagnostic accuracy as high (Table 5). Similar, less robust results were demonstrated in a cohort of critically ill children,[38] in which urine NGAL 48 h prior to a clinical diagnosis of AKI was only moderately accurate at predicting AKI (Table 3). In a different population (post-kidney transplant), urine NGAL was quite successful at predicting the outcome of AKI (dialysis-requiring delayed graft function).[29] Thus, NGAL performed extremely well in two specific populations

Table 3 | Studies of biomarkers for the early diagnosis of AKI

References	Biomarker	Clinical setting	Subjects	Sensitivity/ specificity	Area under ROC	Positive LR	Quality score
Serum							
Herget-Rosenthal *et al.*[9]	Cystatin C	ICU	85	0.82/0.95	NR	16.4	9
Ahlstrom *et al.*[24]	Cystatin C	ICU	202	NR	0.901	N/A	9
Mazul-Sunko *et al.*[25]	Cystatin C	ICU	29	0.5/0.5	NR	1.0	6
Mazul-Sunko *et al.*[25]	ProANP(1-98)	ICU	29	NR/NR	NR	N/A	6
Mishra *et al.*[11]	NGAL	Cardiac surgery in children	71	0.70/0.94	NR	11.6	10
Rinder *et al.*[27]	Neutrophil CD11b	Cardiac surgery	75	NR	NR	62.6†	8
Urine							
Mishra *et al.*[11]	NGAL	Cardiac surgery in children	71	1.0/0.98	0.998	50.0	10
Wagener *et al.*[28]	NGAL	Cardiac surgery	81	0.73/0.78	0.8	3.3	9
Zappitelli *et al.*[38]	NGAL	Critically ill children	140	0.77/0.72	0.78	2.75	10
Parikh *et al.*[29]	NGAL	Post-transplant	63	0.9/0.83	0.9	5.3	10
Parikh *et al.*[10]	IL-18	ICU	138	0.5/0.85	0.73	3.3	9
Parikh *et al.*[29]	IL-18	Post-transplant	63	NR	0.9	N/A	10
Parikh *et al.*[30]	IL-18	Cardiac surgery	71	0.5/0.94	0.75	8.3	10
Han *et al.*[40]	IL-18	Critically ill children	137	0.25/0.81	0.54	1.3	10
Han *et al.*[37]	KIM-1	Cardiac surgery	40	0.74/0.90	0.83	3.16	8
Han *et al.*[37]	NAG	Cardiac surgery	40	1.0/0.3	0.69	1.43	8
Han *et al.*[37]	MMP-9	Cardiac surgery	40	NR	0.50	N/A	8
Eijkenboom *et al.*[31]	GST	Cardiac surgery	84	NR	NR	N/A	9
Westhuyzen *et al.*[32]	γ-GT	ICU	26	1.0/0.9	0.95	10.0	9
Westhuyzen *et al.*[32]	π-GST	ICU	26	1.0/0.9	0.929	10.0	9
Westhuyzen *et al.*[32]	α-GST	ICU	26	0.75/0.9	0.893	7.5	9
Westhuyzen *et al.*[32]	AP	ICU	26	0.5/0.95	0.863	10.0	9
Westhuyzen *et al.*[32]	NAG	ICU	26	1.0/0.81	0.845	5.3	9
Westhuyzen *et al.*[32]	LDH	ICU	26	0.5/0.95	0.688	10.0	9

AKI, acute kidney injury; AP, alkaline phosphatase; GST, glutathione-*S*-transferase; GT, glutamyl transpeptidase; ICU, intensive care unit; IL, interleukin; KIM-1, kidney injury molecule-1; LDH, lactate dehydrogenase; LR, likelihood ratio; MMP-9, matrix metalloproteinase-9; N/A, not applicable; NAG, *N*-acetyl-β-D-glucosaminidase; NGAL, neutrophil gelatinase-associated lipocalin; NR, not reported; ProANP, prohormone of atrial natriuretic peptide; ROC, receiver operating curve.
†Odds ratio rather than true LR.

(children undergoing cardiac surgery and patients receiving kidney transplantation), but did not perform as well in adults undergoing cardiac surgery or in critically ill children.

Four studies[10,29,30,39] of urinary IL-18 as an early predictive biomarker of AKI are included in this review. The AUCs for IL-18 in these four studies varied markedly (0.54–0.9) for the early diagnosis of AKI. In general, IL-18 demonstrated low sensitivity but high specificity (Table 3). As noted for NGAL, the performance of IL-18 also is dependent on the timing of collection in relation to the exposure (for example, the stressor such as cardiac surgery) and the outcome (development of AKI).[10,30] Of note, IL-18 concentration remained an independent predictor for AKI even after adjustment for several clinical variables.[10,39] In summary, elevations in urine IL-18 concentration are rarely false positive elevations, however, many patients with AKI do not have elevations in urine IL-18.

The ability of KIM-1 to predict clinical AKI early has only been evaluated in one published study to date. It appears that KIM-1 is not as accurate in early prediction of AKI, but is better at identifying established AKI (Table 5).[37]

Prognosis of AKI (severity of AKI). Several of the same biomarkers mentioned above were evaluated not only for diagnosis of AKI but also for risk stratification and prognosis of AKI.

Serum biomarkers. In one study, on the same day that the criterion for meeting the RIFLE-R definition of AKI (rise in serum creatinine $\geqslant 50\%$) was observed, the positive predictive value of elevations in serum cystatin C for the need for renal replacement therapy (RRT) was 78% and the AUC was 0.76 (Table 4).[9] Serum NGAL levels were unable to discriminate between children who did not require RRT and those who subsequently required RRT, regardless of various cutoff values examined.[33]

Urine biomarkers. Urinary NGAL within the first 5 days of hospitalization possessed high sensitivity but low specificity for the subsequent need for dialysis in a study of pediatric hemolytic uremic syndrome.[33] Another study[30] demonstrated that urine NGAL at 4 h after cardiac surgery correlated, albeit weakly, with the number of days of AKI ($r^2 = 0.22$; $P = 0.005$). In a different study, however, urine NGAL concentration was minimally useful for predicting persistent AKI (AUC 0.63) or for worsening AKI (AUC 0.61).[38] Thus, the utility of NGAL to predict severity of AKI is certainly mixed.

Similar types of weakly positive results for the ability of urine IL-18 to predict severity of AKI have been demonstrated. Urine IL-18 at 4 h after cardiac surgery weakly correlated with number of days with AKI ($r^2 = 0.18$; $P = 0.01$).[30] In another study, the ability for IL-18 to predict severe AKI (duration >48 h) varied depending on the cutoff used.[39]

While the inflammatory biomarkers (NGAL, IL-18) did not fare extremely well in predicting severe AKI, it appears that tubular proteins and enzymes may be of greater yield for risk-stratification. In 73 patients with non-oliguric AKI, four

Table 4 | Studies of biomarkers for predictors of severity of AKI

References	Biomarker	Clinical setting	Definition of severe AKI	Sensitivity/ specificity	Area under ROC	LR	Quality score
Serum							
Herget-Rosenthal *et al.*[9]	Cystatin C	ICU	RRT	0.82/0.93	0.76	11.7	9
Ahlstrom *et al.*[24]	Cystatin C	ICU	Death	NR	0.624	N/A	9
Trachtman *et al.*[33]	NGAL	HUS	RRT	NR	NR	NR	9
Simmons *et al.*[34]	IL-6	ICU	Death	NR	NR	1.65†	9
Simmons *et al.*[34]	IL-8	ICU	Death	NR	NR	1.54†	9
Simmons *et al.*[34]	IL-10	ICU	Death	NR	NR	1.34†	9
Urine							
Parikh *et al.*[10]	IL-18	ICU	Death	NR	NR	1.6†	9
Han *et al.*[40]	IL1-8	Critically ill children	Duration >48 h	0.53/0.71	NR	1.8	10
Trachtman *et al.*[33]	NGAL	HUS	RRT	0.9/0.54	NR	2.0	9
Zappitelli *et al.*[38]	NGAL	Critically ill children	Duration >48 h	NR	0.63	NR	10
Herget-Rosenthal *et al.*[35]	Cystatin C	ICU	RRT	0.92/0.83	0.92	5.4	10
Herget-Rosenthal *et al.*[35]	α-Microglobulin	ICU	RRT	0.88/0.81	0.86	4.6	10
Herget-Rosenthal *et al.*[35]	Retinol-binding protein	ICU	RRT	NR	0.8	N/A	10
Herget-Rosenthal *et al.*[35]	β-2 microglobulin	ICU	RRT	NR	0.51	N/A	10
Herget-Rosenthal *et al.*[35]	NAG	ICU	RRT	0.85/0.62	0.81	2.2	10
Herget-Rosenthal *et al.*[35]	α-GST	ICU	RRT	NR	0.64	N/A	10
Herget-Rosenthal *et al.*[35]	GGT	ICU	RRT	NR	0.64	N/A	10
Herget-Rosenthal *et al.*[35]	LDH	ICU	RRT	NR	0.59	N/A	10
Han *et al.*[37]	NAG	Hospitalized	RRT or death	NR	0.71	3†	9
Han *et al.*[37]	KIM-1	Hospitalized	RRT or death	NR	0.61	1.4†	9

AKI, acute kidney injury; GGT, γ-glutamyltransferase; GST, glutathione-*S*-transferase; GT, glutamyl transpeptidase; HUS, hemolytic uremic syndrome; ICU, intensive care unit; IL, interleukin; KIM-1, kidney injury molecule-1; LDH, lactate dehydrogenase; LR, likelihood ratio; MMP-9, matrix metalloproteinase-9; N/A, not applicable; NAG, *N*-acetyl-β-D-glucosaminidase; NGAL, neutrophil gelatinase-associated lipocalin; NR, not reported; ROC, receiver operating curve; RRT, renal replacement therapy.

Table 5 | Relationship between timing of biomarker collection and test characteristics (when reported)

References	Biomarker	Cutoff for biomarker	AKI definition (change in creatinine)	Setting	Timing	Sensitivity/ specificity	Area under ROC	LR
Early diagnosis of AKI								
Serum								
Herget-Rosenthal *et al.*[9]	Cystatin C	>50% rise	>50% rise	ICU	2 days prior to clinical AKI	0.55/0.95	0.82	11.0
					1 day prior to clinical AKI	0.82/0.95	0.97	16.4
		>100% rise	>100% rise	ICU	2 days prior to clinical AKI	0.39/1.0	0.92	Infinity
					1 day prior to clinical AKI	0.76/1	0.98	Infinity
		>200% rise	>200% rise	ICU	2 days prior to clinical AKI	0.4/1	0.97	Infinity
					1 day prior to clinical AKI	0.85/1	0.99	Infinity
Ahlstrom *et al.*[24]	Cystatin C	NR	>300% rise	ICU	Day 1 after admission to ICU	NR	0.885	N/A
					Day 2 after admission to ICU	NR	0.893	N/A
					Day 3 after admission to ICU	NR	0.901	N/A
Urine								
Mishra *et al.*[11]	NGAL	50 $\mu g\,l^{-1}$	>50% rise	Cardiac surgery	2 h post-surgery	1.0/0.98	0.998	50.0
					4 h post-surgery	0.95/1.0	1	Infinity
Wagener *et al.*[28]	NGAL	213 ng ml^{-1}	>50% rise	Cardiac surgery	1 h post-surgery	0.8/0.59	0.68	2.0
					3 h post-surgery	0.69/0.65	0.737	2.0
					18 h post-surgery	0.73/0.78	0.8	3.3
					24 h post-surgery	0.8/0.73	0.68	3.0
Parikh *et al.*[30]	IL-18	50 pg ml^{-1}	>50% rise	Cardiac surgery	4 h post-surgery	0.25/0.97	0.61	8.3
					12 h post-surgery	0.5/0.94	0.75	8.3
					24 h post-surgery	0.4/0.94	0.73	6.7
Parikh *et al.*[10]	IL-18	100 pg ml^{-1}	>50% rise	ICU	2 days prior to clinical AKI	NR	0.65	N/A
					1 day prior	0.5/0.85	0.73	3.3
Han *et al.*[37]	KIM-1	—		Cardiac surgery	2 h post-surgery	NR	0.57	N/A
		7 ng mg^{-1} creat	>50% rise		6 h post-surgery	0.85/0.21	0.52	1.1
		7 ng mg^{-1} creat	>50% rise		12 h post-surgery	0.32/0.90	0.83	3.2
		7 ng mg^{-1} creat	>50% rise		24 h post-surgery	0.35/0.90	0.78	3.5
		—			36 h post-surgery	NR	0.84	N/A
		—			48 h post-surgery	NR	0.81	N/A
Han *et al.*[37]	NAG	—		Cardiac surgery	2 h post-surgery	NR	0.65	N/A
		25 U mg^{-1} Ucr	>50% rise		6 h post-surgery	NR	0.58	1.4
		25 U mg^{-1} Ucr	>50% rise		12 h post-surgery	1.0/0.3	0.69	1.4
		25 U mg^{-1} Ucr	>50% rise		24 h post-surgery	0.85/0.25	0.70	1.1
		—	—		36 h post-surgery	NR	0.71	N/A
		—	—		48 h post-surgery	NR	0.66	N/A
Predictors of Severity of AKI								
Herget-Rosenthal *et al.*[9]	Cystatin C	>50% rise	RRT	ICU	2 days prior to clinical AKI	0.53/0.82	0.69	2.9
					1 day prior to clinical AKI	0.76/0.93	0.75	10.9
					Day of diagnosis of clinical AKI	0.82/0.93	0.76	11.7
Ahlstrom *et al.*[24]	Cystatin C	NR	Death	ICU	Day 1 after admission to ICU	NR	0.624	N/A
					Day 2 after admission to ICU	NR	0.598	N/A
					Day 3 after admission to ICU	NR	0.581	N.A

AKI, acute kidney injury; ICU, intensive care unit; IL, interleukin; KIM-1, kidney injury molecule-1; LR, likelihood ratio; N/A, not applicable; NAG, *N*-acetyl-β-D-glucosaminidase; NGAL, neutrophil gelatinase-associated lipocalin; NR, not reported; ROC, receiver operating curve.

urine biomarkers performed well for predicting RRT.[35] The four biomarkers were urinary cystatin C, α-1 microglobulin, NAG, and retinol-binding protein (AUCs 0.92, 0.86, 0.80, and 0.81, respectively). Four other biomarkers that were assessed did not perform as well (Table 4). In a different study of 201 hospitalized patients, NAG again proved to be of prognostic value, as a dose-dependent association for the prediction of the composite end point of dialysis or death was demonstrated.[36] In fact, patients with NAG concentrations in the upper quartile possessed ninefold higher odds for development of the composite end point compared to those in the lowest quartile, and this relationship remained significant after adjustment for co-variates. In this same study,[36] the association between the KIM-1 concentration and the composite end point (dialysis or death) was weaker (odds ratio 3.2 for highest vs lowest quartile) and was no longer significant after adjustment for co-variates.

Prognosis of AKI (AKI-associated death). Risk stratification of AKI can also be defined in terms of death, as opposed to the need for dialysis or duration of AKI. Some studies have tried to predict death in the setting of AKI via biomarkers.

Serum biomarkers. The results of serum cystatin C for predicting AKI-associated death were modest (Table 4).[24] The cytokines IL-6, IL-8, and IL-10 were significantly higher in non-survivors vs survivors in a cohort of adults with AKI, but after adjustment for co-variates, the statistically significant relationship was lost (Table 4).[34]

Urine biomarkers. Urine NGAL was not associated with mortality in the only study that examined this relationship.[38] IL-18 has been shown to predict mortality in adult[10] and pediatric critically ill populations.[39]

Finally, as mentioned above, the association and performance characteristics of NAG activity and KIM-1 level jointly and the composite outcome of dialysis and death were

examined in one study and added a modest amount of prognostic value to traditional clinical severity of illness variables.[36]

DISCUSSION

This systematic review is a novel evaluation of biomarkers for the diagnosis of AKI as reported in the literature. Before a new biomarker can be deemed clinically useful, it requires validation in multiple cohorts, it must be reliable, and it must provide incremental prognostic information over traditional markers and models. This systematic review finds that although some serum and urine biomarkers perform well for the early diagnosis of AKI, the diagnosis of established AKI, or risk stratification of AKI, methodological shortcomings exist in some of the studies, and the reliability and reproducibility of the findings have not yet been demonstrated.

The generalizability of the data to broad patient populations is also not yet manifest, as the test performance varied among different patient types (children vs adult; cardiac surgery vs transplant vs critically ill). Many studies excluded patients with pre-existing renal failure (CKD). In fact, 22 of the 31 studies excluded patients if they possessed underlying CKD. This approach again limits generalizability, because approximately 30% of patients admitted to intensive care units have pre-existing CKD,[41] and patients with CKD are at higher risk for development of AKI than those with normal baseline renal function.[42,43]

In addition, several studies did not provide sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under receiver operator curves of the biomarkers for the diagnosis or risk stratification of AKI. Many studies reported odds ratio or relative risk (RR). Reporting of odds ratio or relative risk is typical in epidemiologic studies to demonstrate the strength of association between risk factor and the disease. Use of these measures, however, is not adequate to determine the performance of a biomarker for classifying disease or predicting risks in people. Strong statistical associations (odds ratio or relative risk) between a biomarker and an outcome do not necessarily imply that the biomarker can discriminate between persons likely to have the outcome and those who do not.[44] Instead, the accuracy or validity of a marker for classifying persons is better summarized by reporting its true-positive fraction (or sensitivity) and its false-positive fraction (also known as 1−specificity). A perfect marker will have true-positive fraction = 1 and false-positive fraction = 0, but the criteria by which the marker is judged useful will depend entirely on the context in which it is to be used. For example, if one is using a biomarker to determine who will need an invasive procedure or risky treatment, then an extremely low false-positive fraction ($<2\%$) is needed to avoid placing patients at risk for these procedures or treatments. Thus, characterization of false-positive and false-negative errors is not equivalent and must be reported separately.

Our review found that some studies of biomarkers were biased in various ways. The most common bias we encountered was 'review bias,' as the examiner and the person who evaluated the experimental and reference test results were not stated to be blinded in 14 of the 31 studies. Second, 'selection bias' was present in nearly 20% of studies, as members of the cohort were enrolled non-consecutively. Third, 'workup detection bias' was a concern overtly in five studies, because the time interval between the experimental test (biomarker) and the disease (clinical AKI) was not stated. Although some studies measured biomarker concentrations at multiple intervals following the stressor (for example, cardiac surgery, intensive care unit admission), the interaction between these times and the degree of accuracy was not thoroughly explored and discussed in several studies. The importance of this issue is critical, as in clinical practice, it is unrealistic for a test to be utilized only at one precise point in time (for example, 24 h before a clinical diagnosis of AKI,[10] or 18 h after cardiac surgery[28]). Thus, when the results of all of the biomarker studies are scrutinized closely, certain patterns of biomarker elevation are evident (Table 5). In general, the biomarkers were most accurate when measured closer to the clinical diagnosis of AKI.

Aside from the limitations of the available studies on serum and urinary biomarkers, there are positive aspects that can be extended in future studies of biomarkers for AKI. Urine IL-18, KIM-1, and NAG performed extremely well in some studies evaluating the diagnosis of established AKI. Serum cystatin C and urine NGAL, IL-18, glutathione-*S*-transferase-π, and γ-glutathione-*S*-transferase performed best for facilitating early diagnosis of AKI and NAG. KIM-1 and IL-18 performed best for mortality risk prediction after AKI. These biomarkers can be considered candidates for phases 3 and 4 clinical biomarker studies.

It is unclear if biomarkers of AKI will have incremental benefit over traditional markers or variables before their adoption in the scientific and clinical community. Only a few studies ($n = 9$[10,25,29,30,32,34,36,38,39]) in this systematic review compared the individual biomarkers to traditional clinical variables or to clinical models that have been used for prediction of AKI, or its severity and outcome; and only six studies assessed the independent association of biomarkers with outcomes via multivariate analyses.[10,27,29,30,34,39] It should be noted, however, that the two most widely used biomarkers in clinical practice (serum troponin for acute myocardial infarction and brain natriuretic peptide for congestive heart failure) are used largely in isolation of other clinical variables.

Biomarkers that have performed well to date will need to be validated in a broad spectrum of patient populations (cardiac surgery, critically ill, post-transplant, underlying CKD). In addition, the patterns of the timing of collection of the biomarker in relationship to the clinical outcome will need to be delineated more clearly. Finally, the combination of biomarker panels (such as the panel for acute MI composed of CK, CK-MB, and troponin) will likely need

creation and validation. We anticipate that a battery of biomarkers will need to be implemented to accurately identify patients with established AKI (for example, serum cystatin C, urine IL-18, and urine KIM-1), diagnose AKI early (for example, serum cystatin C, and urine NGAL), and risk-stratify patients for the need for dialysis or death (for example, urine KIM-1, NAG, and IL-18). If these panels are validated and show robust results, an opportunity may exist to conduct clinical trials of therapeutics for the treatment of AKI and thus improve outcomes of AKI in the near future.

MATERIALS AND METHODS

Studies eligible for review

Studies were eligible if they were prospective or retrospective cohort studies, case–control studies, or randomized controlled trials evaluating the use of serum or urinary biomarkers for the diagnosis or risk stratification of AKI. Studies must have contained 20 or more human subjects.

Finding relevant studies

We searched the MEDLINE and EMBASE databases (2000–May 2007), utilizing the search terms 'acute renal failure' or 'acute kidney injury' or 'creatinine blood level' in conjunction with 'biomarker' or 'biologic marker,' with the following limits: humans, 2000–2007, diagnosis (sensitivity), adults. Trials that attempted to ascertain the diagnostic value of any serum or urinary biomarker were reviewed. We identified potentially relevant studies using a manual search of references from all eligible studies, review articles, Science Citation Index Expanded on the Web of Science, and through searching the top 50 citations for each paper through the 'related articles' feature of PubMed. The search was restricted to the English language.

Quality assessment

Two reviewers (SGC and RY) independently searched trials for inclusion and assessed methodological quality; the two reviewers selected all studies suitable for review. Disagreements were discussed until the decision to include or exclude an individual study was unanimous. Studies were assessed for validity using a modified checklist of the STARD criteria.[45] Studies were scored as 'good' quality if the score was $\geqslant 9$, 'fair' quality if the score was 7–8, and 'poor' quality if the score was $\leqslant 6$. The STARD initiative developed a set of 25 criteria for reporting of studies of diagnostic accuracy during a consensus meeting on September 16 and 17, 2000.[45] The original STARD checklist contained 25 items. For the purposes of our review, however, we limited our quality assessment to 10 out of the 25 criteria that addressed issues of validity (Table 1).

Data abstraction

Data were extracted using prepared data extraction forms. Review Manager software (Revman 2003; Version 4.2) was used for statistical analyses. Studies were separated into three groups: those evaluating biomarkers for the diagnosis of established AKI, those evaluating biomarkers for the early diagnosis of AKI, and those evaluating biomarkers for the prognosis or severity of AKI. Sensitivity and specificity were abstracted from the trials and converted into likelihood ratios, such that the likelihood ratios were equal to sensitivity/(1−specificity). The receiver operator curve, as the prime statistic for evaluation of the accuracy of continuous biomarkers for outcomes, was abstracted, because it describes a whole set of potential result combinations (for example, true-positives, false-positives) for different cutoff values, does not depend on how the marker is coded, and provides a natural common scale for comparing different markers even when they are measured in completely different units. An AUC value of 0.5 is equivalent to a 'coin flip,' and a value of 1.0 indicates a perfect test.

Our research protocol included a plan to pool likelihood ratios as relative risks (and 95% confidence intervals), and to use sensitivity analyses to evaluate the effect of the trial quality and underlying patient type (critically ill, post-cardiac surgery, post-renal transplant, and so on). Pooling of results was not undertaken, however, due to marked heterogeneity in cutoffs for definitions in both the new index test and the reference 'gold standard' definition of AKI among studies of similar biomarkers.

DISCLOSURE

The authors have no conflicts of interest to disclose.

ACKNOWLEDGMENTS

Dr Parikh is supported by NIH grants (K23-DK064689 and R01-HL085757).

REFERENCES

1. Mehta RL, Chertow GM. Acute renal failure definitions and classification: time for change? *J Am Soc Nephrol* 2003; **14**: 2178–2187.
2. Bjornsson TD. Use of serum creatinine concentrations to determine renal function. *Clin Pharmacokinet* 1979; **4**: 200–222.
3. Star RA. Treatment of acute renal failure. *Kidney Int* 1998; **54**: 1817–1831.
4. Bosch JP. Renal reserve: a functional view of glomerular filtration rate. *Semin Nephrol* 1995; **15**: 381–385.
5. Herrera J, Rodriguez-Iturbe B. Stimulation of tubular secretion of creatinine in health and in conditions associated with reduced nephron mass. Evidence for a tubular functional reserve. *Nephrol Dial Transplant* 1998; **13**: 623–629.
6. American society of nephrology. American society of nephrology renal research report. *J Am Soc Nephrol* 2005; **16**: 1886–1903.
7. Mishra J, Ma Q, Prada A *et al.* Identification of neutrophil gelatinase-associated lipocalin as a novel early urinary biomarker for ischemic renal injury. *J Am Soc Nephrol* 2003; **14**: 2534–2543.
8. Mishra J, Mori K, Ma Q *et al.* Neutrophil gelatinase-associated lipocalin: a novel early urinary biomarker for cisplatin nephrotoxicity. *Am J Nephrol* 2004; **24**: 307–315.
9. Herget-Rosenthal S, Marggraf G, Husing J *et al.* Early detection of acute renal failure by serum cystatin C. *Kidney Int* 2004; **66**: 1115–1122.
10. Parikh CR, Abraham E, Ancukiewicz M, Edelstein CL. Urine IL-18 is an early diagnostic marker for acute kidney injury and predicts mortality in the intensive care unit. *J Am Soc Nephrol* 2005; **16**: 3046–3052.
11. Mishra J, Dent C, Tarabishi R *et al.* Neutrophil gelatinase-associated lipocalin (NGAL) as a biomarker for acute renal injury after cardiac surgery. *Lancet* 2005; **365**: 1231–1238.
12. Abu-Omar Y, Mussa S, Naik MJ *et al.* Evaluation of Cystatin C as a marker of renal injury following on-pump and off-pump coronary surgery. *Eur J Cardiothorac Surg* 2005; **27**: 893–898.
13. Biancofiore G, Pucci L, Cerutti E *et al.* Cystatin C as a marker of renal function immediately after liver transplantation. *Liver Transpl* 2006; **12**: 285–291.
14. Benohr P, Grenz A, Hartmann JT *et al.* Cystatin C—a marker for assessment of the glomerular filtration rate in patients with cisplatin chemotherapy. *Kidney Blood Press Res* 2006; **29**: 32–35.
15. Stabuc B, Vrhovec L, Stabuc-Silih M, Cizej TE. Improved prediction of decreased creatinine clearance by serum cystatin C: use in cancer patients before and during chemotherapy. *Clin Chem* 2000; **46**: 193–197.
16. Zhu J, Yin R, Wu H *et al.* Cystatin C as a reliable marker of renal function following heart valve replacement surgery with cardiopulmonary bypass. *Clin Chim Acta* 2006; **374**: 116–121.
17. Villa P, Jimenez M, Soriano M-C *et al.* Serum cystatin C concentration as a marker of acute renal dysfunction in critically ill patients. [see comment]. *Crit Care (London, England)* 2005; **9**: R139–R143.
18. Wynckel A, Randoux C, Millart H *et al.* Kinetics of carbamylated haemoglobin in acute renal failure. *Nephrol Dial Transplant* 2000; **15**: 1183–1188.

19. Mori K, Lee HT, Rapoport D *et al.* Endocytic delivery of lipocalin–siderophore–iron complex rescues the kidney from ischemia-reperfusion injury. *J Clin Invest* 2005; **115**: 610–621.
20. Parikh CR, Jani A, Melnikov VY *et al.* Urinary interleukin-18 is a marker of human acute tubular necrosis. *Am J Kidney Dis* 2004; **43**: 405–414.
21. Boldt J, Brenner T, Lang J *et al.* Kidney-specific proteins in elderly patients undergoing cardiac surgery with cardiopulmonary bypass. *Anesth Analg* 2003; **97**: 1582–1589.
22. Han WK, Bailly V, Abichandani R *et al.* Kidney injury molecule-1 (KIM-1): a novel biomarker for human renal proximal tubule injury. *Kidney Int* 2002; **62**: 237–244.
23. du Cheyron D, Daubin C, Poggioli J *et al.* Urinary measurement of Na+/H+ exchanger isoform 3 (NHE3) protein as new marker of tubule injury in critically ill patients with ARF. *Am J Kidney Dis* 2003; **42**: 497–506.
24. Ahlstrom A, Tallgren M, Peltonen S, Pettila V. Evolution and predictive power of serum cystatin C in acute renal failure. *Clin Nephrol* 2004; **62**: 344–350.
25. Mazul-Sunko B, Zarkovic N, Vrkic N *et al.* Proatrial natriuretic peptide (1–98), but not cystatin C, is predictive for occurrence of acute renal insufficiency in critically ill septic patients. *Nephron Clin Pract* 2004; **97**: c103–c107.
26. Bachorzewska-Gajewska H, Malyszko J, Sitniewska E *et al.* Neutrophil-gelatinase-associated lipocalin and renal function after percutaneous coronary interventions. *Am J Nephrol* 2006; **26**: 287–292.
27. Rinder CS, Fontes M, Mathew JP *et al.* Neutrophil CD11b upregulation during cardiopulmonary bypass is associated with postoperative renal injury. *Ann Thorac Surg* 2003; **75**: 899–905.
28. Wagener G, Jan M, Kim M *et al.* Association between increases in urinary neutrophil gelatinase-associated lipocalin and acute renal dysfunction after adult cardiac surgery. *Anesthesiology* 2006; **105**: 485–491.
29. Parikh CR, Jani A, Mishra J *et al.* Urine NGAL and IL-18 are predictive biomarkers for delayed graft function following kidney transplantation. *Am J Transplant* 2006; **6**: 1639–1645.
30. Parikh CR, Mishra J, Thiessen-Philbrook H *et al.* Urinary IL-18 is an early predictive biomarker of acute kidney injury after cardiac surgery. *Kidney Int* 2006; **70**: 199–203.
31. Eijkenboom JJ, van Eijk LT, Pickkers P *et al.* Small increases in the urinary excretion of glutathione *S*-transferase A1 and P1 after cardiac surgery are not associated with clinically relevant renal injury. *Intensive Care Med* 2005; **31**: 664–667.
32. Westhuyzen J, Endre ZH, Reece G *et al.* Measurement of tubular enzymuria facilitates early detection of acute renal impairment in the intensive care unit. *Nephrol Dial Transplant* 2003; **18**: 543–551.
33. Trachtman H, Christen E, Canaan A *et al.* Urinary NGAL in D+HUS: a novel marker of renal injury. *Pediatr Nephrol* 2006; **21**: 989–994.
34. Simmons EM, Himmelfarb J, Sezer MT *et al.* Plasma cytokine levels predict mortality in patients with acute renal failure. *Kidney Int* 2004; **65**: 1357–1365.
35. Herget-Rosenthal S, Poppen D, Husing J *et al.* Prognostic value of tubular proteinuria and enzymuria in nonoliguric acute tubular necrosis. *Clin Chem* 2004; **50**: 552–558.
36. Liangos O, Perianayagam MC, Vaidya VS *et al.* Urinary *N*-Acetyl-beta-(D)-Glucosaminidase activity and kidney injury molecule-1 level are associated with adverse outcomes in acute renal failure. *J Am Soc Nephrol* 2007; **18**: 904–912.
37. Han WK, Waikar SS, Johnson AM *et al.* Urinary biomarkers for detection of acute kidney injury. *Kidney Int* 2007 (in press).
38. Zappitelli M, Washburn KK, Arikan AA *et al.* Urine neutrophil gelatinase-associated lipocalin is an early marker of acute kidney injury in critically ill children: a prospective cohort study. *Crit Care* 2007; **11**: R84.
39. Washburn KK, Zappitelli M, Arikan AA *et al.* Urinary interleukin-18 is an acute kidney injury biomarker in critically Ill children. *Nephrol Dial Transplant* 2007 [e-pub ahead of print].
40. Han WK, Bailly V, Abichandani R *et al.* Kidney injury molecule-1 (KIM-1): a novel biomarker for human renal proximal tubule injury. *Kidney Int* 2002; **62**: 237–244.
41. Uchino S, Kellum JA, Bellomo R *et al.* Acute renal failure in critically ill patients: a multinational, multicenter study. *JAMA* 2005; **294**: 813–818.
42. Hoste EA, Lameire NH, Vanholder RC *et al.* Acute renal failure in patients with sepsis in a surgical ICU: predictive factors, incidence, comorbidity, and outcome. *J Am Soc Nephrol* 2003; **14**: 1022–1030.
43. Yegenaga I, Hoste E, Van Biesen W *et al.* Clinical characteristics of patients developing ARF due to sepsis/systemic inflammatory response syndrome: results of a prospective study. *Am J Kidney Dis* 2004; **43**: 817–824.
44. Pepe MS, Janes H, Longton G *et al.* Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol* 2004; **159**: 882–890.
45. Bruns DE, Huth EJ, Magid E, Young DS. Toward a checklist for reporting of studies of diagnostic accuracy of medical tests. *Clin Chem* 2000; **46**: 893–895.