HAL
open science

Systematic review and meta-analysis: Prognostic tests of paracetamol-induced acute liver failure

Darren George Norman Craig, Alexander Ford, Peter C Hayes, K J Simpson

To cite this version:

Darren George Norman Craig, Alexander Ford, Peter C Hayes, K J Simpson. Systematic review and meta-analysis: Prognostic tests of paracetamol-induced acute liver failure. Alimentary Pharmacology and Therapeutics, 2010, 10.1111/j.1365-2036.2010.04279.x . hal-00552537

HAL Id: hal-00552537

https://hal.science/hal-00552537v1

Submitted on 6 Jan 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Alimentary Pharmacology & Therapeutics

Systematic review and meta-analysis: Prognostic tests of paracetamol-induced acute liver failure

Journal:	*Alimentary Pharmacology & Therapeutics*
Manuscript ID:	APT-0039-2010.R1
Manuscript Type:	Systematic Review
Date Submitted by the Author:	20-Feb-2010
Complete List of Authors:	Craig, Darren; Royal Infirmary of Edinburgh, Scottish Liver Transplant Unit Ford, Alexander; St James's University Hospital, Department of Academic Medicine Hayes, Peter; Royal Infirmary of Edinburgh, Scottish Liver Transplant Unit Simpson, K; Royal Infirmary of Edinburgh, Scottish Liver Transplant Unit
Keywords:	Drug-induced liver disease < Hepatology, Liver < Organ-based, Meta-analyses < Topics, Liver transplantation < Hepatology

scholarONE™
Manuscript Central

Systematic review and meta-analysis: Prognostic tests of paracetamol-induced acute liver failure

D G N Craig[1], A C Ford[2], P C Hayes[1], K J Simpson[1]

[1]Department of Hepatology, Royal Infirmary of Edinburgh, Edinburgh, UK

[2] Department of Academic Medicine, St James's University Hospital, Leeds, UK

Corresponding author:

Dr Kenneth James Simpson

Scottish Liver Transplantation Unit

Royal Infirmary of Edinburgh

Little France

EDINBURGH

EH16 4SA

Telephone 00441312421717

Fax: 00441312421633

Email: k.simpson@ed.ac.uk

Competing Interest: None to declare

Funding sources: None to declare

Keywords Liver transplantation; King's College Criteria; acetaminophen; acute liver failure

Word count 2939

For Peer Review

ABBREVIATIONS

AFP	Alpha-fetoprotein
ALF	Acute liver failure
ALT	Alanine aminotransferase
APACHE II	Acute Physiology and Chronic Health Evaluation II
AUC	Area under the curve
DOR	Diagnostic odds ratio
HE	Hepatic encephalopathy
IL-6	Interleukin-6
KCC	King's College Criteria
KCH	King's College Hospital
MELD	Model for End-Stage Liver Disease
OLT	Orthotopic liver transplantation
ROC	Receiver operating characteristics
SIRS	Systemic inflammatory response syndrome
SOFA	Sequential Organ Failure Assessment

For Peer Review

ABSTRACT

Background: Paracetamol toxicity remains the leading cause of acute liver failure (ALF) in the developed world. In the UK, the recently modified King's College Criteria (KCC) are used to list patients for emergency liver transplantation, but these criteria have been criticised for their low sensitivity and for spectrum bias in their application.

Aim: To critically evaluate existing prognostic criteria for predicting death without transplantation in paracetamol-induced ALF.

Methods: MEDLINE, EMBASE and CINAHL were searched to identify studies containing adult patients with paracetamol-induced ALF. Selected studies were evaluated, and data were pooled if appropriate, to calculate sensitivity, specificity, and diagnostic odds ratios (DORs) of applied prognostic tests.

Results: Of 6507 studies identified, 14 were eligible for inclusion, evaluating 1960 patients. The original KCC had a pooled sensitivity of 58.2% and specificity of 94.6%, with a DOR of 27.7. Addition of arterial lactate to the KCC reduced the DOR to 26.1. Several other clinical and laboratory variables had higher DORs than the KCC, but were only evaluated in single studies of limited quality.

Conclusion: The original KCC remain well validated criteria with high prognostic accuracy. Other potential prognostic variables should be prospectively assessed in multi-centre studies to further refine the criteria.

For Peer Review

INTRODUCTION

Paracetamol (acetaminophen) toxicity remains the leading cause of acute liver failure (ALF) in the developed world, accounting for over 40% of cases in selected case series.(1,2) Whilst the vast majority of patients recover spontaneously following paracetamol overdose, a small number develop severe acute liver injury, hepatic encephalopathy (HE) and consequently, ALF. Despite significant advances in supportive care, the only effective treatment for the condition remains emergency orthotopic liver transplantation (OLT).(3)

The decision to transplant a patient with paracetamol-induced ALF involves balancing the inherent risks associated with delaying listing for OLT against the potential for spontaneous recovery with medical therapy alone, the risks of surgery in the context of a rapidly evolving critical illness, the scarcity of donor grafts, and the requirement for lifelong immunosuppression. Furthermore, the psychosocial implications of paracetamol overdose cases are considerable, with over 30% of patients that fulfil transplant criteria unsuitable for OLT due to severe psychological illness, or coexistent chronic alcohol or drug dependency.(4) Accurate prognostication in ALF is therefore vital to utilise OLT effectively and prevent unnecessary transplantation, whilst procuring donor organs in a timely fashion for those most likely to benefit.

In 1989 O'Grady and colleagues developed the 'King's College Criteria' (KCC) in an attempt to determine which patients with paracetamol and non-paracetamol-induced ALF have a poor prognosis with medical therapy alone, and will therefore benefit most from OLT.(5) The original KCC for paracetamol-induced ALF were highly specific, but have been criticised for their relatively low negative predictive

For Peer Review

value (6) and, since up to 26% of patients with paracetamol toxicity are medically unfit to undergo surgery at the point they fulfil the KCC,(4) modifications to these criteria now utilise arterial lactate (>3.5 mmol/L after early fluid resuscitation or >3.0 mmol/L after adequate fluid resuscitation) in an attempt to extend the time-window for acquisition of a suitable graft.(7)

Over recent years a plethora of alternative prognostic variables have been proposed, in an attempt to improve, or replace, the KCC. These criteria variously involve radiological,(8) histological,(9) serological,(10,11) or mathematical (12-14) indices, but their assessment is complicated by the use of OLT; due to the inevitably imperfect nature of current listing criteria a small proportion of ALF patients are transplanted who would have spontaneously survived, invalidating further assessment of that particular patient. Studies evaluating the prognostic accuracy of a particular variable should therefore exclude transplanted patients from subsequent analysis. The purpose of this systematic review was to critically evaluate existing prognostic criteria for predicting death without transplantation in paracetamol-induced ALF, to update previous studies in this area,(15,16) and to examine the accuracy of recently described prognostic variables.

For Peer Review

METHODS

Search Strategy and Study Selection

We conducted a systematic review of the medical literature using MEDLINE (1950 to June 2009), EMBASE (1980 to June 2009), and CINAHL (1982 to June 2009) to identify studies recruiting adult (>age 16 years) patients that evaluated prognostic markers of paracetamol-induced ALF. Potential studies were identified by combining the search terms *"acetaminophen"* and *"paracetamol"* (as both Medical Subject Headings (MeSH) and free text terms) using the set operator OR. Prognostic studies of ALF were identified, and combined using the set operator OR, by using the terms *"Kings adj3 College adj3 Criteria"*, *"Clichy$"*, *"APACHE$"*, *"lactate"*, *"Gc-globulin"*, *"alpha-fetoprotein"*, *"phosphate"* and *"MELD"* (all free text terms) and the McMaster expert search strategy for prognostic studies. These studies were combined using the set operator AND with papers evaluating studies on *"liver failure"*, *"hepatic encephalopathy"* or *"liver failure, acute"* (all MeSH). The search was limited to human studies without language restrictions. Abstracts of the studies identified by the initial search were evaluated for appropriateness to the study question by two independent reviewers (DC, KS) and all potentially relevant papers were obtained and evaluated in detail. The bibliographies of these studies and relevant review articles were screened to perform a recursive search of the literature, and abstract books of international liver conferences from the preceding 4 years were hand-searched for potentially relevant studies.

Selected studies were required to report mortality data on cohorts of patients admitted to hospital with acute severe liver injury or ALF secondary to paracetamol overdose (**see Box for study eligibility criteria**). Acute severe liver injury was defined as severe hepatotoxicity (serum alanine aminotransferase (ALT) levels>1000

For Peer Review

U/mL and co-existent coagulopathy), whilst ALF required the additional presence of HE, in patients without pre-existing liver disease. Unselected cohorts of patients with mixed etiologies of acute liver injury were eligible for inclusion if separate data for paracetamol overdose patients were available. Case control studies and studies comparing ALF patients with chronic liver disease patients were excluded due to potential bias in favour of the prognostic variable in question,(17) as were studies where transplantation and death were combined as a single end-point. Studies were required to include more than 25 patients with paracetamol-induced acute liver injury, since the majority of studies in this area include the KCC as a prognostic variable in addition to the main variable studied, and smaller studies may invalidate analysis.(18) All potential articles were assessed independently by two researchers according to eligibility criteria, which were defined prospectively, and disagreements resolved by consensus.

For Peer Review

Data extraction and quality assessment

Data were extracted using predesigned forms by two separate reviewers (DC and KS) on to a Microsoft Excel spreadsheet (Microsoft Corp, Redmond, WA, USA), and any discrepancies were resolved by consensus. The following data were extracted for each study: setting, country and geographical region, year(s) conducted, retrospective or prospective design, inclusion criteria, definitions of paracetamol-induced liver injury, definitions of ALF, total number of subjects included, total number of subjects with paracetamol-induced ALF or acute liver injury, total number of subjects with paracetamol-induced ALF transplanted, duration of follow-up, prognostic score used, timing of application of the prognostic score, and outcome definitions. Study quality was assessed semi-quantitatively according to six

potential sources of bias that can be encountered in studies of prognostic variables (19) and a grade (poor, moderate, good or excellent) allocated (**Table 1**).

Study population	Study population adequately represents the population of interest; source population adequately described; sampling timeframe, place and period of recruitment described; inclusion and exclusion criteria described
Follow-up	Adequate study completion rate; loss to follow-up not associated with key characteristics; reasons for loss to follow-up described; no important differences between participants completing/not completing study
Prognostic factor	Prognostic measure clearly defined; continuous data reported and any cut-points described and appropriate; method and setting of measurement of prognostic factor identical for all patients; adequate proportion of study population has prognostic factor measured
Outcome measurement	Clear definition of outcome measure of interest described and recorded, including duration of follow-up; outcome measure valid; outcome measure identical
Confounding	All important confounders measured and accounted for at all stages of study design, performance and analysis
Analysis	Sufficient presentation of data to permit assessment of analytical analysis; appropriate model-building, study design etc.; no selective reporting of results

Table 1 Quality assessment of included studies (adapted from Hayden 2006 (19))

For Peer Review

Data synthesis and analysis

Prognostic tests evaluating the outcomes of survival without transplantation or death without transplantation following paracetamol-induced liver injury were assessed individually. From each study cohort we extracted total number of subjects dying without OLT, total number of subjects surviving without OLT, and total

numbers of subjects fulfilling or not fulfilling the prognostic variable in question, and constructed 2x2 contingency tables. The sensitivity, specificity, and diagnostic odds ratio (DOR) with their 95% confidence intervals (CI) were calculated for each test using Microsoft Excel (Microsoft Corp, Redmond, Washington, USA) and checked using Meta-DiSc version 1.4 (Universidad Complutense, Madrid, Spain). The value of a DOR ranges from zero to infinity, with higher values indicating better discriminatory test performance, and is calculated from the following formula: sensitivity/(1- sensitivity) / (1- specificity)/specificity.(20) Where required, a zero cell correction of 0.5 was added to all cells to prevent computational problems arising where proportions were equal to zero.(21) Where two or more studies analysed a particular prognostic marker, data were pooled using a random effects model and the pooled DOR calculated. Heterogeneity between studies was assessed using the I^2 statistic with a cut off of 50%, and the χ^2 test with a P value < 0.10, used to define a statistically significant degree of heterogeneity. We explored study setting and study design as potential reasons for heterogeneity.(22) These are exploratory analyses only, and the results should therefore be interpreted with caution.

For Peer Review

RESULTS

The search strategy identified 6507 studies of which 105 were potentially eligible for inclusion and were retrieved for further analysis (**Figure 2**). Of these, 14 were considered eligible for inclusion,(5,7,10,23-33) including two studies published in abstract form only.(32,33) Characteristics of included studies are provided in **Table 2**. Cohen's kappa test for inter-observer agreement was excellent at 0.86 (95% CI 0.75-0.98).

Study	Country	Liver Unit	Study period	Inclusion criteria	Prognostic test(s) evaluated	No paracetamol patients included	Retrospective/ prospective cohort	Retrospective/ prospective test development/ evaluation	Study quality
O'Grady 1989(5)	UK	KCH	1973-1985; 1986-7	FHF as per Trey & Davidson(34) with Grade III-IV HE	pH<7.3 Concurrent PT>100 seconds, serum creatinine> 300 mmol/L, Grade III/IV HE	121 99	Both	Prospective	Moderate
O'Grady 1991(23)	UK	KCH	1988-1990	Severe liver damage	KCC as per O'Grady (5)	60	Prospective	Prospective	Moderate
Izumi 1996 (24)	UK	KCH	Not stated	FHF as per Trey& Davidson (34)	KCC as per O'Grady (5) Factor V ratio <20% Factor V ratio <10%	} 81	Prospective	Prospective (partially)	Poor
Anand 1997(25)	UK	Birmingham	1990-1994	FHF as per Trey & Davidson (34)	pH<7.3 Concurrent PT>100 seconds, serum creatinine> 300 mmol/L, Grade III/IV HE	72 89	Retrospective	Prospective	Poor
Bernal 1998(26)	UK	KCH	1990-1996	Severe hepatotoxicity (KCC); Death whilst not meeting KCC (APACHE III)	KCC as per O'Grady(5) Adapted APACHE III	504 56	Retrospective	Prospective (KCC) Retrospective (APACHE III)	Poor
Mitchell 1998(27)	UK	KCH	1993-1994	Coagulopathy + recent history paracetamol ingestion	KCC as per O'Grady(5) APACHE II>15 at 24 hr post-admission APACHE II>15	} 94	Prospective	Prospective (KCC) Retrospective (APACHE II)	Good
Bernal 2002(7)	UK	KCH	1998-1999 (learning set) 1999-2000 (validation set)	Severe paracetamol-induced hepatotoxicity	KCC as per O'Grady(5) Arterial lactate>3.5 mmol/L at 4 hours Arterial lactate>3.0 mmol/L at 12 hours	99 97 85	Both	Prospective	Moderate

For Peer Review

					KCC + combination of lactate criteria	85			
Bernal 2003(28)	UK	KCH	1998-2000	Acute severe hepatotoxicity	KCC as per O'Grady(5) Phosphate >1.2 mmol/L on admission day +1 or 2	} 170	Retrospective	Prospective	Moderate
Larson 2005(29)	USA	22 academic centres	1998-2003	INR>1.4; HE; jaundice to HE interval<26 weeks	KCC as per O'Grady(5)	252	Prospective	Prospective	Moderate
Schmidt 2005(10)	Denmark	Copenhagen	1999-2002	Peak ALT>1000U/L	KCC as per O'Grady(5) AFP <3.9 at D1 post peak ALT AFP <3.9 & INR>2.4 at D1 post peak ALT	234 188 188	Prospective	Prospective (KCC) Retrospective (AFP)	Moderate
Schmidt 2006(30)	Denmark	Copenhagen	1999-2004	Severe paracetamol-induced FHF	KCC as per O'Grady(5) Arterial lactate Modified KCC SOFA score >8 at admission/ >12 at onset of Grade III/IV HE SIRS at admission/at onset of Grade III/IV HE	95 91 91 95 95	Prospective	Retrospective	Moderate
Zaman 2006(31)	Ireland	Dublin	1994-2005	Paracetamol-induced ALF (jaundice to HE <8 weeks) or rapid bilirubin/ INR/ renal impairment/ hypoglycaemia if no HE	KCC as per O'Grady(5) MELD >30	} 60	Retrospective	Prospective	Poor
Bates 2007(32)	UK	Edinburgh	2004-2007	Acute severe liver injury	KCC as per O'Grady(5) Lactate modifications to KCC as per Bernal(7)	69 32	Retrospective	Prospective	Poor
Bernal 2007(33)	UK	KCH	Not stated	Acute liver failure	IL-6 within 24 hours of admission	31	Prospective	Retrospective	Poor

Table 2 Characteristics of included studies

For Peer Review

PT, prothrombin time; FHF, fulminant hepatic failure

For Peer Review

The eligible studies evaluated a total of 1960 patients with paracetamol-induced acute liver injury or ALF. Three (10,24,27) studies had complete temporal overlap with other studies from the same unit; however, these studies were included for their evaluation of unique prognostic markers separate to the KCC. There was only one multicentre study.(29) Five (25,26,28,31,32) studies evaluated patient cohorts retrospectively, whilst five (10,26,27,30,33) studies developed prognostic test thresholds retrospectively (usually from receiver operator characteristics (ROC) curve analysis). No studies blinded observers to patient outcome or other potentially confounding prognostic data; and only two (5,7) validated their prognostic marker in a separate cohort. Consequently, six (24-26,31-33) of the 14 studies were graded as poor and seven (5,7,10,23,28-30) as moderate quality (**Table 2**). The 14 eligible studies analysed a total of 22 different prognostic markers or variations thereof. The sensitivity, specificity, and DORs of these are reported in **Table 3**.

For Peer Review

Prognostic test	Study	N/deaths	Test +ve/ deaths	Sensitivity (95% CIs)	Specificity (95% CIs)	Diagnostic Odds Ratio (95% CIs)	Heterogeneity	
							X^2	I^2
pH<7.3	O'Grady 1989(5)	121/43	22/21	48.8 (33.3-65.5)	98.7 (93.1-100)	73.5 (9.4-577.4)	4.91 p=0.027	80%
	Anand 1997(25)	72/39	31/24	61.5 (44.6-76.6)	78.8 (61.1-91.0)	5.9 (2.1-17.1)		
	POOLED			**54.9 (43.5-65.9)**	**92.8 (86.3-96.8)**	**18.0 (1.1-229.6)**		
Concurrent PT>100s, serum creatinine>300 mmol/L, Grade III/IV HE	O'Grady 1989(5)	99/22	15/10	45.5 (24.4-67.8)	93.5 (85.5-97.9)	12.0 (3.5-41.3)	0.8 p=0.374	0%
	Anand 1997(25)	89/45	24/19	42.2 (27.7-57.9)	88.6 (75.4-96.2)	5.7 (1.9-17.2)		
	POOLED			**43.3 (31.2-56.0)**	**91.7 (85.3-96.0)**	**7.9 (3.5-18.1)**		
KCC (combined)†	O'Grady 1989(5)	220/65	37/31	47.7 (35.1-60.5)	96.1 (91.8-98.6)	22.6 (8.8-58.6)	8.4 p=0.038	64%
	O'Grady 1991(23)	60/26	23/19	73.1 (52.2-88.4)	88.2 (72.6-96.7)	20.4 (5.2-79.0)		
	Bernal 1998(26)	504/99	80/71	71.7 (61.8-80.3)	97.8 (95.8-99.0)	111.6 (50.5-246.4)		
	Bernal 2002(7)	99/21	20/16	76.2 (52.8-91.8)	94.9 (87.4-98.6)	59.2 (14.3-245.3)		
	POOLED (KCH)			**65.9 (59.0-72.3)**	**96.1 (94.4-97.5)**	**43.9 (17.6-109.3)**		
	Larson 2005(29)							
	(at admission)	252/74	34/19	25.7 (16.2-37.2)	91.6 (86.5-95.2)	3.8 (1.8-7.9)		
	Schmidt 2006(30)	95/48	36/27	77.1 (62.7-88.0)	83.0 (69.2-92.4)	16.4 (5.9-45.3)		
	Zaman 2006(31)	60/29	21/21	72.4 (52.8-87.3)	100.0 (88.8-100.0)	159.4 (8.7-2908.9)		
	Bates 2007(32)	69/17	18/15	88.2 (63.6-98.5)	94.2 (84.1-98.8)	122.5 (18.7-803.1)		
	POOLED (non-							

For Peer Review

	KCH)			**48.8 (41.0-56.6)**	**91.2 (87.5-94.1)**	**16.5 (3.5-77.8)**	17.0 p=0.001	82%
	POOLED (overall)			**58.2 (53.1-63.3)**	**94.6 (93.0-95.9)**	**27.7 (9.2-83.5)**	53.0 p<0.001	87%
Arterial lactate >3.5 at admission	Bernal 2002(7)	97/21	18/14	66.7 (43.0-85.4)	94.7 (87.1-98.6)	36.0 (9.3-139.6)	4.54 p=0.103	56%
	Schmidt 2006(30)	91/46	61/39	84.8 (71.1-93.7)	51.1 (35.8-66.3)	5.8 (2.2-15.8)		
	Bates 2007(32)	69/17	37/15	88.2 (63.6-98.5)	57.7 (43.2-71.3)	10.2 (2.1-49.4)		
	POOLED			**81.0 (70.9-88.7)**	**72.3 (64.9-78.8)**	**12.2 (4.0-37.4)**		
Arterial lactate >3.0 following resuscitation	Bernal 2002(7)	85/21	18/16	76.2 (52.8-91.8)	96.9 (89.2-99.6)	99.2 (17.6-559.3)	9.83 p=0.007	80%
	Schmidt 2006*(30)	91/46	55/36	78.3 (63.6-89.1)	57.8 (42.2-72.3)	4.9 (2.0-12.3)		
	Bates 2007(32)	32/12	20/12	100.0 (73.5-100.0)	60.0 (36.1-80.9)	36.8 (1.9-708.0)		
	*at onset of Grade III/IV HE							
	POOLED			**81.0 (70.6-89.0)**	**77.5 (69.3-84.4)**	**22.8 (2.5-210.0)**		
Arterial lactate >4.0 at admission	Schmidt 2006 (30)	91/46	51/34	73.9 (58.9-85.7)	62.2 (46.5-76.2)	4.7 (1.9-11.4)	-	-
Arterial lactate >4.0 at onset of Grade III/IV HE			40/31	67.4 (52.0-80.5)	80.0 (65.4-90.4)	8.3 (3.2-21.5)		
KCC + arterial lactate >3.0 following resuscitation	Bernal 2002(7)	85/21	24/19	90.5 (69.6-98.8)	92.2 (82.7-97.4)	112.1 (20.1-625.7)	8.83 p=0.003	89%
	Schmidt 2006(30)	91/46	69/42	91.3 (79.2-97.6)	40.0 (25.7-55.7)	7.0 (2.1-22.9)		
	POOLED			**91.0 (81.5-96.6)**	**70.6 (61.2-79.0)**	**26.1 (1.7-393.7)**		

For Peer Review

Factor V ratio <20% Factor V ratio <10%	Izumi 1996(24)	81/35	69/34 51/29	97.1 (85.1-99.9) 63.0 (47.6-76.8)	23.9 (12.6-38.8) 37.1 (21.5-55.1)	10.7 (1.3-87.3) 5.3 (1.8-15.1)	-	-
Adapted APACHE III	Bernal 1998(26)	56/28	19/16	57.1 (37.2-75.5)	89.3 (71.8-97.7)	11.1 (2.7-45.6)	-	-
APACHE II>15 at 24 hr post-admission APACHE II>15 within 5 days of admission	Mitchell 1998(27)	94/14	13/11 23/13	78.6 (49.2-95.3) 92.9 (66.1-99.8)	97.5 (91.3-99.7) 87.5 (78.2-93.8)	143.0 (21.4 -953.5) 91.0 (10.7-772.8)	-	-
Phosphate >1.2 mmol/L on admission day +1 or 2	Bernal 2003(28)	170/52	55/42	80.8 (67.5-90.4)	89.0 (81.9-94.0)	33.9 (13.8-83.3)	-	-
AFP <3.9 at D1 post peak ALT AFP <3.9 & INR>2.4 at D1 post peak ALT	Schmidt 2005(10)	188/33	54/33 74/33	100.0 (89.4-100.0) 100.0 (89.4-100.0)	73.6 (65.9-80.3) 86.5 (80.0-91.4)	184.9 (11.1-3085.6) 419.1 (24.8-7097.5)	-	-
SOFA score >8 at admission SOFA score >12 at onset of Grade III/IV HE	Schmidt 2006(30)	95/48	48/32 54/39	66.7 (51.6-79.6) 81.3 (67.4-91.1)	66.0 (50.7-79.1) 68.1 (52.9-80.9)	3.9 (1.7-9.1) 9.2 (3.6-23.9)	-	-

For Peer Review

SIRS at admission SIRS at onset of Grade III/IV HE	Schmidt 2006(30)	95/48	52/33 41/34	68.8 (53.7-81.3) 70.8 (55.9-83.0)	59.6 (44.3-73.6) 85.1 (71.7-93.8)	3.2 (1.4-7.5) 13.9 (5.0-38.3)	-	-
MELD >30	Zaman 2006(31)	60/29	38/28	96.6 (82.2-99.9)	67.7 (48.6-83.3)	58.8 (7.0-495.8)	-	-
IL-6 within 24 hours of admission	Bernal 2007(33)	31/8	7/6	75.0 (34.9-96.8)	95.7 (78.1-99.9)	66.0 (5.1-857.7)	-	-

Table 3 Sensitivity, specificity, and diagnostic odds ratios of individual prognostic markers.

† Anand 1997 (25) excluded as 2x2 table not reconstructable; Mitchell 1998,(27) Bernal 2003,(28) and Schmidt 2005(10) excluded due to complete temporal overlap with previous studies from same unit

For Peer Review

King's College Criteria

A total of 13 studies evaluated the original KCC for paracetamol-induced ALF, either as separate elements (arterial pH<7.3 or concurrent grade III/IV HE, serum creatinine>300µmol/L and prothrombin time (PT)>100 seconds) (5,25) or as a whole, (7,10,23,24,26-32) including a total of 1929 patients. One study (29) evaluated the KCC only on admission. After exclusion of studies with complete temporal overlap with other studies,(10,24,27) pooled specificity of the KCC was high, at 94.6% (95% CI 93.0-95.9), but the pooled sensitivity was relatively poor at 58.2% (95% CI 53.1-63.3). The pooled DOR for the KCC was 27.7 (95% CI 9.2-83.5) (**Table 3**). The summary ROC curve is shown in **Figure 3**. The area under the curve (AUC) was calculated as 0.91 (95% CI 0.79-0.99), suggesting good performance of the KCC overall.(35) The accuracy of the KCC improved when studies originating from KCH were analysed separately from those outside KCH (DOR 43.9 (95% CI 17.6-109.3) for KCH-based studies vs. DOR 16.5 (95% CI 3.5-77.8) for non-KCH-based studies). Statistically significant heterogeneity (I^2=87%) existed between individual study results, though this fell to 79% after exclusion of the single multi-centre study,(29) and to 64% after exclusion of studies from outside KCH.

Lactate modifications to the KCC

The original study (7) that evaluated arterial lactate in the prognosis of paracetamol-induced ALF reported similar specificity, but improved sensitivity, when compared with the original KCC. However, two (30,32) subsequent studies evaluating arterial lactate alone failed to replicate this high prognostic accuracy (**Table 3**), with pooled DORs of 12.2 (95% CI 4.0-37.4) and 22.8 (95% CI 2.5-210.0)

For Peer Review

for the early and post-resuscitation lactate values respectively. This was mainly due to the reduced specificity seen in the two studies from outside KCH. Combination of the KCC with a post-resuscitation lactate value >3.0 offered higher prognostic accuracy in the original lactate study (7) but this was not replicated in a subsequent evaluation.(30)

Other prognostic markers

Several other markers, all evaluated in single studies, appeared to offer improved prognostic accuracy when compared with the KCC, with low serum alpha-fetoprotein (AFP) (DOR 419.1),(10) 24-hour post-admission Acute Physiology and Chronic Health Evaluation II (APACHE II) score (DOR 143.0),(27) serum interleukin-6 (IL-6) levels (DOR 66.0),(33) Model for End-Stage Liver Disease (MELD) score (DOR 58.8),(31) and serum phosphate (DOR 33.9) (28) all outperforming the KCC (**Table 3**). However, further studies would be required to replicate these findings to confirm or refute whether these other markers are indeed superior to the KCC.

DISCUSSION

This systematic review and meta-analysis has demonstrated that the original KCC for paracetamol-induced ALF have high pooled specificity (94.6%), but low pooled sensitivity (58.2%) in determining prognosis in patients with paracetamol-induced ALF. Additionally, the benefit of the arterial lactate modifications to the KCC (7) is questionable according to these data. Other proposed prognostic markers, in

For Peer Review

particular AFP, APACHE II scores, and serum IL-6 levels, showed encouraging prognostic accuracy but were only evaluated in single studies of variable quality.

This review is limited by the quality of the included studies, which had significant heterogeneity and were generally of poor or moderate quality, with only two (5,7) studies validating their prognostic model prospectively in a separate cohort. Acute liver failure is a rare syndrome and, as a result, many studies were small and retrospective in nature. Continuous variables were frequently analysed retrospectively using cut-off values designed to maximise AUC values, a method which assumes constant risk amongst the 'high' and 'low' risk groups.(18) Several studies attempted to evaluate multiple prognostic markers retrospectively using the same cohort, an approach which increases the risk of obtaining a statistically significant result by chance. Another potentially confounding issue is the lack of international standardisation of laboratory variables, such as prothrombin time and creatinine, and changes to assay reagents over the time course of these studies.(36) We also appreciate that exclusion of studies where transplantation and death were deemed equivalent may have introduced spectrum bias, since prognostic tests are usually applied in settings where OLT is available. This problem is difficult to circumvent given that emergency OLT will never be subjected to a clinical trial, but, given the reduced quality-of-life seen following transplantation for paracetamol-induced ALF, the accumulated risks of immunosuppression, and the scarcity of liver donors,(16) accurate calculation of the specificity of each prognostic test is vital in order to minimise inappropriate transplantation.

This study expands upon a recent systematic review (16) evaluating the KCC for paracetamol-induced ALF by including additional prognostic markers. The former study found spectrum bias in studies originating from KCH, with increased

For Peer Review

spontaneous survival amongst patients listed for transplantation, but not receiving a graft, compared with those meeting criteria but never listed, suggesting that a 'healthier' cohort may be preferentially transplanted in KCH. The current data confirm the increased prognostic accuracy of the original KCC in studies originating from KCH compared with studies from other units, with a higher pooled DOR in KCH-based studies compared with that from studies conducted in other units. This raises further questions regarding the overall generalisability of the KCC to patients with paracetamol-induced ALF treated outside KCH.

The heterogeneity of the included studies evaluating the KCC was partly due to a multicentre US study (29) which explicitly applied the KCC solely at the time of admission, rather than dynamically throughout admission as originally intended.(5,37) Whilst early and accurate prognostication in ALF is vital in order to permit timely listing for OLT, application of the KCC solely at admission reduces the time available for the disease to evolve and may introduce spectrum bias. Given that the median time taken to fulfil the KCC following admission to a tertiary liver transplant unit is 12 hours,(7) application of the KCC solely at admission may explain the low sensitivity of the criteria in the US study, whilst concerns have also been expressed about the prophylactic use of fresh frozen plasma in this US study and the potential confounding effects upon prothrombin time.(38) Furthermore, 48% of cases resulted from unintentional overdoses in this US study, a type of overdose seen less frequently in the UK.(39) It may be that repeated ingestion of supratherapeutic doses of paracetamol over a protracted time course disrupts liver function subacutely, so that patients are less likely to develop the profound acidosis or concurrent severe coagulopathy, renal dysfunction, and encephalopathy required to fulfil the KCC, but have at least as poor a prognosis as an intentional overdose at a

For Peer Review

single time point. Future prognostic scoring systems may therefore need to take the pattern of paracetamol overdose into account in addition to traditional biochemical parameters.

The addition of post-resuscitation arterial lactate to the KCC (7) is questioned by this study. Only two additional (30,32) studies, one of which was reported in abstract-form, that reported this modification fulfilled eligibility criteria for inclusion, but these studies suggest little benefit from this. Furthermore, the reduced specificity of the lactate criteria undermines the traditional benefit of the KCC in 'ruling in' a hopeless prognosis. The presence of systemic inflammation, with or without sepsis, is increasingly recognised as important in ALF,(40) but hyperlactataemia can result from numerous other organ sources and therefore, perhaps not surprisingly, arterial lactate is a relatively non-specific prognostic indicator in paracetamol-induced ALF. Within critical care settings, the use of APACHE II and Sequential Organ Failure Assessment (SOFA) scores to monitor organ dysfunction have greater recognition than the KCC, and are attractive as early prognostic markers in ALF.(41) APACHE II, in particular, showed encouraging prognostic accuracy but was only evaluated in a single eligible study.(27) One (29) additional study evaluating APACHE II scores was excluded from this particular analysis as the prognostic scoring reported in the study could not be reconstructed into a 2x2 table; this study is also notable for the atypical nature of the patient cohort as outlined above. Given that APACHE III, systemic inflammatory response syndrome (SIRS) and SOFA scores all performed less well than the KCC (albeit in single studies), the role of these markers may be to permit earlier identification of a high-risk cohort requiring transfer to tertiary centres that offer liver transplantation, rather than as definitive transplant listing criteria.

For Peer Review

Persistently elevated serum phosphate or reduced AFP levels may reflect failure of hepatic regeneration, and therefore could help predict a poorer outcome in paracetamol-induced ALF. Serum AFP showed high prognostic performance in one (10) study, whilst serum phosphate showed equivalence with the KCC in one (28) retrospective study from KCH. Other authors, (in studies where transplantation was equated with death and hence excluded from this study), have demonstrated conflicting results with serum phosphate,(42-45) so further evaluation of both serum AFP and phosphate in future studies would be worthwhile. The MELD scoring system has been widely adopted for organ allocation in chronic liver disease and has shown encouraging prognostic accuracy in non-paracetamol ALF.(12,46-48) More limited data exist regarding the use of MELD in paracetamol-induced ALF, but one additional study (excluded from this analysis as the 2x2 table was not reconstructable) suggested that its use may be limited by a high false positive rate.(49)

In summary, this systematic review and meta-analysis has demonstrated that the original KCC for paracetamol-induced ALF have high pooled specificity, but low pooled sensitivity in determining prognosis in patients with paracetamol-induced ALF. The KCC had reduced prognostic accuracy when applied outside of KCH and were occasionally applied only at admission. The reduced specificity of the KCC following the addition of arterial lactate calls into question the benefit of this modification, suggesting that re-evaluation of this as a prognostic marker is required. Urgent consideration should be given to the design of a high-quality, prospective study evaluating the KCC, APACHE II scores, and markers of hepatic regeneration, such as serum AFP and phosphate, in paracetamol-induced ALF. Given the relatively rare nature of ALF, such a study is likely to require cooperation between

For Peer Review

several large centres, and argues the need for a collaborative network of tertiary hospitals experienced in the management and prognostication of ALF, similar to that developed by the Acute Respiratory Distress Syndrome Network program.(50)

For Peer Review

Figure legends

Figure 1 Eligibility criteria for included studies

Figure 2 Flow diagram of assessment of studies identified in the systematic review

Figure 3 Summary receiver operator curve (SROC) of studies evaluating the original KCC

AUC: Area under the curve; SE: Standard Error; Q: heterogeneity

For Peer Review

Competing Interest: None to declare

Funding sources: None to declare

For Peer Review

(1) Ostapowicz G, Fontana RJ, Schiodt FV, Larson A, Davern TJ, Han SH, et al. Results of a prospective study of acute liver failure at 17 tertiary care centers in the United States. Ann.Intern.Med. 2002 Dec 17;137(12):947-954.

(2) Bernal W, Cross TJ, Auzinger G, Sizer E, Heneghan MA, Bowles M, et al. Outcome after wait-listing for emergency liver transplantation in acute liver failure: a single centre experience. J.Hepatol. 2009 Feb;50(2):306-313.

(3) Craig DG, Lee A, Hayes PC, Simpson KJ. Review article: the current management of acute liver failure. Aliment.Pharmacol.Ther. 2010;31(3):345-358.

(4) Simpson KJ, Bates CM, Henderson NC, Wigmore SJ, Garden OJ, Lee A, et al. The utilization of liver transplantation in the management of acute liver failure: comparison between acetaminophen and non-acetaminophen etiologies. Liver Transplantation 2009;15:600-609.

(5) O'Grady JG, Alexander GJ, Hayllar KM, Williams R. Early indicators of prognosis in fulminant hepatic failure. Gastroenterology 1989;97(2):439.

(6) Renner EL. How to decide when to list a patient with acute liver failure for liver transplantation? Clichy or King's College criteria, or something else? J.Hepatol. 2007 Apr;46(4):554-557.

(7) Bernal W, Donaldson N, Wyncoll D, Wendon J. Blood lactate as an early predictor of outcome in paracetamol-induced acute liver failure: a cohort study. Lancet 2002 Feb 16;359(9306):558-563.

(8) Yamagishi Y, Saito H, Tada S, Horie Y, Kato S, Ishii H, et al. Value of computed tomography-derived estimated liver volume/standard liver volume ratio for predicting the prognosis of adult fulminant hepatic failure in Japan. J.Gastroenterol.Hepatol. 2005 Dec;20(12):1843-1849.

(9) Baeg N-, Bodenheimer J, H.C, Burchard K. Long-term sequellae of acetaminophen-associated fulminant hepatic failure: Relevance of early histology. Am.J.Gastroenterol. 1988;83(5):569-571.

(10) Schmidt LE, Dalhoff K. Alpha-fetoprotein is a predictor of outcome in acetaminophen-induced liver injury. Hepatology 2005 Jan;41(1):26-31.

(11) Schiodt FV, Rossaro L, Stravitz RT, Shakil AO, Chung RT, Lee WL, et al. Gc-globulin and prognosis in acute liver failure. Liver Transplantation 2005 Oct;11(10):1223-1227.

(12) Kremers WK, Van Ijperen M, Kim WR, Freeman RB, Harper AM, Kamath PS, et al. MELD Score as a Predictor of Pretransplant and Posttransplant Survival in OPTN/UNOS Status 1 Patients. Hepatology 2004 Mar;39(3):764-769.

(13) Yantorno SE, Kremers WK, Ruf AE, Trentadue JJ, Podestá LG, Villamil FG. MELD is superior to King's college and Clichy's criteria to assess prognosis in fulminant hepatic failure. Liver transplantation 2007;13(6):822.

(14) Cholongitas EB, Betrossian A, Leandro G, Shaw S, Patch D, Burroughs AK. King's criteria, APACHE II, and SOFA scores in acute liver failure. Hepatology 2006 author reply 882; Apr;43(4):881.

(15) Bailey B, Amre DK, Gaudreault P. Fulminant hepatic failure secondary to acetaminophen poisoning: a systematic review and meta-analysis of prognostic criteria determining the need for liver transplantation. Crit.Care Med. 2003 Jan;31(1):299-305.

For Peer Review

(16) Ding GK, Buckley NA. Evidence and consequences of spectrum bias in studies of criteria for liver transplant in paracetamol hepatotoxicity. QJM 2008 Sep;101(9):723-729.

(17) Altman DG. Systematic reviews of evaluations of prognostic variables. BMJ 2001 Jul 28;323(7306):224-228.

(18) Perel P, Edwards P, Wentz R, Roberts I. Systematic review of prognostic models in traumatic brain injury. BMC Medical Informatics & Decision Making 2006;6:38.

(19) Hayden JA, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Annals of internal medicine 2006;144(6):427.

(20) Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. J.Clin.Epidemiol. 2003 Nov;56(11):1129-1135.

(21) Dinnes J, Deeks J, Kirby J, Roderick P. A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy. Health Technol.Assess. 2005 iii; Mar;9(12):1-113.

(22) Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 6;327(7414):557-560.

(23) O'Grady JG, Wendon J, Tan KC, Potter D, Cottam S, Cohen AT, et al. Liver transplantation after paracetamol overdose. BMJ 1991 Jul 27;303(6796):221-223.

(24) Izumi S, Langley PG, Wendon J, Ellis AJ, Pernambuco RB, Hughes RD, et al. Coagulation factor V levels as a prognostic indicator in fulminant hepatic failure. Hepatology 1996 Jun;23(6):1507-1511.

(25) Anand AC, Nightingale P, Neuberger JM. Early indicators of prognosis in fulminant hepatic failure: an assessment of the King's criteria. Journal of Hepatology 1997;26(1):62.

(26) Bernal W, Wendon J, Rela M, Heaton N, Williams R. Use and outcome of liver transplantation in acetaminophen-induced acute liver failure. Hepatology 1998 Apr;27(4):1050-1055.

(27) Mitchell I, Bihari D, Chang R, Wendon J, Williams R. Earlier identification of patients at risk from acetaminophen-induced acute liver failure. Crit.Care Med. 1998 Feb;26(2):279-284.

(28) Bernal W, Wendon J. More on serum phosphate and prognosis of acute liver failure [4]. Hepatology 2003 01 Aug;38(2):533-534.

(29) Larson AM, Polson J, Fontana RJ, Davern TJ, Lalani E, Hynan LS, et al. Acetaminophen-induced acute liver failure: Results of a United States multicenter, prospective study. Hepatology 2005 Dec;42(6):1364-1372.

(30) Schmidt LE, Larsen FS. Prognostic implications of hyperlactatemia, multiple organ failure, and systemic inflammatory response syndrome in patients with acetaminophen-induced acute liver failure. Crit.Care Med. 2006 Feb;34(2):337-343.

(31) Zaman MB, Hoti E, Qasim A, Maguire D, McCormick PA, Hegarty JE, et al. MELD score as a prognostic model for listing acute liver failure patients for liver transplantation. Transplant.Proc. 2006 Sep;38(7):2097-2098.

For Peer Review

(32) Bates CM, Kochar N, Davidson JS, Hayes PC, Simpson KJ. Blood lactate as a predictor of poor prognosis in acute liver failure in acetaminophen overdose. Hepatology 2007;46(4):615A.

(33) Bernal W, Auzinger G, Sizer E, Wendon J. Early prediction of outcome of acute liver failure using bedside measurement of interleukin-6. Hepatology 2007;46(4 Supplement 1):617A.

(34) Trey C, Davidson CS. The management of fulminant hepatic failure. Prog.Liver Dis. 1970;3:282-298.

(35) Jones CM, Athanasiou T. Summary receiver operating characteristic curve analysis techniques in the evaluation of diagnostic tests. Ann.Thorac.Surg. 2005 Jan;79(1):16-20.

(36) Newsome PN, Henderson NC, Germain L, Ludlam CA, Simpson KJ. Prothrombin time to assess fulminant hepatic failure. Lancet 2001 Dec 22-29;358(9299):2172.

(37) O'Grady JG. Prognostication in acute liver failure: A tool or an anchor?. Liver Transplantation 2007 Jun;13(6):786-787.

(38) O'Grady JG. Broadening the view of acetaminophen hepatotoxicity. Hepatology 2005 Dec;42(6):1252-1254.

(39) Makin A, Williams R. Paracetamol hepatotoxicity and alcohol consumption in deliberate and accidental overdose. QJM 2000 Jun;93(6):341-349.

(40) Rolando N, Wade J, Davalos M, Wendon J, Philpott-Howard J, Williams R. The systemic inflammatory response syndrome in acute liver failure. Hepatology 2000 Oct;32(4 Pt 1):734-739.

(41) Cholongitas EB, Betrossian A, Leandro G, Shaw S, Patch D, Burroughs AK. King's criteria, APACHE II, and SOFA scores in acute liver failure. Hepatology 2006 author reply 882; Apr;43(4):881.

(42) Schmidt LE, Dalhoff K. Serum phosphate is an early predictor of outcome in severe a cetaminophen-induced hepatotoxicity. Hepatology 2002 Sep;36(3):659-665.

(43) Gow PJ, Sood S, Angus PW. Serum phosphate as a predictor of outcome in acetaminophen-induced fulminant hepatic failure. Hepatology 2003 author reply 712; Mar;37(3):711-712.

(44) Ng KL, Davidson JS, Bathgate AJ. Serum phosphate is not a reliable early predictor of outcome in paracetamol induced hepatotoxicity. Liver Transplantation 2004 Jan;10(1):158-159.

(45) Macquillan GC, Seyam MS, Nightingale P, Neuberger JM, Murphy N. Blood lactate but not serum phosphate levels can predict patient outcome in fulminant hepatic failure. Liver Transplantation 2005 Sep;11(9):1073-1079.

(46) Wiesner RH. MELD/PELD and the allocation of deceased donor livers for status 1 recipients with acute fulminant hepatic failure, primary nonfunction, hepatic artery thrombosis, and acute Wilson's disease. Liver Transplantation 2004 Oct;10(10 SUPPL. 2):S17-S22.

(47) Dhiman RK, Jain S, Maheshwari U, Bhalla A, Sharma N, Ahluwalia J, et al. Early indicators of prognosis in fulminant hepatic failure: an assessment of the Model for End-Stage Liver Disease (MELD) and King's College Hospital criteria. Liver Transplantation 2007 Jun;13(6):814-821.

For Peer Review

(48) Katoonizadeh A, Decaestecker J, Wilmer A, Aerts R, Verslype C, Vansteenbergen W, et al. MELD score to predict outcome in adult patients with non-acetaminophen-induced acute liver failure. Liver International 2007 Apr;27(3):329-334.

(49) Wei G, Bergquist A, Broome U, Lindgren S, Wallerstedt S, Almer S, et al. Acute liver failure in Sweden: Etiology and outcome. J.Intern.Med. 2007 Sep;262(3):393-401.

(50) Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. The Acute Respiratory Distress Syndrome Network. N.Engl.J.Med. 2000 May 4;342(18):1301-1308.

For Peer Review

PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4-5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5,6
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	Figure 1
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	6
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6-7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7-9

For Peer Review

Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7-9
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	12, Table 3
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	12, Table 3
Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	19
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	17
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	Figure 2
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Table 2
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Table 2
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Table 3
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Table 3
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Table 3
DISCUSSION			

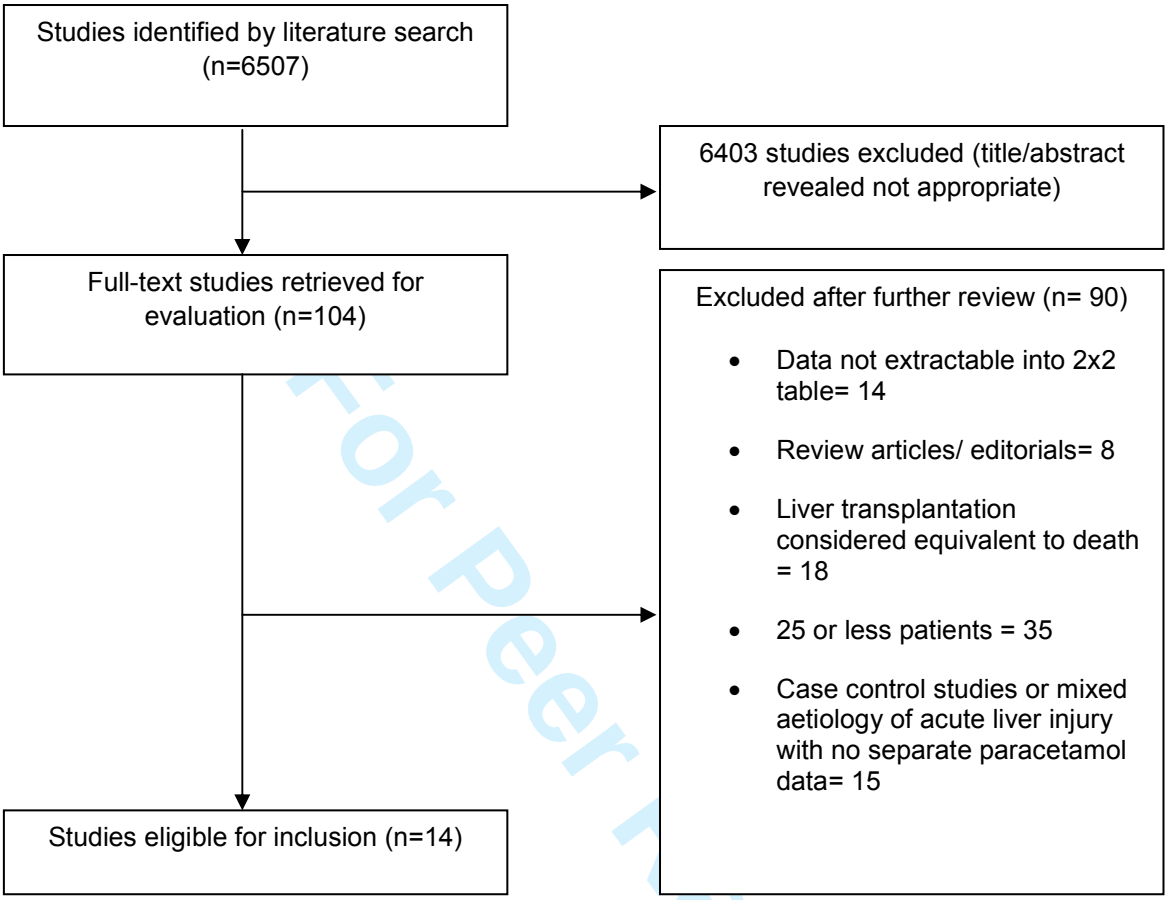
For Peer Review

Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	18
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	19
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	18-21
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	1

For Peer Review

- Adult patients (aged > 16 years) presenting with acute severe liver injury or ALF attributed to paracetamol hepatotoxicity
- Cohort studies
- Included >25 patients with paracetamol hepatotoxicity
- Mortality data recorded
- Data for death and liver transplantation considered separately

For Peer Review



For Peer Review

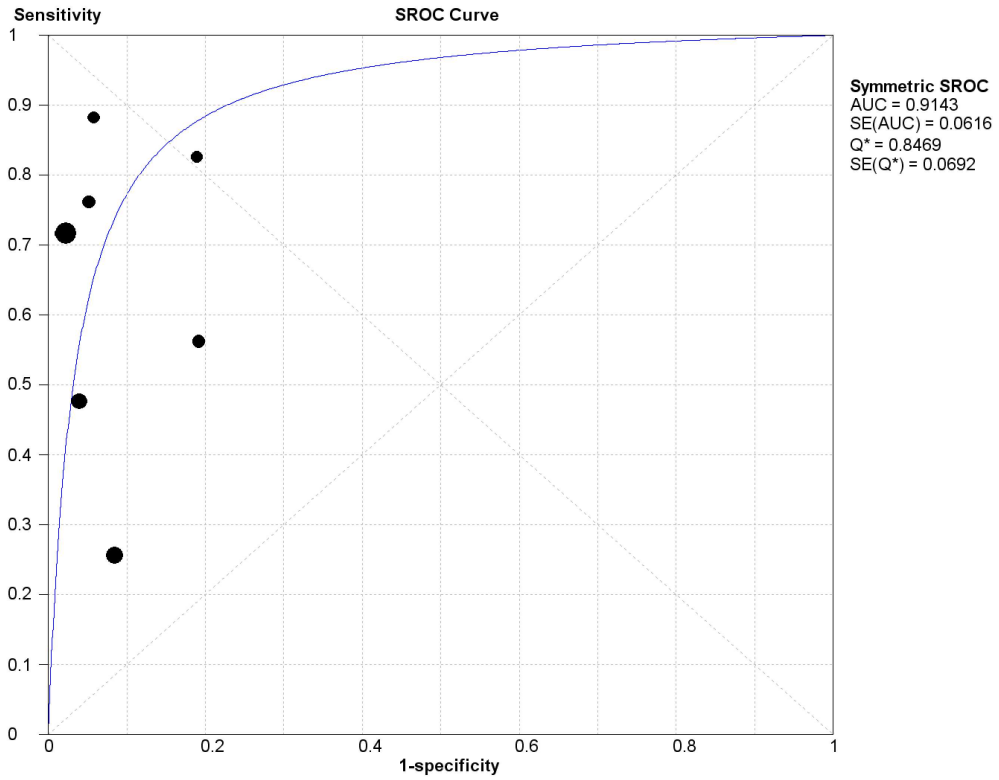


Figure 3 Summary receiver operator curve (SROC) of studies evaluating the original KCC
AUC: Area under the curve; SE: Standard Error; Q: heterogeneity

710x564mm (96 x 96 DPI)

View