Appendix for:

Biomarkers for Prediction of Renal Replacement Therapy in Acute Kidney Injury: A Systematic Review and Meta-Analysis

1. Search Strategies

Embase

Embase search strategy:

1 biomarker (139901)
2 NGAL OR neutrophil gelatinase-associated lipocalin OR neutrophil gelatinase associated lipocalin (6910)
3 KIM-1 OR kidney injury molecule-1 (1462)
4 cystatin c OR cystatin-c (9633)
5 L-FABP OR fatty acid-binding protein 1 (1049)
6 IL-18 OR interleukin-18 or interleukin 18 (17305)
7 IGFBP7 OR IGF-binding protein-7 OR IGF binding protein 7 (433)
8 TIMP2 OR tissue inhibitor metalloproteinase-2 (1029)
9 calprotectin (3997)
10 CAF OR c-terminal agrin fragment OR c terminal agrin fragment (4427)
11 L1 OR L2 OR L3 OR L4 OR L5 OR L6 OR L7 OR L8 OR L9 OR L10
12 AKI OR acute kidney injury OR acute kidney failure OR acute renal failure (80015)
13 RRT or renal replacement therapy or ?dialysis OR CVVH? OR (hemofiltration OR haemofiltration) OR (CRRT OR continuous renal replacement therapy) (279067)
14 L11 AND L12 AND L13 (947)
15 L14 NOT MEDLINE/FS (913)
16 L14 NOT REVIEW/DT (752)

752 Search results
Results obtained September 19, 2017

Pubmed

Pubmed/NIH search strategy:

(biomarker
OR (NGAL OR „Neutrophil gelatinase-associated lipocalin“ OR „Neutrophil gelatinase associated lipocalin“)
OR (KIM-1 OR „Kidney Injury Molecule-1“)
OR („Cystatin C“ OR „Cystatin-C“)
OR (L-FABP OR „Fatty acid-binding protein 1„)
OR (IL-18 OR „Interleukin-18“ OR „Interleukin 18“)
OR (IGFBP7 OR „IGF-Binding Protein-7“ OR „IGF Binding Protein 7“)
OR (TIMP2 OR „Tissue Inhibitor Metalloproteinase-2“)
OR Calprotectin
OR (CAF OR „c-terminal agrin fragment“ OR „c terminal agrin fragment“))
AND (aki OR „acute kidney injury“ OR „acute kidney failure“ OR „acute renal failure“)
AND (rrt OR „renal replacement therapy“ OR *dialysis OR CVVH* OR (hemofiltration OR haemofiltration) OR (CRRT OR „continous renal replacement therapy“))
NOT Review[ptyp]

656 Search results
Results obtained September 19, 2017

CENTRAL

CENTRAL search strategy:

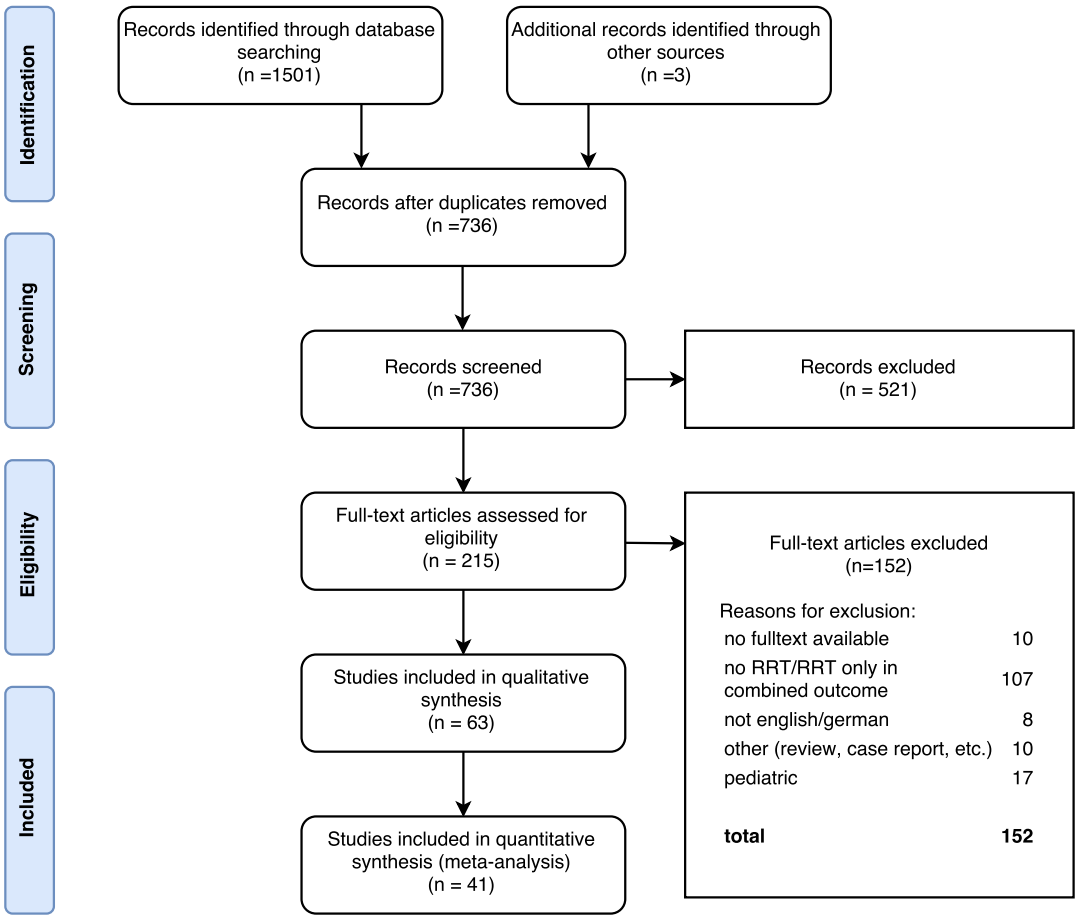
ID	Search	Hits
#1	biomarker	5980
#2	(NGAL or "Neutrophil gelatinase-associated lipocalin" or "Neutrophil gelatinase associated lipocalin")	370
#3	(KIM-1 or "Kidney Injury Molecule-1")	113
#4	("Cystatin C" or "Cystatin-C")	516
#5	(L-FABP or "Fatty acid-binding protein 1")	54
#6	(IL-18 or "Interleukin-18" or "Interleukin 18")	392
#7	(IGFBP7 or "IGF-Binding Protein-7" or "IGF Binding Protein 7")	17
#8	(TIMP2 or "Tissue Inhibitor Metalloproteinase-2" or TIMP-2)	70
#9	Calprotectin	325
#10	(CAF or "c-terminal agrin fragment" or "c terminal agrin fragment")	541
#11	(aki or "acute kidney injury" or "acute kidney failure" or "acute renal failure")	3179
#12	(rrt or "renal replacement therapy" or *dialysis or CVVH* or (hemofiltration or haemofiltration) or (CRRT or "continous renal replacement therapy"))	15562
#13	#2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10	2163
#14	#1 or #13	7962
#15	#14 and #11 and #12	115

115 Search results (93 Trials, 1 Econ. Eval., 21 Reviews excluded)
Results obtained September 19, 2017

2. Supplemental figures and tables

2.1. Supplemental figures and tables for the study selection and QUADAS-2 risk of bias assessment

Fig. 7 Flow chart of study selection



QUADAS-2 risk of bias assessment (Fig.8 and Tab.2)

QUADAS-2 risk of bias assessment is performed in four domains:

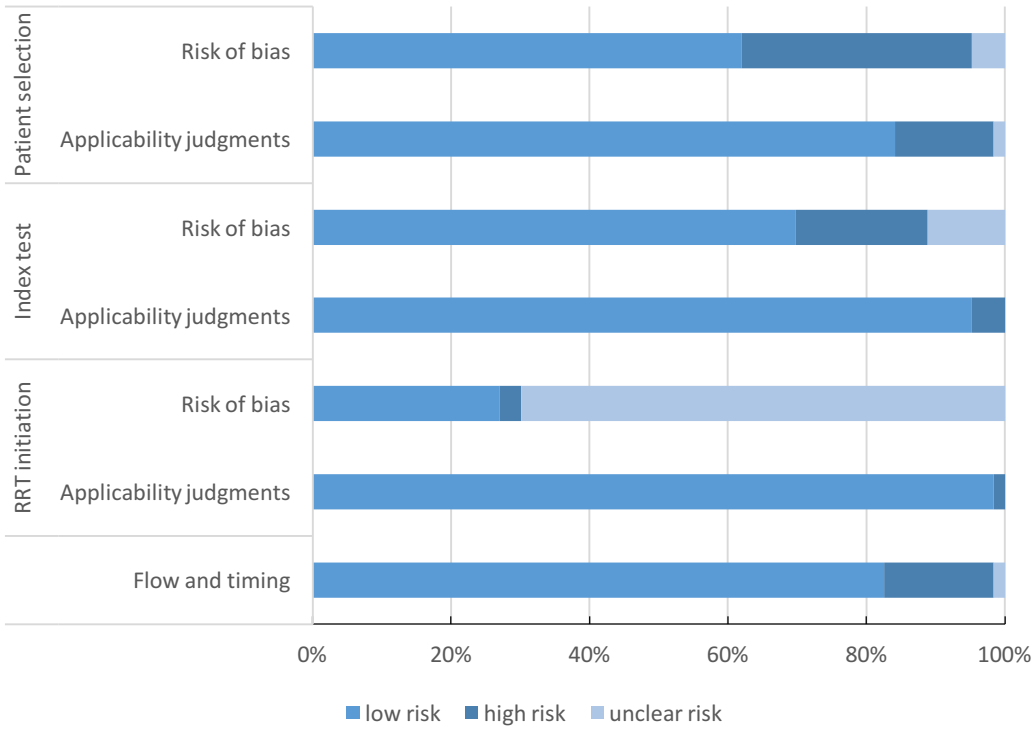
Risk of bias in the domain ‘patient selection’ evaluated the methods of patient selection (e.g. if a consecutive or random sample of patients was enrolled or if the study avoided inappropriate exclusions); the applicability judgment was based on whether there was concern that the included patients matched the review question or not.

Risk of bias in the domain ‘index test’ states, whether there was concern that the conduct and interpretation (e.g. if the laboratory personal was blinded to the patient’s condition) of the index test lead to possible bias, while the applicability judgment was based on whether there was concern that the index test, its conduct or the interpretation differed from the review question or not.

The domain ‘RRT initiation’ states whether there was a possible risk of bias regarding the initiation of RRT (eg. what criteria were used to classify the need for RRT initiation and if the decision to initiate RRT was made without knowledge of the biomarker results). The applicability judgment in this domain was based on whether there was concern that the initiation of RRT matched the review question.

In the domain ‘flow and timing’ judgment for risk of bias was based on whether there was concern if e.g. some patients were excluded from the analysis or if there was an appropriate interval between critical illness and RRT.

Fig. 8 Results of the QUADAS-2 risk of bias assessment.



Tab. 2 Results of the QUADAS-2 risk of bias assessment

	Patient selection		Index Test		RRT Initiation		
Author/Year	Risk of bias	Applicability judgements	Risk of bias	Applicability judgements	Risk of bias	Applicability judgements	Flow & Timing
Albeladi et al. 2017 [28]	low	low	low	low	low	low	Low
Alge et al. 2013 [59]	low	low	high	high	high	low	Low
Bagshaw et al. 2010 [29]	low	low	high	low	high	low	low
Cemil et al. 2014 [19]	high	low	high	low	low	low	high
Chun et al. 2017 [80]	low	low	unclear	low	unclear	low	low
Constantin et al. 2010 [30]	low	low	low	low	unclear	low	low
Cruz et al. 2009 [31]	low	low	low	low	unclear	low	low
de Geus et al. 2011 [32]	low	low	low	low	unclear	low	low
Dihazi et al. 2016 [33]	low	low	low	low	low	low	low
Drey et al. 2015 [75]	low	low	high	low	unclear	low	low
Du et al. 2013 [60]	high	low	high	low	unclear	low	low
Dusse et al. 2016 [64]	low	low	low	low	unclear	low	low
Endre et al. 2010 [35]	low	low	low	high	unclear	low	low
Endre et al. 2011 [34]	low	low	low	high	unclear	low	low
Gaipov et al. 2015 [67]	low	low	high	low	unclear	low	low
Garcia-Alvarez et al. 2015 [61]	low	low	high	low	unclear	low	high
Glassford et al. 2013 [36]	low	low	low	low	unclear	low	low
Gocze et al. 2015 [68]	low	low	low	low	unclear	low	low
Haase-Fielitz et al. 2009 [65]	high	low	low	low	low	low	high
Haase-Fielitz et al. 2011 [62]	high	low	low	low	low	low	low
Haines et al. 2017 [37]	low	low	low	low	unclear	low	low
Hanson et al. 2011 [77]	high	high	high	low	low	low	low
Herget-Rosenthal et al. 2004 [38]	low	low	low	low	unclear	low	low
Herget-Rosenthal et al. 2004 [20]	low	low	low	low	low	low	low
Hjortrup et al. 2015 [74]	high	unclear	low	low	unclear	low	low
Ho et al. 2017 [21]	low	low	unclear	low	unclear	low	low
Hu et al. 2017 [70]	low	low	low	low	unclear	low	low
Itenov et al. 2016 [39]	high	low	low	low	unclear	low	low
Jalkanen et al. 2013[40]	low	low	high	low	unclear	low	low
Kiessling et al. 2014[58]	high	high	low	low	low	low	low
Kim et al. 2017 [76]	low	low	low	low	unclear	low	low
Koyner et al. 2015 [41]	high	low	low	low	unclear	low	low
Koziolek et al. 2012 [42]	low	low	high	low	low	low	low
Linko et al. 2013 [43]	high	high	low	low	unclear	low	high
Lukasz et al. 2014 [79]	low	high	low	low	unclear	low	low
Mahdavi-Mazdeh et al. 2012 [73]	high	high	low	low	low	high	low
Maisel et al. 2016 [22]	low	low	low	low	unclear	low	low
Mårtensson et al. 2017 [44]	low	low	low	low	unclear	low	low
McIlroy et al. 2015 [66]	high	low	low	low	unclear	low	low
Nejat et al. 2010 [45]	high	high	low	low	unclear	low	high
Nisula et al. 2014 [46]	low	low	low	low	unclear	low	low
Nisula et al. 2015 [47]	low	low	low	low	unclear	low	low
O'Sullivan et al. 2017 [48]	high	low	low	low	unclear	low	low
Park et al. 2013 [23]	low	low	low	low	low	low	low
Pianta et al. 2015 [72]	high	high	unclear	low	low	low	low
Pickering et al. 2012 [50]	low	low	low	low	unclear	low	unclear
Pickering et al. 2013 [49]	high	low	low	low	unclear	low	high
Pipili et al. 2014 [51]	high	low	unclear	low	low	low	low
Plewes et al. 2017 [78]	low	high	low	low	low	low	low
Ralib et al. 2012 [52]	low	low	low	low	unclear	low	high
Renhua et al. 2014 [24]	high	low	low	low	unclear	low	low
Rewa et al. 2015 [25]	low	low	high	low	unclear	low	high
Royakkers et al. 2011 [54]	high	low	low	low	unclear	low	low
Royakkers et al. 2012 [53]	high	low	low	low	low	low	low
Shum et al. 2015 [69]	low	low	high	low	low	low	low
Siew et al. 2010 [55]	low	low	low	low	unclear	low	high
Siew et al. 2013 [56]	low	low	low	low	unclear	low	high

	Patient selection		Index Test		RRT Initiation		
Author/Year	Risk of bias	Applicability judgements	Risk of bias	Applicability judgements	Risk of bias	Applicability judgements	Flow & Timing
Skinner et al. 2017 [81]	low	low	low	low	unclear	low	low
Srisawat et al. 2011 [26]	high	low	unclear	low	unclear	low	low
Sumida et al. 2014 [63]	unclear	low	unclear	low	unclear	low	low
Susantitaphong et al. 2012 [27]	low	high	low	low	unclear	low	low
Tiranathanagul et al. 2013 [57]	unclear	low	unclear	low	low	low	low
Valette et al. 2013 [71]	unclear	low	low	low	unclear	low	low

2.2. Supplemental forest plots

Fig. 9 Forest plots of urinary IL-18 predicting RRT.
Fig. 9a Urinary concentration of IL-18.

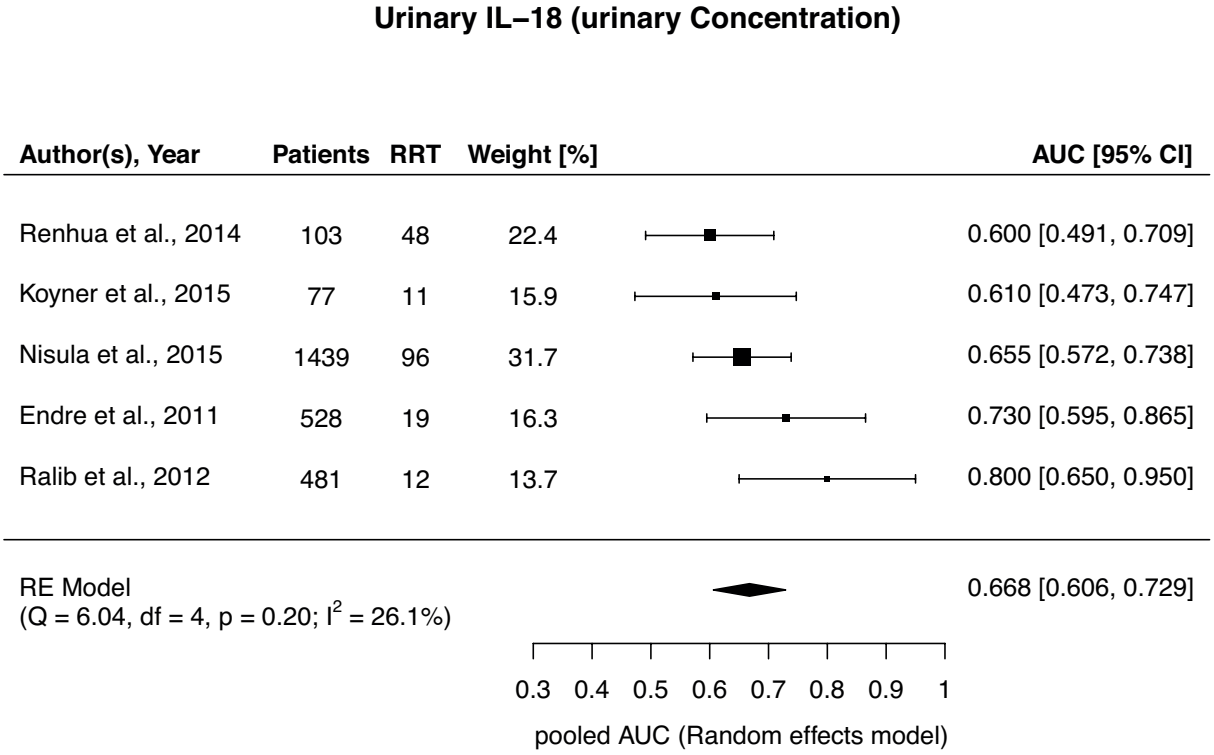


Fig. 9b Urinary IL-18 normalized to urinary creatinine.

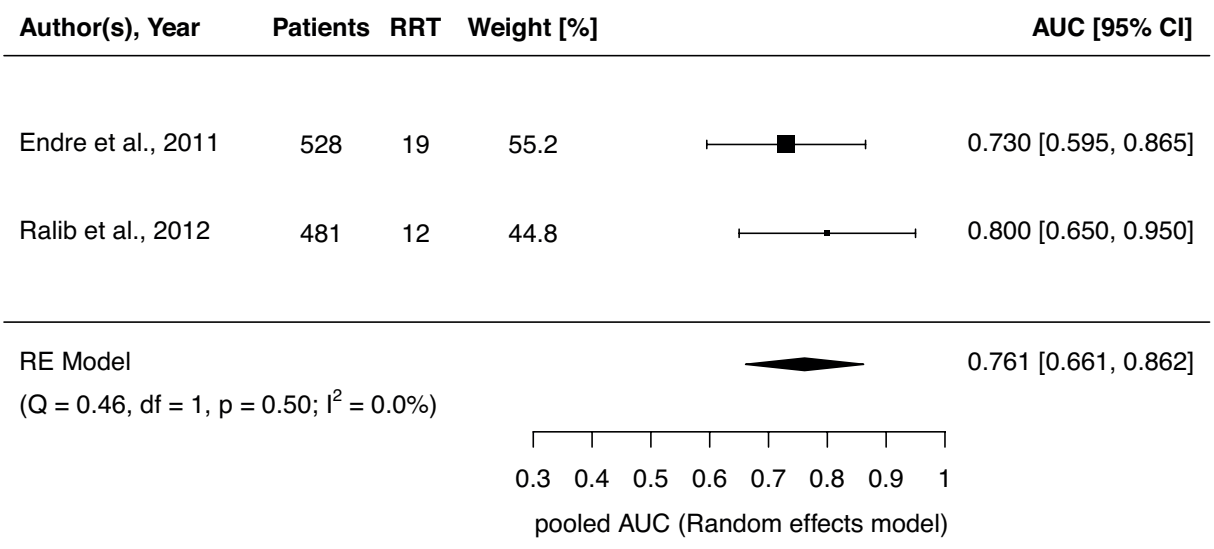


Fig. 10 Forest plots of urinary cystatin C (conc.) predicting RRT

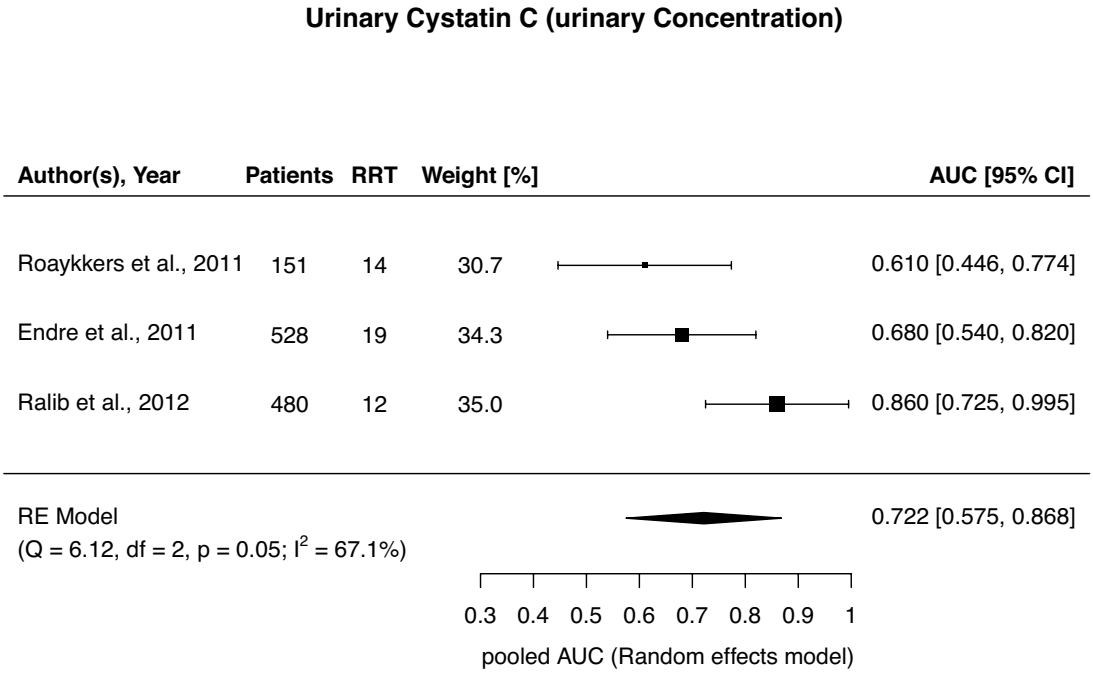


Fig. 11 Forest plots of urinary KIM-1 predicting RRT.
Fig.11a Urinary concentration of KIM-1.

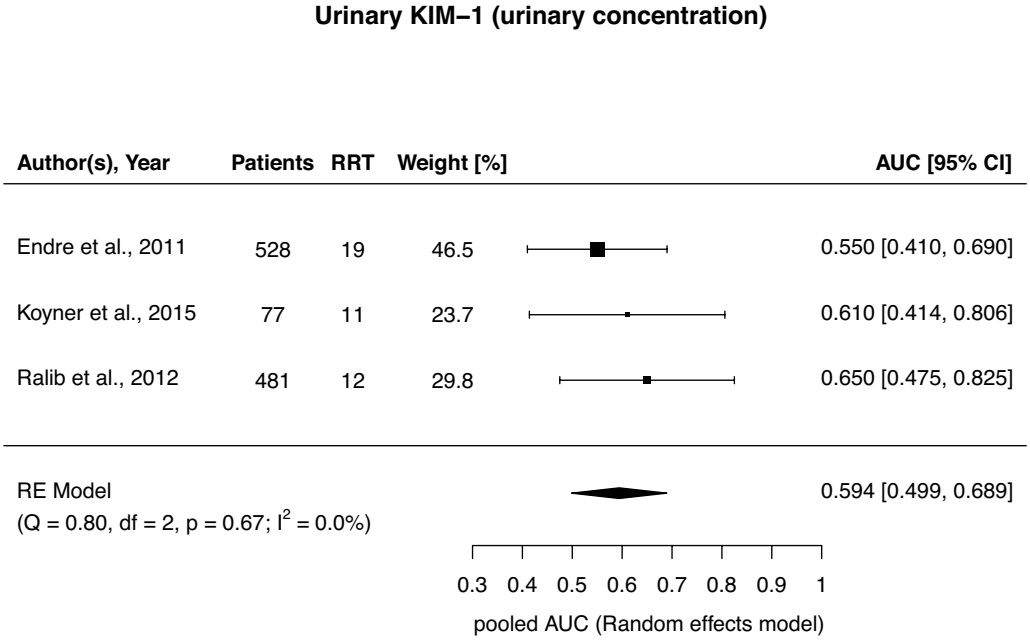


Fig. 11b Urinary KIM-1 normalized to urinary creatinine

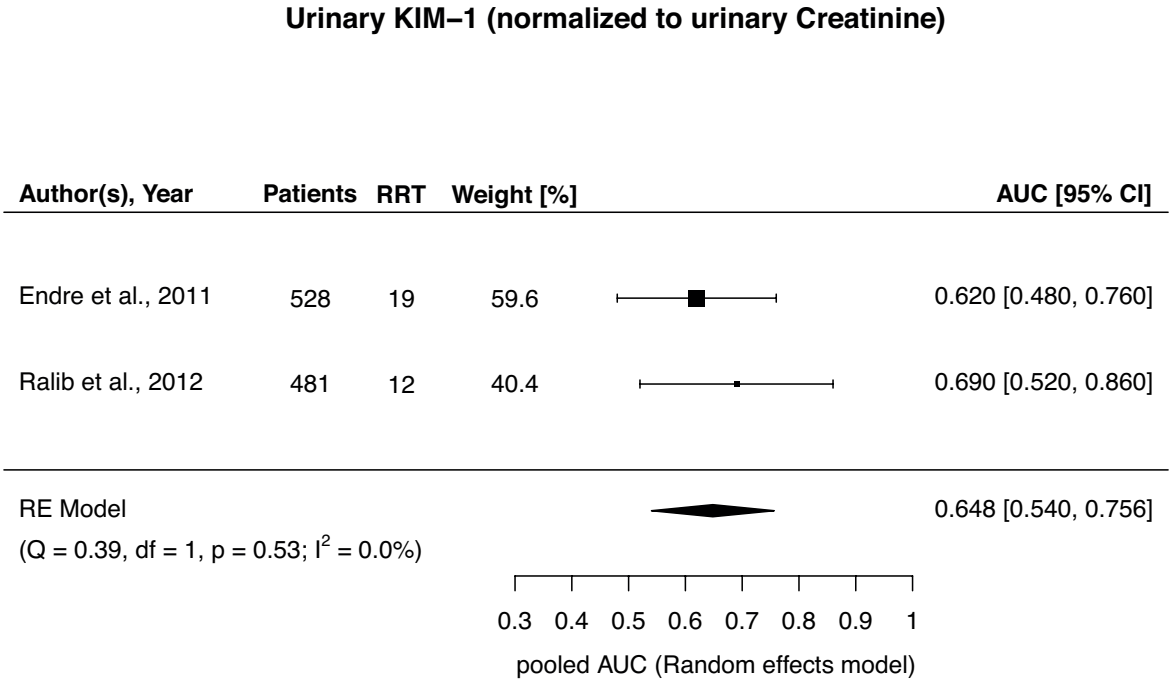
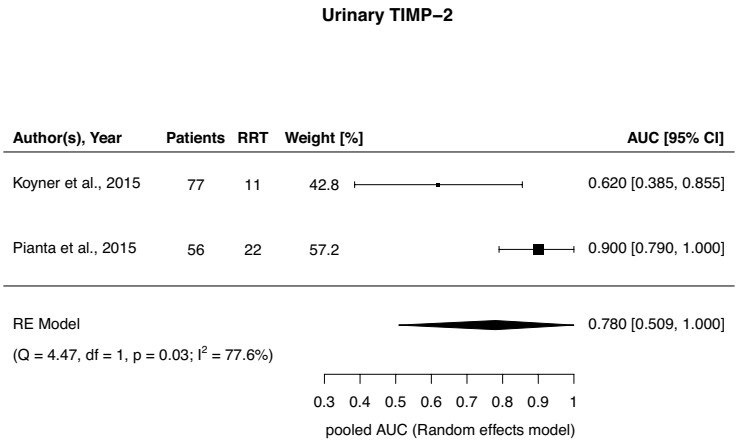


Fig. 12a Forest plot of urinary TIMP-2 predicting RRT. Fig. 12b Forest plot of urinary IGFBP-7 predicting RRT



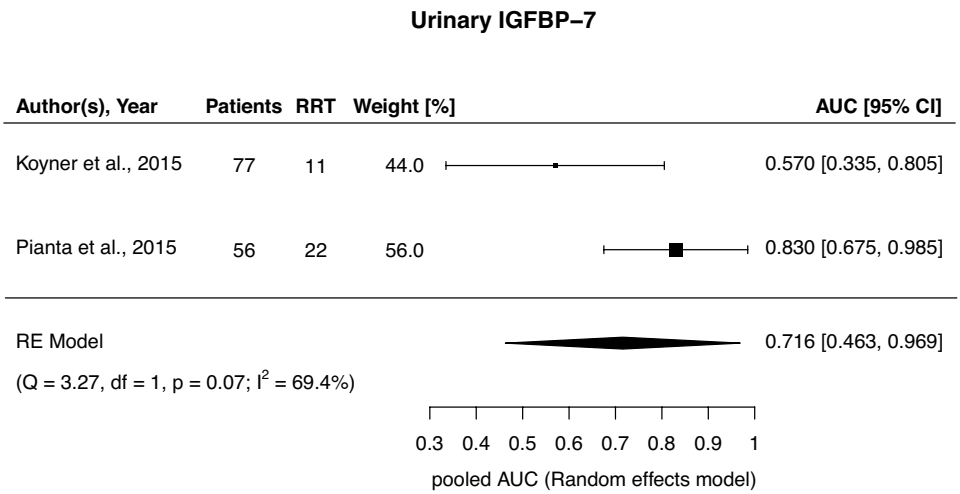


Fig. 13a Forest plot of urine output predicting RRT

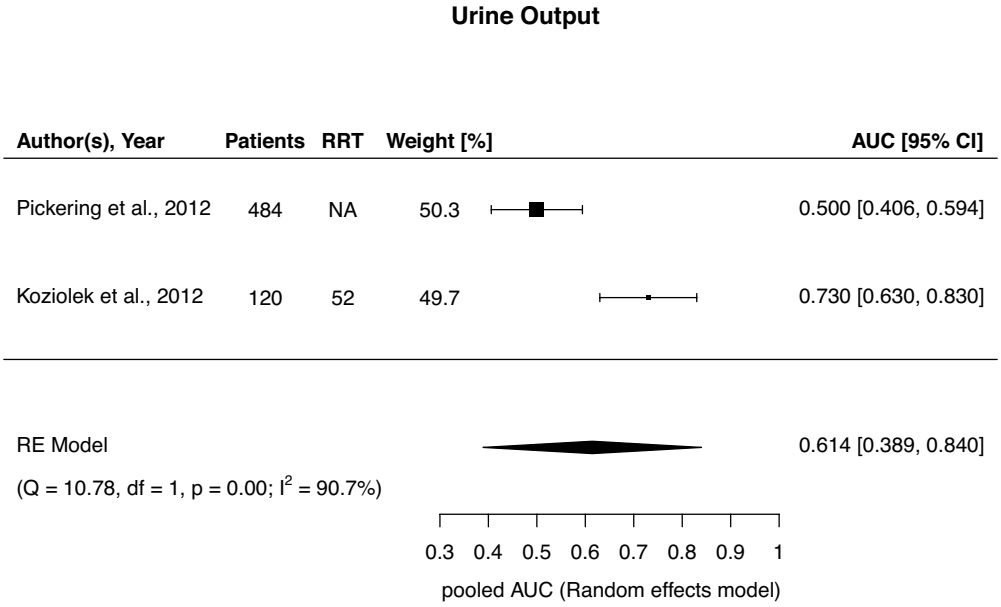


Fig. 13b Forest plot of urinary NAG normalized to urinary creatinine predicting RRT

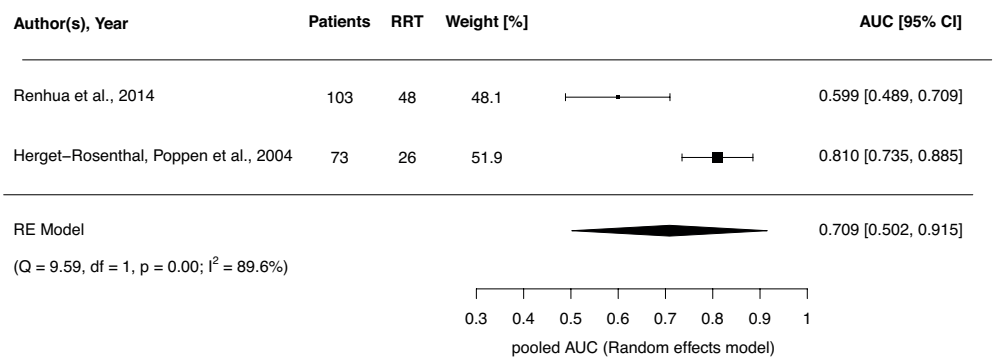


Fig. 13c Forest plot of the Fractional Excretion of Sodium predicting RRT

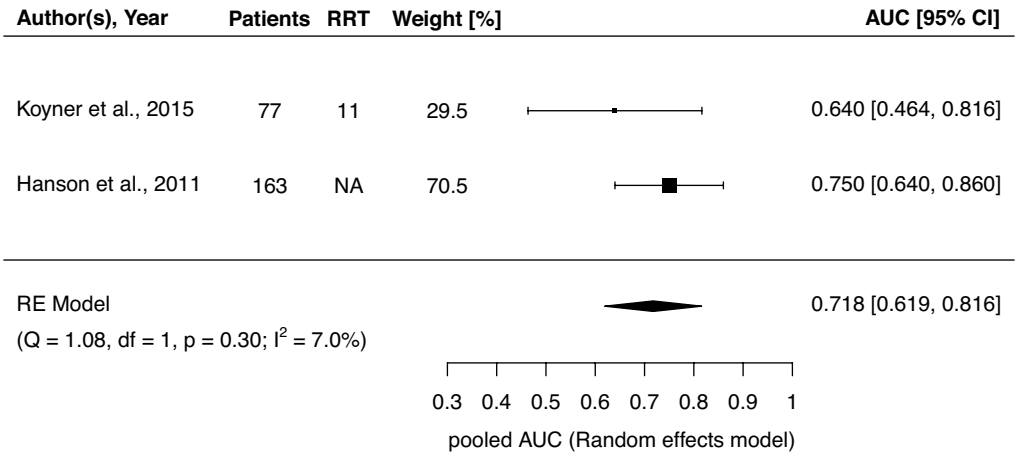


Fig. 13d *Forest plot of blood urea nitrogen predicting RRT*

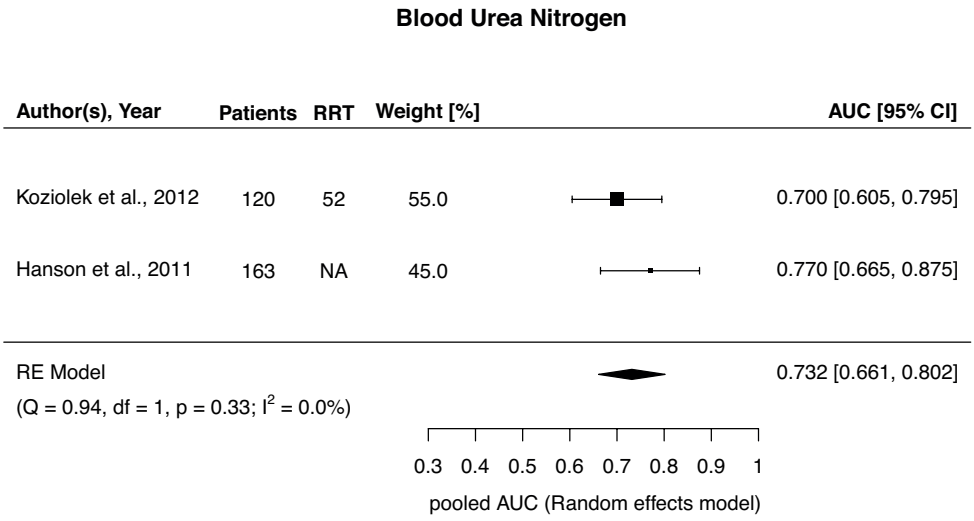


Fig. 14 Funnel plots of biomarkers evaluated in at least 10 studies

Fig. 14a Funnel plot for the urinary concentration of urinary NGAL

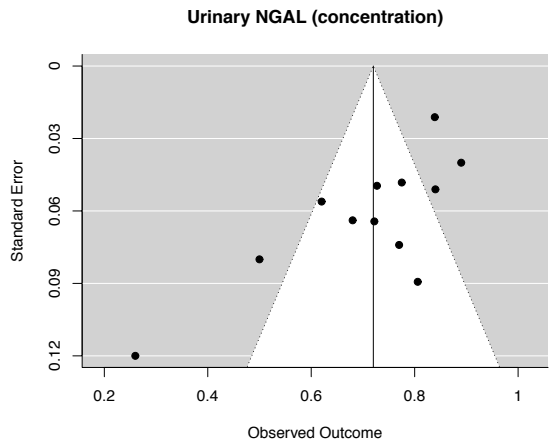


Fig. 14b Funnel plot for plasma, serum and whole blood NGAL

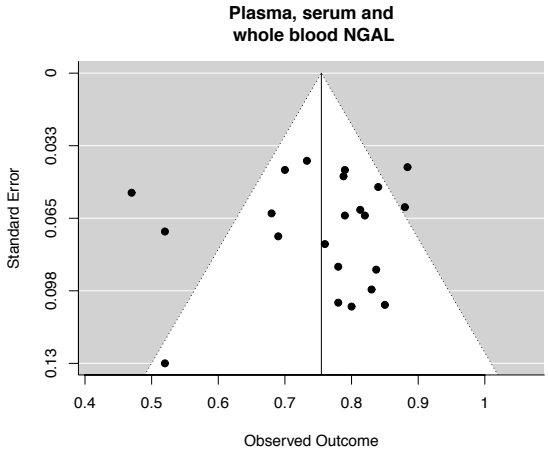
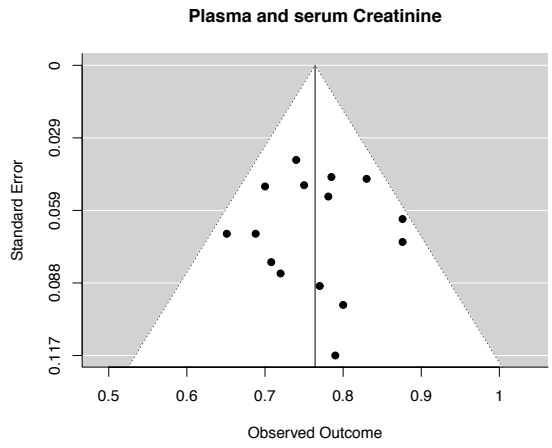


Fig. 14c Funnel plot for plasma and serum creatinine



Tab. 3 Differences between pooled AUCs for urinary biomarkers

	IL-18 urinary conc. vs. normalized to urinary creatinine	**Cystatin C** urinary conc. vs. normalized to urinary creatinine	**KIM-1** urinary conc. vs. normalized to urinary creatinine
Difference	0.007	0.068	0.054
p-value	0.8863	0.5172	0.4618

Tab. 4 Differences between pooled AUCs for NGAL

	NGAL plasma vs. serum	**blood NGAL** vs. urinary NGAL (concentration)	**blood NGAL** vs. urinary NGAL (normalized)
Difference	0.114	0.035	0.028
p-value	0.1593	0.4735	0.4260

Tab. 5 Differences between pooled AUCs for blood creatinine and cystatin C

	Creatinine plasma vs. serum	**Cystatin C** plasma vs. serum	p+s **Cystatin C** vs. urinary CysC (concentration)	p+s **Cystatin C** vs. urinary CysC (normalized)
Difference	0.013	0.077	0.046	0.022
p-value	0.7178	0.3310	0.5520	0.7733

3. Sensitivity analyses

3.1. Inclusion of moderators into the Random Effects Model

To address the between-trial heterogeneity, moderators (number of patients receiving RRT) were added to the random-effects model on an exploratory basis, but no significant findings were noticed during this process.

For *plasma/serum/whole blood NGAL*, the test for residual heterogeneity and for moderators shows no significant influence of the incidence of RRT. The residual heterogeneity decreases from Q=69.18 (df=21) to Q=69.1 (df=20). The effect of the incidence of RRT is not statistical significant: Q=0.02 (df=1).

For *plasma/serum creatinine*, the test for residual heterogeneity and for moderators shows no significant influence of the incidence of RRT. The residual heterogeneity decreases from Q=15.23 (df=14) to Q=8.83 (df=9). The effect of the incidence of RRT is not statistical significant: Q=0.07 (df=1).

For *plasma/serum cystatin C*, the test for residual heterogeneity and for moderators shows no significant influence of the incidence of RRT. The residual heterogeneity decreases from Q=4.9 (df=6) to Q=2.75 (df=4). The effect of the incidence of RRT is not statistical significant: Q=0.01 (df=1).

For *urinary TIMP-2*IGFBP-7*, the test for residual heterogeneity and for moderators shows no significant influence of the incidence of RRT. The residual heterogeneity decreases from Q= 5.7 (df=3) to Q=5.66 (df=2). The effect of the incidence of RRT is not statistical significant: Q=0.00 (df=1).

For *urinary cystatin C (conc.)*, the test for residual heterogeneity and for moderators shows no significant influence of the incidence of RRT. The residual heterogeneity decreases from Q=6.12 (df=2) to Q=4.02 (df=1). The effect of the incidence of RRT is not statistical significant: Q=0.38 (df=1).

For *urinary cystatin C (norm.)*, the test for residual heterogeneity and for moderators shows no significant influence of the incidence of RRT. The residual heterogeneity decreases from Q=19.70 (df=3) to Q=11.62 (df=2). The effect of the incidence of RRT is not statistical significant: Q=0.48 (df=1).

For *urinary NGAL (conc.)*, the test for residual heterogeneity and for moderators shows no significant influence of the incidence of RRT. The residual heterogeneity decreases from Q=60.25 (df=11) to Q=45.58 (df=9). The effect of the incidence of RRT is not statistical significant: Q=1.45 (df=1).
For *urinary NGAL (norm.)*, the test for residual heterogeneity and for moderators shows no significant influence of the incidence of RRT. The residual heterogeneity decreases from Q=12.94 (df=6) to Q=12.84 (df=5). The effect of the incidence of RRT is not statistical significant: Q=0.14 (df=1).
For *urinary IL-18 (conc.)*, the test for residual heterogeneity and for moderators shows no significant influence of the incidence of RRT. The residual heterogeneity decreases from Q=6.04 (df=4) to Q=5.44 (df=3). The effect of the incidence of RRT is not statistical significant: Q=0.41 (df=1).
For *urinary KIM-1 (conc.)*, the test for residual heterogeneity and for moderators shows no significant influence of the incidence of RRT. The residual heterogeneity decreases from Q=0.80 (df=2) to Q=0.14 (df=1). The effect of the incidence of RRT is not statistical significant: Q=0.65 (df=1).

It was not possible to add moderators to the RE models of BUN, FeNa, TIMP-2, IL-18 (norm.), IGFBP-7, NAG, KIM-1 (norm.) and UO due to the limited number of studies included in these RE models.

3.2. Paired analysis of biomarkers reported in the same study

Another process applied to address between-trial heterogeneity was to pair biomarkers reported in the same studies and create an estimate of the average AUC improvement for those biomarkers.
When comparing plasma, serum and whole blood NGAL and plasma/serum creatinine in studies that provide results for both biomarkers, the performance of NGAL is slightly better than creatinine (AUC improvement Hjortrup -0.04, Valette +0.04, Pickering +0.07, Tiranathanagul +0.10, Sumida +0.06, Maisel +0.01, Gaipov -0.13, Mahdavi-Mazdeh +0.00, average AUC improvement 0.013) as opposed to the pooled AUC from the unpaired analysis.
The trend for plasma/serum cystatin C of outperforming plasma/serum creatinine still holds true after pairing those biomarkers (Nejat +0.07, Koziolek +0.04, Pipili -0.02, Renhua -0.013, Kiessling -0.03, average AUC improvement 0.01).
Urinary TIMP-2 which is outperformed by TIMP-2*IGFBP-7 in the pooled analysis, shows improved performance, indicating a slightly better predictive performance than TIMP-2 and IGFBP-7 combined, when only comparing the reported values by Koyner and Pianta (AUC improvement +0.01 and +0.02, respectively, average AUC improvement 0.015).
Urinary cystatin C (conc.) outperforms urinary IL-18 (conc.) in the pooled analysis, this gap diminishes when pairing those biomarkers (AUC improvement for Endre and Ralib -0.05 and +0.06, respectively, average AUC improvement 0.005). When urinary cystatin C and IL-18 were normalized to urinary creatinine, cystatin C still performs slightly better than IL-18 (AUC improvement -0.02, +0.08, respectively, average AUC improvement 0.03).
For urinary NGAL (norm.), which is outperformed by urinary cystatin C (norm.), a slight improvement can be noted when pairing it with cystatin C (norm.) (AUC improvement Endre +0.08, Ralib -0.04, average AUC improvement 0.02; note: only 2 of 7 studies for urinary NGAL considered for paired analysis, omitting 5 studies).
While urinary NGAL (norm.) is outperformed by urinary IL-18 (norm.) the opposite holds true for NGAL (conc.) and IL-18 (conc.). When pairing NGAL (conc.) and IL-18 (conc.), NGAL slightly outperforms IL-18 (AUC improvement Renhua +0.17, Koyner -0.11, Endre +0.04, average AUC improvement 0.035), while IL-18 (norm.) is slightly better than NGAL (norm.) (AUC improvement Endre +0.06, Ralib +0.04, average AUC improvement 0.05).

3.3. Subgroup analysis for biomarkers included in the meta-analysis

Tab. 6 Subgroup analysis for plasma, serum and whole blood NGAL. Mixed ICU populations are ICU cohorts which include medical and surgical patients. Excluded from this category are studies only investigating specific patient cohorts for example after cardiac surgery, after renal transplantation, suffering from malaria etc. Difference means difference between pooled AUCs

	p/s/wb NGAL (all studies)	mixed ICU populations only	cut-off 150-350 ng/ml	cut-off >600 ng/ml	cut-off 150-350 vs. >600 ng/ml
AUC	0.755	0.747	0.742	0.779	
(95%-CI)	(0.706-0.803)	(0.685-0.808)	(0.678-0.805)	(0.689-0.870)	
Difference		0.008	0.013	0.024	0.037
p-value		0.8413	0.7964	0.6469	0.5613

Fig. 15 Plasma, serum and whole blood NGAL. Subgroup analysis including only mixed ICU populations

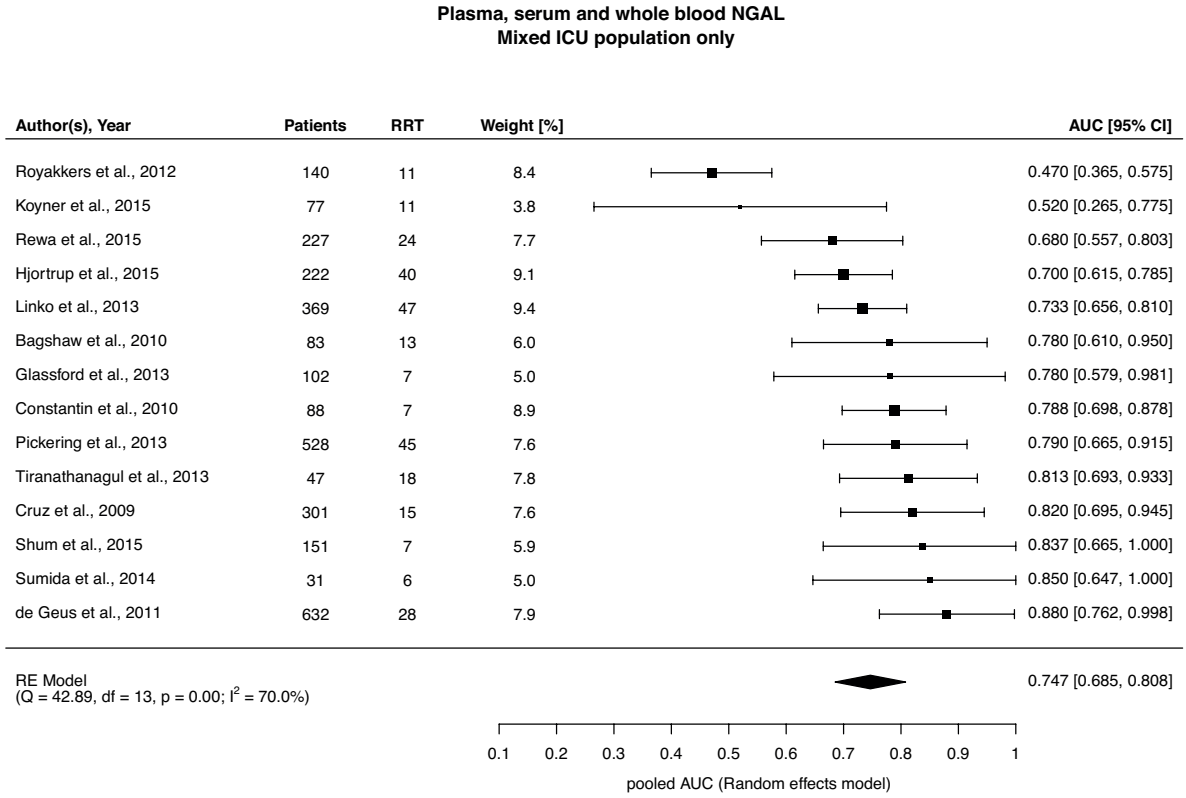


Fig. 16 Subgroup analysis for different NGAL thresholds

Fig. 16a Plasma, serum and whole blood NGAL. Subgroup analysis including only studies with stated cut-off between 150 and 350 ng/ml

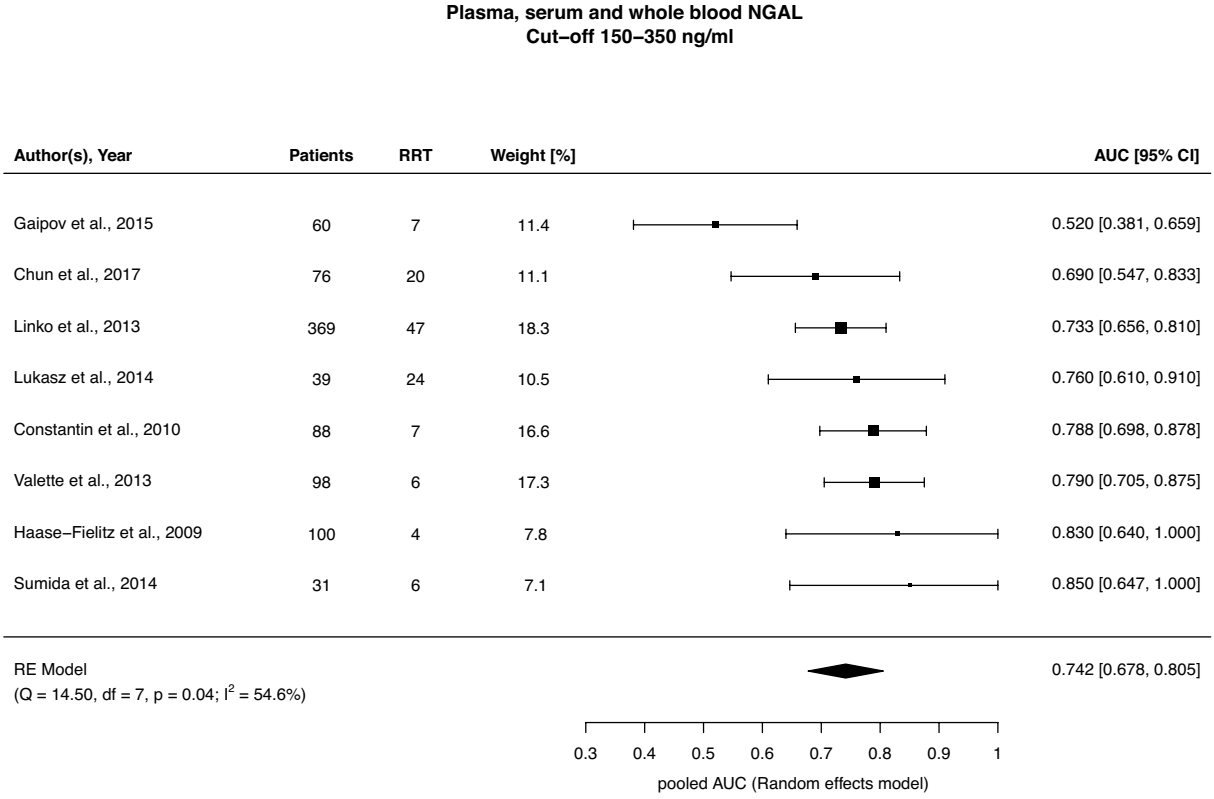
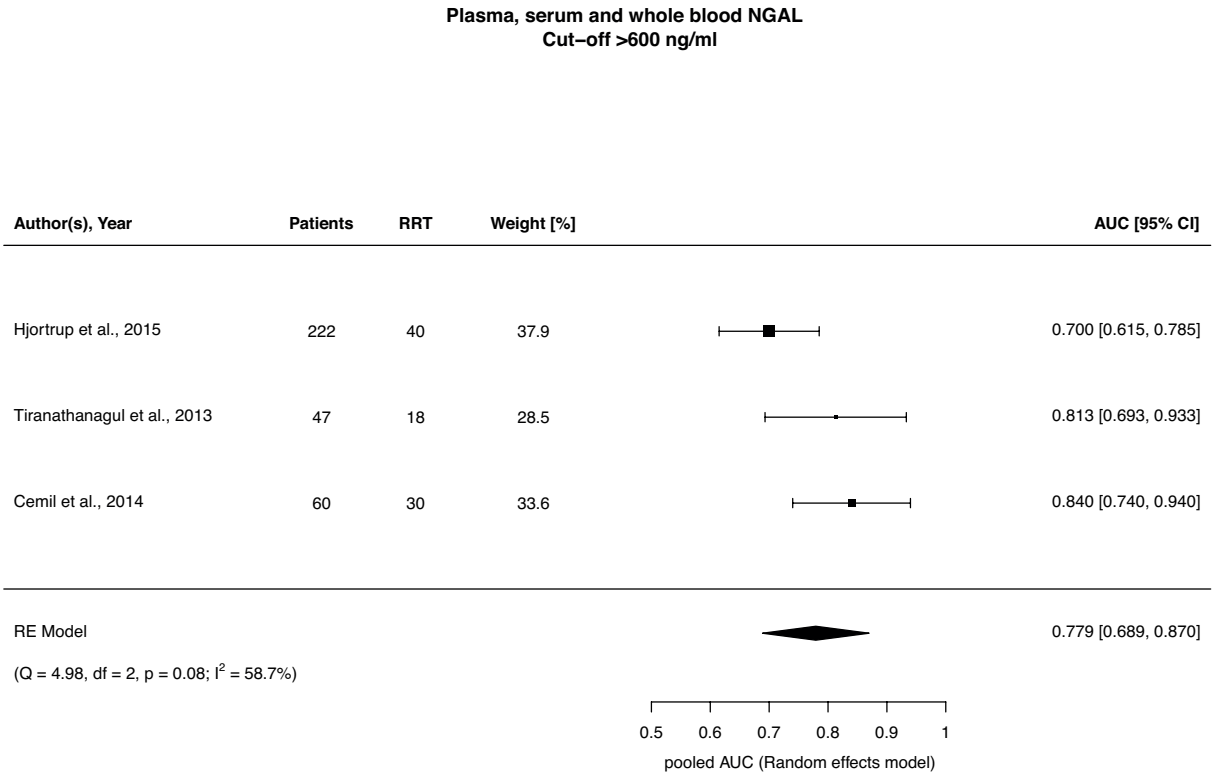


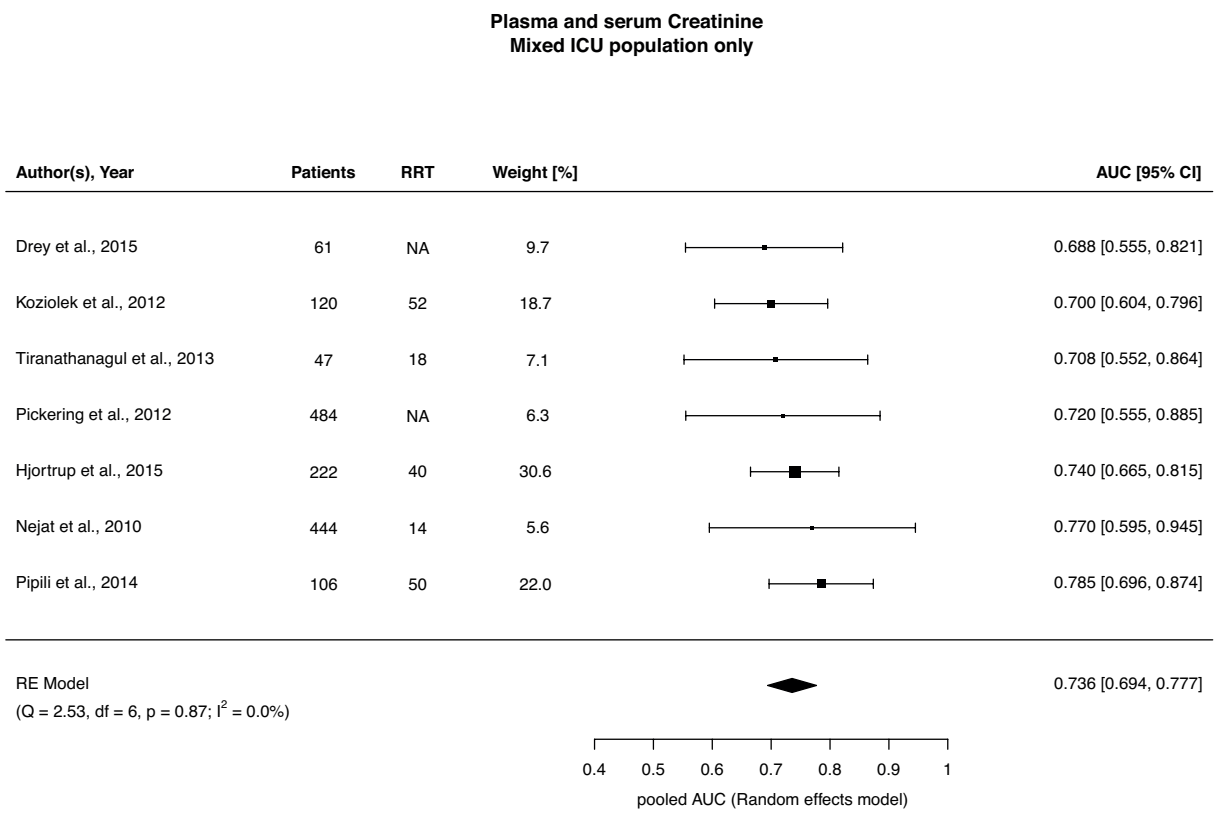
Fig. 16b Plasma, serum and whole blood NGAL. Subgroup analysis including studies with stated cut-off ≥600 ng/ml



Tab. 7 Subgroup analysis for plasma and serum creatinine. Differences between pooled AUCs

	p/s Cr (all studies)	mixed ICU populations only
AUC	0.764	0.736
(95%-CI)	(0.732-0.796)	(0.694-0.777)
Difference		0.028
p-value		0.2950

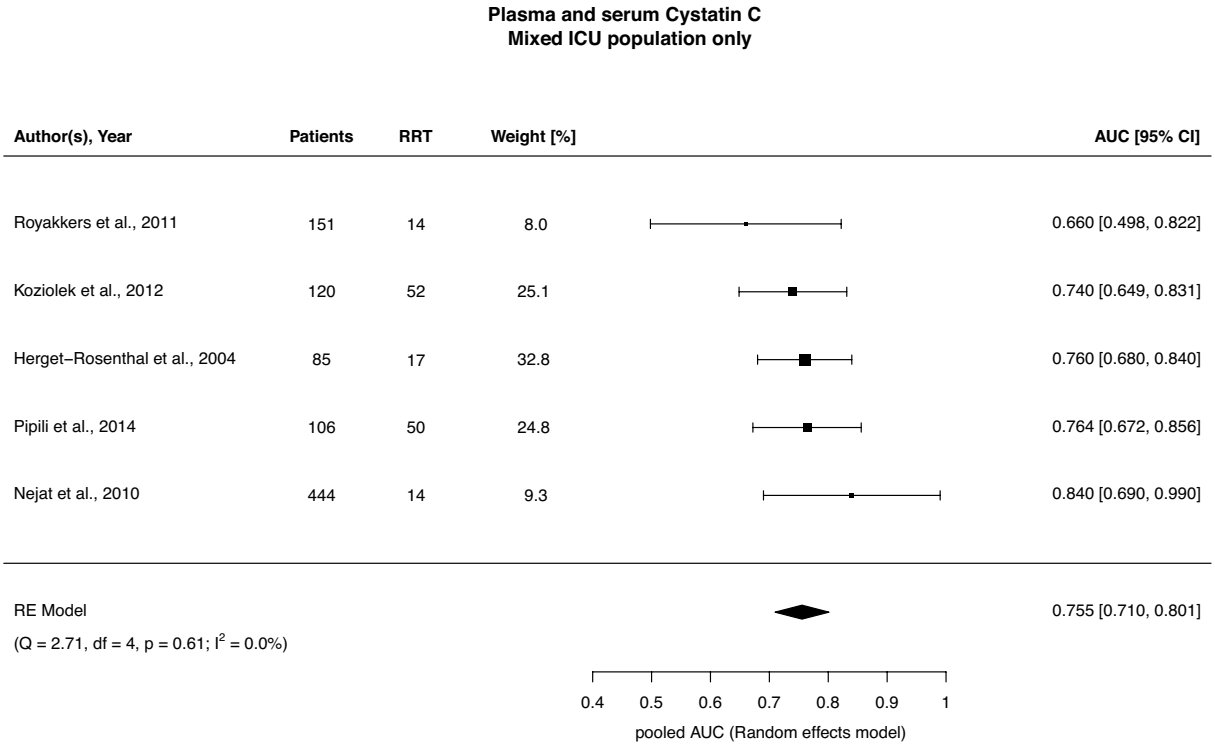
Fig. 17 Plasma and serum creatinine. Subgroup analysis including only mixed ICU populations



Tab. 8 Subgroup analysis for plasma and serum Cystatin C. Differences between pooled AUCs

	p/s CysC (all studies)	mixed ICU populations only
AUC	0.768	0.755
(95%-CI)	(0.729-0.807)	(0.710-0.801)
Difference		0.013
p-value		0.6706

Fig. 18 Plasma and serum Cystatin C. Subgroup analysis including only mixed ICU populations



Tab. 9 Subgroup analysis for urinary NGAL (concentration and normalized to urinary creatinine). Differences between pooled AUCs

	urinary NGAL (concentration) (all studies)	mixed ICU populations only	urinary NGAL (normalized) (all studies)	mixed ICU populations only
AUC	0.720	0.721	0.727	0.710
(95%-CI)	(0.638-0.803)	(0.589-0.853)	(0.678-0.776)	(0.605-0.815)
Difference		0.001		0.017
p-value		0.9900		0.7737

Fig. 19 Subgroup analysis for urinary NGAL (conc.) including only mixed ICU populations

Fig. 19a Subgroup analysis including only mixed ICU populations for the urinary concentration of urinary NGAL

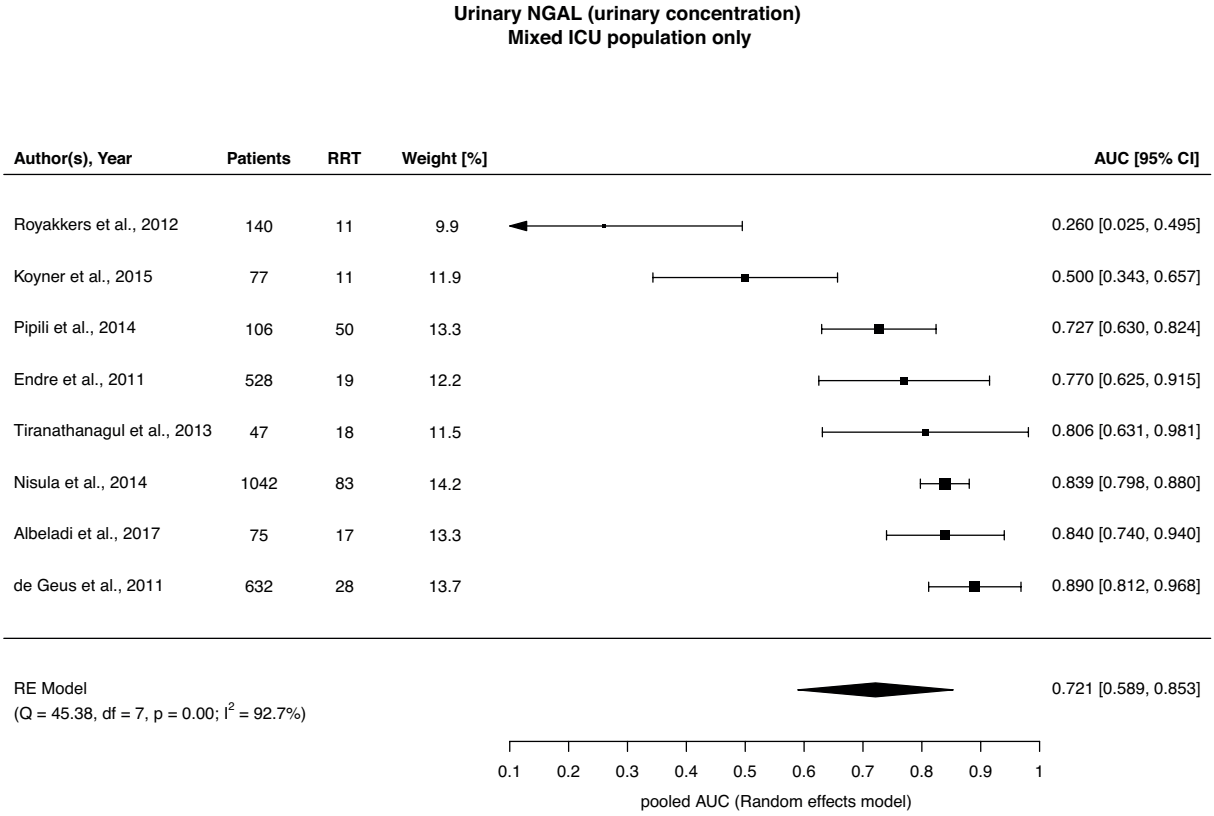


Fig. 19b Subgroup analysis including only mixed ICU populations for urinary NGAL normalized to urinary creatinine