

Different Methods of Creatinine Measurement Significantly Affect MELD Scores

Evangelos Cholongitas,^{1*} Laura Marelli,¹ Andrew Kerry,² Marco Senzolo,¹ David W. Goodier,² Devaki Nair,² Michael Thomas,² David Patch,¹ and Andrew K. Burroughs¹

¹Liver Transplantation and Hepatobiliary Medicine Unit and ²Department of Clinical Biochemistry, Royal Free Hospital, London, UK

Bilirubin (Bil) interferes with creatinine (Cr) measurement. Different laboratory methods are used to overcome this problem. Model for end-stage liver disease (MELD) scoring incorporates Cr and is used to prioritize patients for liver transplantation. Thus, MELD scores may vary with different Cr measurements influencing patients' priority. Our aim was to evaluate 4 different Cr assays (O'Leary modified Jaffe [mJCr], compensated [rate blanked] kinetic Jaffe [cJCr], enzymatic [ECr], and standard kinetic Jaffe [JCr]) in patients with abnormal liver function tests and assess changes in MELD score. A total of 403 consecutive samples from 158 patients' Cr assays were evaluated. Bland-Altman plots and MELD scores were also evaluated for each assay. Agreement was found to be poor among all Cr assays. Increased variability in Cr occurred with increasing Bil concentrations: Bil <100 $\mu\text{mol/L}$ \leq 3-point MELD variation - 3-point difference in 2%; Bil \geq 400 $\mu\text{mol/L}$ \leq 7-point MELD variation - \geq 3-point difference in 78%. When MELD was \geq 25 (mJCr as reference; mean, 30.5 points), MELD variation was greatest: mean, 28 (MELD cJCr), 27.5 (MELD ECr), and 28.4 (MELD JCr) ($P < 0.001$). In conclusion, there is poor agreement among different assays for Cr. As Bil concentration rises, there is greater variability in each creatinine measurements and thus greater variability in MELD scores that, this affect prioritization for liver transplantation. *Liver Transpl* 13:523-529, 2007.

© 2007 AASLD.

Received April 24, 2006; accepted September 9, 2006.

See Editorial on Page 486

The Model for end-stage liver disease (MELD) rates the severity of liver disease and is used to determine prioritization for liver transplantation (LT) in the United States.¹ The initial model was based on survival predictions for 231 cirrhotics who had undergone transjugular intrahepatic portosystemic shunt.² Subsequently, the MELD score was used to assess severity of liver disease and risk of dying within 3 months in patients awaiting LT.³

Several studies have confirmed the predictive ability of MELD. After the activation of the new allocation system in the United States, there were higher LT rates, fewer new registrations, and reduced mortality rate on the waiting list without an increasing death rate after LT.^{4,5} MELD has also been used prognostically in cirrhosis outside of transplantation.⁶

The advantage of the MELD score is that it uses objective, verifiable, and widely available laboratory tests: serum bilirubin (Bil), serum creatinine (Cr), and the international normalized ratio (INR) of prothrombin time.^{1,7,8} However, use of different laboratory methodologies for estimating INR and analysis of the same specimens for INR in different hospital laboratories⁹ gave significant disparity of results, such that the calculated MELD scores and priority for LT varied and thus may have some advantaged patients with less severe disease.^{9,10}

Chromogens, such as glucose, uric acid, antibiotics, keto acids and Bil, interfere with Cr measurement. The interference from Bil results in lower Cr values and is greater with increasing serum Bil concentrations, typically found in the sickest patients—namely, those with the greatest priority for LT. The negative interference of Bil is an unresolved problem.¹¹⁻¹³ Various modifica-

Abbreviations: MELD, model for end-stage liver disease; LT, liver transplantation; Bil, bilirubin; Cr, creatinine; INR, international normalized ratio; mJCr, O'Leary modified Jaffe; cJCr, compensated kinetic Jaffe; ECr, enzymatic; JCr, standard kinetic Jaffe. Address reprint requests to Andrew K. Burroughs, FRCP, Professor of Hepatology, Liver Transplantation and Hepatobiliary Medicine, Royal Free Hospital, Pond St., Hampstead, London NW3 2QG, UK. Telephone: 0044 20 74726229; FAX: 0044 20 74726226; E-mail: Andrew.Burroughs@royalfree.nhs.uk

DOI 10.1002/lt.20994

Published online in Wiley InterScience (www.interscience.wiley.com).

tions of the original Jaffe and enzymatic methods have been introduced to overcome this interference,^{11,14} which might lead to further discrepancy in Cr measurements.

The Jaffe reaction is most frequently used. Cr reacts with alkaline picrate, forming an orange-red complex that Janovski measured spectrophotometrically at 520 nm. Deproteinization of samples prior to analysis, increasing the picrate concentration, the use of Bil oxidase, or the addition of potassium ferricyanide, and kinetic alkaline picrate methods are modifications used to overcome Bil interference. Recently, Roche Diagnostics has introduced new calibrator for automated systems (cfas) calibrators to improve the correlation between the Jaffe method and high-performance liquid chromatography. A compensated value of Cr, created by adding 29 $\mu\text{mol/L}$ (0.33 mg/dL), is made independently of Bil concentration. Several enzymatic methods using creatininase and creatinine hydrolase convert creatinine to creatine. Dry chemistry systems such as Kodak Ektachem are not prone to interference due to a step that filtrates Bil.

The variety of methods reflects that in clinical practice no "gold standard," or even a reference standard, has been achieved for Cr measurement to compensate for Bil interference.

In the only previous study published as a full paper and related to LT candidates, 2 different Cr methods (Jaffe colorimetric and enzymatic) on 3 different analyzers were evaluated in 29 patients with a mean serum Bil concentration of 3.6 mg/dL (62 $\mu\text{mol/L}$); Cr values were not significantly affected. However, in a small study of 50 patients with liver disease, we showed that interference in Cr assays by different methods led to different MELD scores.¹⁵ Another preliminary report, published as an abstract, showed significant variation of MELD scores among 20 LT candidates when 2 different methods of Cr (enzymatic and nonenzymatic) were used.¹⁶

The aims of our study were to assess (1) the degree of agreement among 4 different Cr assays, (2) the potential changes in MELD scores using different methodologies for Cr measurement, (3) the cut-off points above which serum Bil significantly affects serum Cr and, therefore, MELD score estimation, and (4) the magnitude of changes in MELD scores (i.e., potential clinical impact) in patients with abnormal liver function tests.

PATIENTS AND METHODS

We assessed 403 consecutive blood samples obtained prospectively during routine clinical care from 158 outpatients or inpatients with abnormal liver function tests, seen at the Royal Free Hospital. None of the patients were receiving renal support measures. The samples were stored at -20°C prior to analysis. Serum creatinine concentration was determined on each sample using 4 different methods (Table 1). The first 3 used Roche Modular P unit (Roche Diagnostics, Ltd., Lewes, UK) calibrated using a lyophilized human-serum-based Cfas calibrator (Roche Diagnostics) standardized

TABLE 1. Types of Serum Creatinine Assay Used in the Study

Laboratory	Creatinine, mg/dL ($\mu\text{mol/L}$)
Method	O'Leary modified Jaffe
Instrument	Roche Modular P unit (Roche Diagnostics)
Reference range	Male: 0.68-1.36 (60-120) Female: 0.68-1.10 (60-97)
Method	Compensated (rate blanked) kinetic Jaffe
Instrument	Roche Modular P unit (Roche Diagnostics)
Reference range	Male: 0.70-1.20 (62-106) Female: 0.5-0.9 (44-80)
Method	Enzymatic creatinine
Instrument	Roche Modular P unit (Roche Diagnostics)
Reference range	Male: 0.5-1.1 (44-97) Female: 0.5-0.9 (44-80)
Method	Standard kinetic Jaffe
Instrument	Olympus AU2700 analyzer (Olympus Diagnostic Systems)
Reference range	Male <50 y: 0.84-1.24 (74-110), Male >50 y: 0.81-1.44 (72-127), Female: 0.66-1.09 (58-96)

against an isotope dilution mass spectrometry (ID-MS) method. The methods used were the following:

O'Leary modified Jaffe (mJCr): Potassium ferricyanide is used to oxidize Bil to biliverdin (pre-step). An increase in absorbance is measured at 505 nm and blanking at 570 nm. Bil is reported not to cause interference if $<400 \mu\text{mol/L}$.

Compensated (rate blanked) kinetic Jaffe (cJCr): This method measures increase in absorbance at 505 nm with blanking at 570 nm. No significant Bil interference is reported up to 171 $\mu\text{mol/L}$.

Enzymatic creatinine (ECr): This method uses a creatininase/creatinase/sarcosine oxidase system with detection at 546 nm and absorbance blanking at 700 nm. Interference from Bil is reported to be insignificant up to 85 $\mu\text{mol/L}$.

Standard kinetic Jaffe (JCr): This method is performed on an Olympus AU2700 analyzer (Olympus Diagnostic Systems, Southall, UK). The assay was calibrated using the manufacturer's recommended Olympus System calibrator, which is traceable to the National Institute of Standards and Technology Standard Reference Material (909b level 2). It measures an increase in absorbance at 520 nm and blanking at 800 nm. Interference is reported less than 10% up to 684 $\mu\text{mol/L}$ Bil.

These methods were developed to overcome the interference of Bil on Cr measurement, and clinical laboratories use them either solely in jaundiced patients or as a routine test. The 4 methods discussed are among the most widely used for estimating Cr in hospital laboratories in the United Kingdom and the United States: the

kinetic Jaffe (or one of its modifications) in 75% (10% the O'Leary modified Jaffe) and 73%, respectively, and the enzymatic method in 7% and 27%, respectively.¹⁷

Serum Bil was measured using a diazotized sulfanilic method, with caffeine accelerator on the Roche Modular P unit. INR was measured using the ACL Futura Coagulometer (Instrumentation Laboratory, Ltd., UK) and rabbit brain thromboplastin (PT-Fib HS Plus, Instrumentation Laboratory.), with an international sensitivity index of 1.12. MELD scores were derived separately using a Cr value from each method and calculated from the official United Network for Organ Sharing website (e.g., all patients with Cr <1 mg/dL were automatically awarded a creatinine of equal to 1 mg/dL).^{1,4}

The 4 groups were defined with interference to Bil: (1) Bil <100 $\mu\text{mol/L}$ (n = 164, 41%), (2) between 100 and 199 $\mu\text{mol/L}$ (n = 103, 25%), (3) between ≥ 200 and 399 $\mu\text{mol/L}$ (n = 81, 20%), or (4) ≥ 400 $\mu\text{mol/L}$ (n = 55, 14%).

Statistical analysis

All data were analyzed using the statistical package SPSS version 10.0 (SPSS Inc., Chicago, IL). Quantitative variables were expressed as mean values \pm 1 SD, and/or median values (range). Significance testing was 2-sided and set to less than 0.05. The Wilcoxon signed rank test was used for nonparametric evaluation between paired Cr values and paired MELD scores. The degree of agreement among the different Cr methods was evaluated using the Bland-Altman method, by plotting the difference between the 2 scores against their mean.¹⁸ This method is the correct statistical test, compared to any simple correlation test resulting in a correlation value, as such a test measures only the strength of a relation between 2 variables, with a test of the hypothesis of no relationship. Thus, if one evaluates different methods of measuring the same variable, in our case Cr, there will always be, a priori, a strong correlation, as all methods have a sound methodological basis, but they will not necessarily result in the same numerical values. As our aim was to assess the comparability between different laboratory methods of Cr measurement, a test assessing the degree of agreement between methods is necessary—the Bland-Altman plot. This test does not give a number or a percentage for the degree of agreement, so we evaluated the absolute number and percentage of values outside 1.96 SD. If the latter was $\geq 10\%$, agreement was considered to be poor. The correlation between different Cr values and different MELD scores with Bil were evaluated by the Spearman correlation.

RESULTS

PATIENT CHARACTERISTICS

The 158 patients had a median age of 52 years (range, 19-92), and 94 (58%) were men; The etiology of liver disease was alcoholic in 29 patients, viral hepatitis (B or C) in 43, cryptogenic/nonalcoholic steatohepatitis in

TABLE 2. Etiology of Liver Disease in Royal Free Hospital Cohort of Outpatients or Admitted to the Ward for Complications of Their Liver Disease

Cause of Liver Disease	Patients (n = 158) n (%)	Samples (n = 403) n (%)
Alcohol	29 (18.5)	129 (32)
Viral (hepatitis B or C)	43 (27)	69 (17.5)
Autoimmune	4 (2.5)	6 (1.5)
Cryptogenic/NASH	45 (28)	89 (22)
PBC/PSC	18 (12)	50 (12)
Others	19 (12)	60 (15)

Abbreviations: NASH, nonalcoholic steatohepatitis; PBC, primary biliary cirrhosis; PSC, primary sclerosing cholangitis.

45, primary biliary cirrhosis/ primary sclerosing cholangitis in 18, autoimmune hepatitis in 4, and other causes in 19 (Table 2). Sixty-five (41%) were cirrhotics, and the vast majority of them were on the list for LT. Also, 104 had 1 blood sample, and 54 had 2 or more.

Laboratory Measurements

The median values of Bil and INR (among 403 blood samples) were 129 (range, 4-913) $\mu\text{mol/L}$ and 1.5 (range, 0.9-8), respectively. The median mJCr, cJCr, ECr, and JCr values were 100 $\mu\text{mol/L}$ (range, 56-1,280), 73 $\mu\text{mol/L}$ (range, 33-1,339), 67 $\mu\text{mol/L}$ (range, 32-1,146), and 83 $\mu\text{mol/L}$ (range, 40-1,212), respectively ($P < 0.001$ for all comparisons). The median scores for MELD mJCr, MELD cJCr, MELD ECr, and MELD JCr were 21, 20, 19, and 20, respectively ($P < 0.001$ for all comparisons).

The mean values for mJCr, cJCr, ECr, and JCr according to Bil concentrations were Bil <100 $\mu\text{mol/L}$: 86 ± 21 , 77 ± 25 , 73 ± 25 , and 84 ± 22 , respectively; Bil ≥ 100 to <200 $\mu\text{mol/L}$: 112 ± 55 , 86 ± 68 , 83 ± 69 , and 96 ± 61 , respectively; Bil ≥ 200 to < 400 $\mu\text{mol/L}$: 127 ± 40 , 86 ± 45 , 79 ± 41 , and 99 ± 43 , respectively; and Bil ≥ 400 $\mu\text{mol/L}$: 228 ± 210 , 171 ± 233 , 146 ± 202 , and 176 ± 206 , respectively ($P < 0.001$ for all Bil group comparisons and for comparisons within each Bil group of the 4 Cr measurements (Table 3).

In addition, mJCr was significantly different compared to cJCr, ECr, and JCr in 392 (97%), 402 (99.9%), and 396 (98%) samples, respectively ($P < 0.001$); cJCr was significantly different compared to ECr and JCr in 387 (96%) and 398 (99%), respectively, ($P < 0.001$); and ECr was significantly different compared to JCr in 402 (99.9%) samples ($P < 0.001$). These differences remained significant in all 4 groups according to Bil concentrations, even when Bil was normal (≤ 17 $\mu\text{mol/L}$) (n = 48). In this subgroup, mJCr was significantly different compared to cJCr, ECr, and JCr in 48 (100%), 48 (100%), and 46 (96%), respectively ($P < 0.001$); cJCr was significantly different compared to ECr and JCr in 47 (98%) of samples ($P < 0.001$); and ECr was signifi-

TABLE 3. Mean Values of Serum Creatinine ($\mu\text{mol/L}$) as Measured by mJCr, cJCr, ECr, and JCr for Different Categories of Raised Bilirubin Concentrations

	mJCr	cJCr	ECr	JCr
Bilirubin mg/dL ($\mu\text{mol/L}$)				
<5.85 mg/dL (<100 $\mu\text{mol/L}$)	86 \pm 21	77 \pm 25	73 \pm 25	84 \pm 22
\geq 5.85 and <11.6 mg/dL (\geq 100 and <200 $\mu\text{mol/L}$)	112 \pm 55	86 \pm 68	83 \pm 69	96 \pm 61
\geq 11.6 and <23.4 mg/dL (\geq 200 and <400 $\mu\text{mol/L}$)	127 \pm 40	86 \pm 45	79 \pm 41	99 \pm 43
\geq 23.4 mg/dL (\geq 400 $\mu\text{mol/L}$)	228 \pm 210	171 \pm 233	146 \pm 202	176 \pm 206

The differences among mJCr, cJCr, ECr, and JCr were all significant ($P < 0.05$) for all comparisons in the 4 groups of bilirubin categories.

cantly different compared to JCr in 47 (98%) of samples ($P < 0.001$).

Correlation Between Bil and Cr, and Agreement among Cr Values Using Different Methods of Cr Measurement

There was reasonable correlation between Bil and mJCr ($r = 0.69$, $P < 0.001$). Other correlations were weak: Bil and cJCr or ECr ($r = 0.14$, $P = 0.004$) and Bil and JCr ($r = 0.26$, $P < 0.001$). Agreement according to Bland and Altman of the mJCr, which had the best correlation with Bil, and the other 3 creatinine methods was better for the mJCr and JCr, compared to mJCr and cJCr or ECr, but in each case more than 10% were outside 2 SD, and differences were also in opposite directions. The agreement between the other 3 methods was better, but approximately 8% were outside 2 SD, with differences in opposite directions (Fig. 1A and 1B).

Correlation Between Bil and MELD Scores Using Different Methods of Cr Measurement

The correlation between Bil and MELD scores was always good ($r > 0.80$, $P < 0.001$), but it was better between Bil and MELD mJCr ($r = 0.87$), followed by MELD cJCr ($r = 0.82$), MELD JCr ($r = 0.82$), and MELD ECr ($r = 0.81$).

MELD Scores Calculated Using Different Cr Laboratory Measurements

The mean scores for MELD mJCr, MELD cJCr, MELD ECr, and MELD JCr according to Bil concentrations were Bil <100 $\mu\text{mol/L}$: 13.3 \pm 4.7, 13 \pm 4.5, 13 \pm 4.4, and 13.2 \pm 4.5, respectively ($P = 0.14$); Bil \geq 100 to <200 $\mu\text{mol/L}$: 23 \pm 4.6, 22 \pm 4.8, 21.9 \pm 4.8, and 22.2 \pm 4.7 respectively ($P < 0.035$); Bil \geq 200 to <400 $\mu\text{mol/L}$: 26.7 \pm 5.4, 24.8 \pm 5.2, 24.5 \pm 5.4 and 25 \pm 5.4, respectively ($P < 0.001$); and Bil \geq 400 $\mu\text{mol/L}$: 31.3 \pm 5.3, 27.7 \pm 6, 27.1 \pm 6, and 28.3 \pm 6, respectively ($P < 0.001$). In the last group, 10 had a MELD score >40 (capped at 40, the upper value according to the United Network for Organ Sharing formula (Table 4).

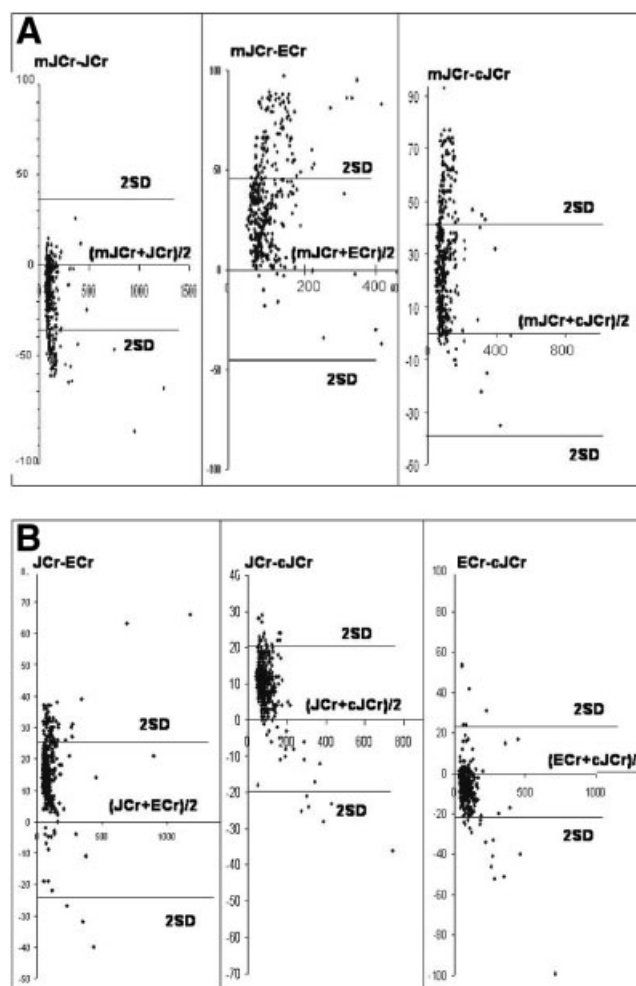


Figure 1. Degree of agreement between Cr values using different methods of Cr measurement (Bland-Altman plots). (A) . (B) .

The MELD mJCr scores were significantly different ($P < 0.05$) compared to MELD cJCr, MELD ECr, and MELD JCr when Bil ($\mu\text{mol/L}$) was \geq 33, \geq 36, and \geq 62, respectively. The MELD cJCr scores were significantly different compared to MELD ECr and MELD JCr when Bil was \geq 62 and \geq 38 $\mu\text{mol/L}$, respectively. Finally, the

TABLE 4. Mean Values of MELD Scores for Different Concentrations of Bilirubin and Severity of Liver Disease (mJCr is Reference Creatinine)*

	MELD mJCr	MELD cJCr	MELD ECr	MELD JCr
Bil mg/dL ($\mu\text{mol/L}$)				
<5.85 mg/dL (<100 $\mu\text{mol/L}$)	13.3 \pm 4.7	13 \pm 4.5	13 \pm 4.4	13.2 \pm 4.5
\geq 5.85 and <11.6 mg/dL (\geq 100 and <200 $\mu\text{mol/L}$)	23 \pm 4.6	22 \pm 4.8	21.9 \pm 4.8	22.2 \pm 4.7
\geq 11.6 and <23.4 mg/dL (\geq 200 and <400 $\mu\text{mol/L}$)	26.7 \pm 5.4	24.8 \pm 5.2	24.5 \pm 5.4	25 \pm 5.4
\geq 23.4 mg/dL (>400 $\mu\text{mol/L}$)	31.3 \pm 5.3	27.7 \pm 6	27.1 \pm 6	28.3 \pm 6
MELD score (using mJCr)				
0-15	10.2 \pm 1.2	10.2 \pm 1.2	10.2 \pm 1.2	10.2 \pm 1.2
15-19	17 \pm 3.1	16.5 \pm 2.8	16.5 \pm 2.7	16.7 \pm 2.7
20-24	22.2 \pm 2	21.2 \pm 1.4	21 \pm 1.5	21.5 \pm 1.3
\geq 25	30.5 \pm 4.5	28 \pm 4.9	27.5 \pm 5	28.4 \pm 4.9

*The differences among MELD cJCr, MELD mJCr, MELD ECr, and MELD JCr scores were all significant for all comparisons when Bil \geq 3.6 mg/dL (62 $\mu\text{mol/L}$) ($P < 0.05$).

MELD ECr scores were significantly different compared to MELD JCr when Bil was \geq 36 $\mu\text{mol/L}$.

The proportion of MELD mJCr scores, which were different compared to MELD cJCr, MELD ECr, and MELD JCr scores, progressively increased with higher Bil values. For Bil <100 $\mu\text{mol/L}$ ($n = 164$), only 5 (3%) had a difference \geq 2 points and 3 (2%) had a difference of 3 points (range of difference, 0-3 points). For Bil \geq 100 to <200 $\mu\text{mol/L}$ ($n = 103$), 17 (17%) had a difference \geq 2 points and 7 (7%) had a difference \geq 3 points (range of difference, 0-4 points). For Bil \geq 200 to <400 $\mu\text{mol/L}$ ($n = 81$), 46 (57%) had a difference \geq 2 points and 25 (31%) had a difference \geq 3 points (range of difference, 0-5 points). For Bil \geq 400 $\mu\text{mol/L}$ ($n = 55$), 48 (87%) had a difference \geq 2 points and 43 (78%) had a difference \geq 3 points (range of difference, 0-7 points). This trend was similar for the other Cr methods (cJCr, ECr, and JCr) when each was considered, in turn, as the reference method.

In terms of MELD scores, using MELD mJCr as the reference score, the mean scores of MELD using the other creatinine methods (MELD cJCr, MELD ECr, and MELD JCr) were as follows: 0-14 points (all means, 10.2 points) ($n = 93$); 15-19 points (mean, 17; means were 16.5, 16.5, and 16.7, respectively) ($n = 99$); 20-24 points (mean, 22.2; means were 21.2, 21, and 21.5, respectively) ($n = 82$); and \geq 25 points (mean, 30.5; means were 28, 27.5, and 28.4, respectively) ($n = 129$) (Table 4).

DISCUSSION

It has been recognized for many years that Bil is the major chromogen resulting in negative interference when measuring Cr, usually resulting in lower values.^{11,14} Several laboratory assays have been developed to overcome this problem, leading to different Cr measurements.^{11,13,14} MELD score is a liver-specific prognostic scoring system based on Bil, Cr, and INR.^{19,20} The MELD score was adopted to determine

prioritization for LT in the United States in February of 2002¹ replacing waiting time and Child-Turcotte-Pugh score, as this system had failed to prioritize according to severity of liver disease.^{19,21}

Although the 3 components of the MELD score are widely assumed to be reproducible, different laboratory methods for Cr measurement may result in widely variable values, particularly when Bil concentrations are high. Indeed, our study shows that significant differences in MELD scores occur, that are most divergent in the most jaundiced patients—in other words, those with the highest priority for LT. In addition, significantly different MELD scores were derived even with slightly abnormal values of Bil. For example, MELD mJCr scores were not significantly different compared to MELD cJCr only when serum Bil <33 $\mu\text{mol/L}$ and MELD ECr scores were similar to MELD JCr, only when Bil was <36 $\mu\text{mol/L}$. Differences in MELD scores of 1 or 2 points may not significantly alter prognosis, but they could affect ranking of patients. The clinical impact could be low, but it would need to be formally studied. A difference of 3 points or more is more likely to be clinically relevant for both prognosis and ranking.

In a recent preliminary report¹⁶ of 20 LT candidates, 2 different methods of creatinine measurement (ECr and an unidentified nonenzymatic method) were evaluated in 147 blood samples. When Bil was >10 mg/dL (170 $\mu\text{mol/L}$), MELD ECr scores were significantly higher than those of the unidentified nonenzymatic method. Similar to our study, there was also a progressive increase in the difference between mean MELD ECr and MELD nonenzymatic ECr scores for higher values of Bil. Bil had a stronger correlation with ECr, compared to nonenzymatic ECr; in other words, there was a less negative interference of Bil using the enzymatic method. However, in our study, the modified standard Jaffe method, which is used by several laboratories (including our own center) to overcome the interference of Bil, gave Cr values with the best correlation with Bil. The other 2 nonenzymatic methods (compensated ki-

netic Jaffe and standard Jaffe) as well as the enzymatic method, gave Cr values with weak correlation with Bil. Nevertheless, all 4 methods had poor or very poor agreement, resulting in significant differences in MELD scores.

In contrast to the aforementioned study¹⁶ and ours, another study¹⁰ evaluated 29 samples with Bil concentrations between 0.1 and 22.7 mg/dL (1.7-388.2 μ mol/L) and did not find a significant difference between 2 methods of creatinine measurement (enzymatic [Ortho Vitros 950] and Jaffe colorimetric on 2 different analyzers [Roche Hitachi 917 and Dade Dimension RXL]) when deriving the MELD score. This study¹⁰ concluded that significant differences in the MELD score between laboratories were primarily attributable to INR differences and not Bil or Cr values. However, the patients evaluated had a relatively low mean value of Bil (3.6 mg/dL, or 62 μ mol/L), which could explain the authors' results. Indeed, in our study, we found that the 4 MELD scores were all significantly different from each other only when Bil was \geq 62 μ mol/L. In our cohort, the mean MELD mJCr score in the subgroup with Bil \geq 62 μ mol/L was 24.8, similar to the mean MELD score of current candidates at LT in the United States (mean, 25 points)^{1,22} and much higher than the range of mean MELD scores (13.6-17.1 points) among candidates in the study by Trotter et al.,¹⁰ most of whom probably might not be listed currently if a MELD score \geq 15 became a minimum listing criterion.²³ In our study, we found no differences among the mean values of the 4 MELD scores when the MELD mJCr was between 0 and 14 points. On the other hand, for MELD mJCr \geq 25 points, the mean difference was up to 3 points (mean MELD mJCr of 30.5 vs. mean MELD ECr of 27.5) and range was between 0 and 7 points.

Our findings may have important implications, since Cr is the most widely used index for evaluation of renal function in cirrhotic patients and is often accepted as a useful surrogate marker of kidney reserve by hepatologists,²⁴ as well being used in MELD scoring for prioritization for LT.⁵ The aim of our study was not to evaluate the predictive accuracy of MELD score, but to show that individual patients may receive significantly different MELD scores and, thus, priority for LT using different methods for Cr measurement. We found that these differences in MELD scores ranged from 3 points (in the group with Bil <100 μ mol/L) to 7 points (in the group with Bil \geq 400 μ mol/L); the proportion of \geq 2 and \geq 3 points difference was greater in higher Bil concentrations (Table 5). The clinical significance of these findings is relevant. At Royal Free Hospital, for example, the proportion of candidates with Bil \geq 100 μ mol/L and \geq 200 μ mol/L on the day of LT was 20% and 15%, respectively (none on dialysis). In other centers, these proportions could be lower, or indeed higher.

In addition, the numerical value of Cr is also influenced by gender, age, ethnicity, and muscle mass,²⁵⁻²⁷ compounding the variations in measurement, which we have shown. As a result, patients with a relatively lower muscle mass, such as women or elderly people, may have Cr values that significantly underestimate their

renal dysfunction.²⁴ As the MELD allocation is a justice system, it is important that there are no systematic or other biases that could influence prioritization of some patients vs. others. Our study shows that the patients with the highest priority and Bil values would be those most affected by different assays for Cr. These variations will also affect the use of MELD to assess prognosis in liver disease outside of LT.⁶ Standardization of Cr measurement in jaundiced patients would eliminate some of these problems.

In conclusion, we found that utilization of different laboratory assays for Cr gives rise to significantly different MELD scores that would lead to inequalities in prioritization of candidates, especially those with the highest priority who are more jaundiced and have higher MELD scores. MELD differences may have little clinical impact particularly if differences are only 1-2 points, but this needs to be evaluated formally. Differences of 3 or 4 points, which are not uncommon with Bil concentrations between 100 μ mol/L (5.85 mg/dL) and 200 μ mol/L (11.6 mg/dL), are likely to be clinically relevant. Thus, although the MELD score is an evidence-based system for prioritizing candidates for LT, the results of our study show that the particular creatinine assay used in MELD scoring should be taken into account, as some patients inadvertently will be discriminated against with respect to others, in terms of priority on an LT waiting list. In addition to this lack of assay standardization for Cr, other factors affecting Cr concentrations, such as gender²⁸ and ethnicity, require further study to render the MELD system closer to an unbiased justice system for prioritizing LT candidates.²⁹

REFERENCES

1. Wiesner RH. Patient selection in an era of donor liver shortage: current US policy. *Nat Clin Pract Gastroenterol Hepatol* 2005;2:24-30.
2. Malinchoc M, Kamath PS, Gordon FD, Peine CJ, Rank J, ter Borg PC. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. *Hepatology* 2000;31:864-871.
3. Kamath PS, Wiesner RH, Malinchoc M, Kremers W, Therneau TM, Kosberg CL, et al. A model to predict survival in patients with end-stage liver disease. *Hepatology* 2001;33:464-470.
4. Wiesner R, Edwards E, Freeman R, Harper A, Kim R, Kamath P, et al. Model for end-stage liver disease (MELD) and allocation of donor livers. *Gastroenterology* 2003;124:91-96.
5. Freeman RB, Wiesner RH, Edwards E, Harper A, Merion R, Wolfe R. Results of the first year of the new liver allocation plan. *Liver Transpl* 2004;10:7-15.
6. Cholongitas E, Papatheodoridis GV, Vangeli M, Terreni N, Patch D, Burroughs AK. Systematic review: the model for end-stage liver disease—should it replace Child-Pugh's classification for assessing prognosis in cirrhosis? *Aliment Pharmacol Ther* 2005;22:1079-1089.
7. Freeman RB. MELD: the holy grail of organ allocation? *J Hepatol* 2005;42:16-20.
8. Said A, Williams J, Holden J, Remington P, Gangnon R, Musat A, Lucey MR. Model for end stage liver disease score predicts mortality across a broad spectrum of liver disease. *J Hepatol* 2004;40:897-903.

9. Arjal R, Mcashland T, Smith D, Brimhall B, Trotter J. Variation in model for end liver disease (MELD) score between transplant centres based on laboratory method for determination of international normalized ratio (INR). *Hepatology* 2005;42:205A.
10. Trotter JF, Brimhall B, Arjal R, Phillips C. Specific laboratory methodologies achieve higher model for endstage liver disease (MELD) scores for patients listed for liver transplantation. *Liver Transpl* 2004;10:995-1000.
11. Badiou S, Dupuy AM, Descomps B, Cristolead JP. Comparison between the enzymatic vitros assay for creatinine determination and three other methods adapted on the Olympus analyzer. *J Clin Lab Anal* 2003;17:235-240.
12. Boot S, LaRoche N, Legg EF. Elimination of bilirubin interference in creatinine assays by routine techniques: comparisons with a high performance liquid chromatography method. *Ann Clin Biochem* 1994;31(Pt 3):262-266.
13. Spencer K. Analytical reviews in clinical biochemistry: the estimation of creatinine. *Ann Clin Biochem* 1986;23(Pt 1):1-25.
14. Lolekha PH, Jaruthunyaluck S, Srisawasdi P. Deproteinization of serum: another best approach to eliminate all forms of bilirubin interference on serum creatinine by the kinetic Jaffe reaction. *J Clin Lab Anal* 2001;15:116-121.
15. Kerry A, Cholongitas E, Marelli L, Goodier D, Nair D, Thomas M, Burroughs A. MELD score and creatinine measurements. *Clin Chim Acta* 2005;355:A153.
16. Bzowej N, Huilgol V, Seban H, Bonacini M, Rabkin J, Doo E, et al. MELD underestimated unless enzymatic creatinine used. *J Hepatol* 2004;40(Suppl 2):40A.
17. Miller WG, Myers GL, Ashwood ER, Killeen AA, Wang E, Thienpont LM, Siekmann L. Creatinine measurement: state of the art in accuracy and interlaboratory harmonization. *Arch Pathol Lab Med* 2005;129:297-304.
18. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-310.
19. Durand F, Valla D. Assessment of the prognosis of cirrhosis: Child-Pugh versus MELD. *J Hepatol* 2005;42(Suppl): S100-S107.
20. Wang VS, Saab S. Liver transplantation in the era of model for end-stage liver disease. *Liver Int* 2004;24:1-8.
21. Christensen E. Prognostic models including the Child-Pugh, MELD and Mayo risk scores—where are we and where should we go? *J Hepatol* 2004;41:344-350.
22. Freeman R. Trends in MELD score and transplantation rates under the MELD system [abstract]. *Hepatology* 2005;42:203A.
23. Merion RM, Schaubel DE, Dykstra DM, Freeman RB, Port FK, Wolfe RA. The survival benefit of liver transplantation. *Am J Transplant* 2005;5:307-313.
24. Perrone RD, Madias NE, Levey AS. Serum creatinine as an index of renal function: new insights into old concepts. *Clin Chem* 1992;38:1933-1953.
25. Mpio I, Laville M, Hadj-Aissa A, Fauvel JP. Predicted creatinine clearance to evaluate glomerular filtration rate in black Caribbean subjects. *Nephrol Dial Transplant* 2003;18:1307-1310.
26. Swedko PJ, Clark HD, Paramsothy K, Akbari A. Serum creatinine is an inadequate screening test for renal failure in elderly patients. *Arch Intern Med* 2003;163:356-360.
27. Davies KM, Heaney RP, Rafferty K. Decline in muscle mass with age in women: a longitudinal study using an indirect measure. *Metabolism* 2002;51:935-939.
28. Cholongitas E, Marelli L, Kerry K, Goodier D, Nair D, Thomas M, et al. Female liver transplant recipients with the same GFR as male recipients have lower MELD scores—a systematic bias. *AM J Transpl* (in press).
29. Cholongitas E, Senzolo M, Triantos C, Samonakis D, Patch D, Burroughs AK. MELD is not enough—enough of MELD? *J Hepatol* 2005;42:475-477.