

# Desempenho de Seis Modelos de Predição Prognóstica em Pacientes Críticos que Receberam Suporte Renal Extracorpóreo\*

## *Performance of Six Prognostic Scores in Critically ILL Patients Receiving Renal Replacement Therapy*

Elizabeth R. Maccariello<sup>1</sup>, Carla Valente<sup>2</sup>, Lina Nogueira<sup>3</sup>, Márcia Ismael<sup>3</sup>,  
Ricardo V. R. Valença<sup>2</sup>, José E. S. Machado<sup>2</sup>, Eduardo Rocha<sup>4</sup>, Márcio Soares<sup>5</sup>

### RESUMO

**JUSTIFICATIVA E OBJETIVOS:** Não existe consenso sobre qual modelo prognóstico deva ser utilizado em pacientes com disfunção renal aguda (DRA). O objetivo deste estudo foi avaliar o desempenho de seis escores de prognóstico em pacientes que necessitaram de suporte renal.

**MÉTODO:** Coorte prospectiva de pacientes internados nas unidades de terapia intensiva (UTI) de três hospitais terciários que necessitaram de suporte re-

nal por DRA durante 32 meses. Foram excluídos os pacientes crônicos em programa de diálise ou com < 24h de internação na UTI. Os dados das primeiras 24h de UTI foram utilizados no cálculo do SAPS II e do APACHE II, e os dados das primeiras 24h de suporte renal foram utilizados no cálculo dos escores LOD, ODIN, Liaño e Mehta. A discriminação foi avaliada através da área sobre a curva ROC (AUROC) e a calibração através do teste do *goodness-of-fit* de Hosmer-Lemeshow. A letalidade hospitalar foi o desfecho de interesse.

**RESULTADOS:** Quatrocentos e sessenta e sete pacientes foram incluídos e a letalidade hospitalar foi 75%. Os valores dos escores SAPS II, APACHE II e LOD foram  $48,5 \pm 11,2$ ,  $27,4 \pm 6,3$ , 7 (5-8) pontos, respectivamente. A calibração foi adequada para todos os escores, com exceção do Mehta ( $p = 0,001$ ). Entretanto, a discriminação foi ruim para todos os modelos, com AUROC variando entre 0,60 para o ODIN e 0,72 para o SAPS II e Mehta. Com exceção do Mehta, todos os modelos subestimaram a letalidade.

**CONCLUSÕES:** Todos os seis modelos estudados foram inadequados na predição prognóstica de pacientes graves com DRA e necessidade de suporte renal.

**Unitermos:** hemodiálise, índices de gravidade de doença, insuficiência renal aguda, mortalidade, prognóstico, UTI

1. Médica Nefrologista, Hospital Universitário Antônio Pedro, HUFF, Niterói; Laboratório Multidisciplinar de Pesquisa do Hospital Universitário Clementino Fraga Filho, UFRJ; Nephro Consultoria; Rede D'Or de Hospitais, Doutora em Nefrologia pela UNIFESP; Pós-Doutora pela *Harvard Medical School*

2. Médico Nefrologista, Nephro Consultoria; Rede D'Or de Hospitais.

3. Enfermeira, Nephro Consultoria, RJ.

4. Professor Adjunto de Nefrologia, Faculdade de Medicina e Laboratório Multidisciplinar de Pesquisa do Hospital Universitário Clementino Fraga Filho, UFRJ; Nephro Consultoria, RJ; Rede D'Or de Hospitais, Pós-Doutorado em Nefrologia pela UFRJ.

5. Médico do Centro de Tratamento Intensivo, Instituto Nacional de Câncer; Doutor em Clínica Médica pela UFRJ, Bolsista de Produtividade em Pesquisa do CNPq, Título de Especialista em Medicina Intensiva pela AMIB; Pesquisador Clínico da Nephro Consultoria, RJ.

\*Recebido dos Hospitais Barra D'Or, Quinta D'Or, Copa D'Or, Rio de Janeiro, RJ

Apresentado em 03 de abril de 2008

Aceito para publicação em 06 de maio de 2008

Endereço para correspondência:

Dra. Elizabeth R. Maccariello  
Rua Jardim Botânico, 674 – Sala 405  
22461-000 Rio de Janeiro, RJ

Fone: +55 21 2249-9131; Fax: +55 21 3205-7034

E-mail: emaccariello@yahoo.com.br, emaccariello@hucff.ufrj.br

©Associação de Medicina Intensiva Brasileira, 2008

### SUMMARY

**BACKGROUND AND OBJECTIVES:** There is no consensus about prognostic scores for use in patients with acute kidney injury (AKI). The aim of this study was to evaluate the performance of six prognostic scores in predicting hospital mortality in patients with AKI and need for renal replacement therapy (RRT).

**METHODS:** Prospective cohort of patients admitted to the intensive care units (ICU) of three tertiary care hospitals that required RRT for AKI over a 32-month period. Patients with end-stage renal disease and those with ICU stay < 24h were excluded. Data from the first 24h of ICU admission were used to calculate SAPS II and APACHE II scores, and data from the first 24h of RRT were used in the calculation of LOD, ODIN, Liaño and Mehta scores. Discrimination was evaluated using the area under ROC curve (AUROC) and calibration using the Hosmer-Lemeshow goodness-of-fit test. The hospital mortality was the end-point of interest.

**RESULTS:** 467 patients were evaluated. Hospital mortality rate was 75%. Mean SAPS II and APACHE II scores were  $48.5 \pm 11.2$  and  $27.4 \pm 6.3$  points, and median LOD score was 7 (5-8) points. Except for Mehta score ( $p = 0.001$ ), calibration was appropriate in all models. However, discrimination was uniformly unsatisfactory; AUROC ranged from 0.60 for ODIN to 0.72 for SAPS II and Mehta scores. In addition, except for Mehta, all models tended to underestimate hospital mortality.

**CONCLUSIONS:** Organ dysfunction, general and renal-specific severity-of-illness scores were inaccurate in predicting outcome in ICU patients in need for RRT.

**Key Words:** acute kidney injury, dialysis, ICU, mortality, prognosis, severity-of-illness scores

## INTRODUÇÃO

A avaliação prognóstica é rotina na prática médica. A seleção dos pacientes, o tipo e a intensidade dos tratamentos, além de outras decisões relacionadas ao cuidado dos pacientes são influenciados pelo seu prognóstico<sup>1</sup>. Os escores prognósticos são compostos por variáveis clínicas e laboratoriais relevantes dos pacientes, que estão associados com o seu desfecho clínico. Embora estes modelos não devam ser utilizados para a predição individual, eles podem ser úteis nas discussões clínicas sobre prognóstico, na avaliação da qualidade de uma UTI e, principalmente, para a estratificação de pacientes em estudos clínicos<sup>1,2</sup>. Entretanto, os escores devem ser validados antes da sua utilização rotineira em uma população de pacientes.

Diversos modelos probabilísticos têm sido propostos para a avaliação do prognóstico de pacientes com disfunção renal aguda (DRA)<sup>3-8</sup>. No entanto, não existe consenso sobre qual modelo seja o mais apropriado para estes pacientes<sup>9-12</sup>.

O objetivo deste estudo foi avaliar o desempenho de seis modelos probabilísticos (dois escores gerais<sup>7,8</sup>,

dois escores específicos para pacientes com disfunção renal aguda<sup>3-6</sup> e dois escores de disfunções orgânicas agudas<sup>13,14</sup>), na predição da letalidade hospitalar de grande coorte de pacientes internados em unidades de terapia intensiva (UTI) com DRA e necessidade de suporte renal.

## MÉTODOS

Estudo prospectivo de coorte realizado entre dezembro de 2004 e julho de 2007 nas UTI clínicas e cirúrgicas de três hospitais terciários privados no Rio de Janeiro, RJ. O Hospital Barra D'Or possui 180 leitos e quatro UTI com um total de 42 leitos de terapia intensiva. O Hospital Quinta D'Or possui 200 leitos, sendo 60 de terapia intensiva distribuídos em três unidades. O Hospital Copa D'Or possui 200 leitos, com um total de 42 leitos de terapia intensiva distribuídos em cinco UTI. As UTI individuais são compostas por seis a 14 leitos. Médicos, enfermeiros, fisioterapeutas, nutricionistas e demais profissionais com experiência no cuidado de pacientes graves trabalham em todas as unidades. Pelo menos um intensivista, duas enfermeiras e cinco técnicos de enfermagem trabalham em regime de plantão (12h) nas unidades. Visitas clínicas com toda a equipe multiprofissional são realizadas diariamente em cada UTI. As decisões relacionadas ao cuidado dos pacientes são tomadas em conjunto pela equipe das UTI e pelo(s) médico(s) responsável(is) pelo paciente. Este estudo foi aprovado pelos Comitês de Ética em Pesquisa das instituições participantes, sem necessidade da obtenção do Termo de Consentimento Livre e Esclarecido, pela sua natureza observacional.

### **Crterios de Elegibilidade, Coleta de Dados e Definições de Termos e Variáveis**

Durante o estudo, todos os pacientes adultos (idade  $\geq 18$  anos) com diagnóstico de DRA ou disfunção renal crônica agudizada (DRCA) que necessitaram de suporte renal foram estudados. Os pacientes terminais, com insuficiência renal crônica em programa de diálise, submetidos a suporte renal por outra razão que não a disfunção renal ou com tempo de internação na UTI < 24h foram excluídos. Para o diagnóstico de DRCA considerou-se uma taxa de filtração glomerular (TFG) < 60 mL/min por pelo menos três meses<sup>15</sup>. Quando a TFG prévia era desconhecida, a fórmula do MDRD (*modification of diet in renal disease*) foi utilizada para a sua estimativa<sup>16</sup>. No caso de re-internações, somente a primeira foi considerada nas análises.

A DRA foi classificada de acordo com os critérios do RIFLE no momento do início do suporte renal<sup>17</sup>. Oligúria foi definida como um débito urinário < 400 mL/dia. As decisões para indicação, interrupção e método do suporte renal foram tomadas em conjunto pelo nefrologista e pelos médicos responsáveis pelo paciente. Os métodos de suporte renal utilizados foram a diálise convencional diária, a diálise estendida ou prolongada e os métodos contínuos de suporte renal. Os métodos foram indicados levando em consideração o quadro clínico do paciente e a sua hemodinâmica. Os métodos contínuos foram utilizados nos pacientes em uso de fármacos vasoativos ou com potencial de instabilidade hemodinâmica<sup>18</sup>.

No momento da internação e durante a permanência na UTI, dados demográficos, clínicos e laboratoriais foram coletados. Os seguintes escores de prognóstico foram calculados: as segundas versões do *Acute Physiology and Chronic Health Evaluation* (APACHE II)<sup>7</sup> e do *Simplified Acute Physiology Score* (SAPS II)<sup>8</sup>, o *Logistic Organ Dysfunction*<sup>14</sup>, o *Organ Dysfunction and Infection*<sup>14</sup>, e os escores de Liaño<sup>3</sup> e de Mehta<sup>6</sup>. Os escores APACHE II e SAPS II foram calculados no primeiro dia de internação na UTI e os escores de disfunção orgânica e os específicos para pacientes com DRA foram calculados no primeiro dia de suporte renal de acordo com as suas publicações originais. Para pacientes sedados, a escala de coma de Glasgow (ECG) prévia à sedação foi considerada. Quando uma variável não estava disponível, o valor zero ou valores normais foram atribuídos de acordo com as instruções para o cálculo de cada modelo prognóstico. No presente estudo, as variáveis demográficas e fisiológicas foram obtidas em todos os pacientes. Dentre as variáveis laboratoriais, valores normais foram atribuídos para o tempo de protrombina ativada em 102 (22%) e para a dosagem de bilirrubinas em 42 (9%) pacientes. As dosagens de bilirrubinas estavam disponíveis em todos os pacientes com icterícia ao exame clínico. As comorbidades foram diagnosticadas conforme os critérios de cada escore. Os pacientes foram classificados de acordo com o motivo da internação em clínicos, cirúrgicos eletivos e cirúrgicos de urgência.

Os critérios da Conferência de Consenso do *American College of Chest Physicians* e da *Society of Critical Care Medicine* foram utilizados para o diagnóstico de sepse<sup>19</sup>. Nas UTI estudadas, os critérios para o diagnóstico de infecção são os do *Centers for Diseases Control*<sup>20</sup>. A letalidade hospitalar foi o desfecho de interesse.

### Análise Estatística

O dados foram digitados em planilha eletrônica por uma única pessoa. A consistência dos dados foi observada por um autor e, ao final do estudo, foi realizada uma checagem em duplicata de amostra aleatória de 10% dos pacientes. Houve também uma verificação final de valores implausíveis e díspares. Estatísticas descritivas habituais foram utilizadas na caracterização da população. As variáveis contínuas foram apresentadas na forma de média  $\pm$  desvio-padrão ou de mediana (intervalo interquartil, 25%-75%) de acordo com a distribuição. A avaliação do desempenho dos escores foi realizada através das análises da discriminação e da calibração. A discriminação dos escores, ou seja, a capacidade do escore prognóstico de diferenciar os pacientes que sobreviveram daqueles que faleceram, foi avaliada através da área sob a curva recebedora das características dos observadores (*curva receiver operating characteristic*, AUROC)<sup>21</sup>. A análise da calibração (a concordância entre o número de pacientes observados e preditos a falecer no hospital através de todo espectro da gravidade) foi realizada pela estatística "C" do Qui-quadrado do *goodness-of-fit* (GOF) de Hosmer-Lemeshow<sup>22</sup>. Neste caso um valor elevado de  $p > 0,05$  foi considerado como critério para boa calibração. Também foi calculada a taxa de letalidade padronizada (TLP – óbitos observados / óbitos preditos) para cada modelo. Curvas de calibração foram construídas. Nelas, as taxas de letalidades observadas (eixo "y") e predita (eixo "x") foram confrontadas; os pacientes (colunas) foram distribuídos no eixo das abscissas ("x") em decis de letalidade predita. Um p-valor bicaudado < 0,05 foi considerado para significância estatística. Para as análises estatísticas foi utilizado o pacote SPSS versão 11.0 (SPSS Inc., Chicago, IL, USA).

### RESULTADOS

Durante o período do estudo, 467 pacientes preencheram os critérios de elegibilidade e foram incluídos. Dentre eles, 382 (82%) pacientes foram internados nas UTI por complicações clínicas, 49 (10%) em pós-operatório de cirurgias eletivas e 39 (8%), em pós-operatório de cirurgias de urgência. As principais características demográficas e clínicas estão apresentadas na tabela 1 e as principais razões para a internação na UTI são listadas na tabela 2. Os principais fatores associados com o desenvolvimento de DRA foram: sepse (n = 354, 82%), choque/hipoperfusão (n = 333, 71%), contraste/fármacos (n = 129, 28%), rabiomiólise (n = 22, 5%) e obstrução das vias urinárias (n = 15, 3%). A DRA foi multifatorial em 385 (82%) pacientes.

No dia do início do suporte renal, a média da concentração sérica de uréia foi 106 (62-164) mg/dL, a mediana da concentração sérica de creatinina foi 1,7 (1,1-2,6) mg/dL e a mediana do débito urinário foi 860 (406-1500) mL/24h. Cento e cinco (22%) pacientes apresentavam oligúria. O suporte renal foi iniciado no primeiro dia de UTI em 227 (49%) pacientes. Em função de elevada frequência de instabilidade hemodinâmica, os métodos contínuos de suporte renal foram utilizados em 83% dos pacientes; 14% dos pacientes foram tratados inicialmente com hemodiálise prolongada e 3% com hemodiálise convencional diária. As taxas de letalidade na UTI e hospitalar foram 71% e 75%, respectivamente.

Os resultados das análises de desempenho dos seis modelos são sumarizados na tabela 3. De modo geral, a discriminação foi ruim a regular para todos os modelos, sendo maior para o SAPS II [AUROC = 0,72 (intervalo de confiança de 95% (IC 95%) = 0,66 – 0,77)] e para o escore de Mehta [AUROC = 0,72 (IC 95% = 0,67 – 0,78)]. Com exceção do escore de Mehta, todos os demais modelos apresentaram uma boa calibração (GOF,  $p > 0,05$ ) e tenderam a subestimação da letalidade hospitalar (TLP > 1). As curvas de calibração para os modelos específicos para DRA, prognósticos gerais e de disfunção orgânica estão representadas nas figuras 1a, 1b e 1c, respectivamente.

Tabela 1 – Principais Características Demográficas e Clínicas (n = 467).

Variáveis	n (%), Média ± DP, Mediana (IIQ)
Idade (anos)	71,1 ± 15,3
Sexo masculino	241 (52%)
Internação hospitalar antes da UTI (dias - n)	0 (0 – 1)
Capacidade funcional comprometida (Knaus C ou D)	186 (40%)
Índice de comorbidade de Charlson (pontos)	3 (1 – 4)
Índice de comorbidade de Charlson ≥ 1 ponto	382 (82%)
SAPS II (à internação na UTI, pontos)	48,5 ± 11,2
APACHE II (à internação na UTI, pontos)	27,4 ± 6,3
LOD no D1 de suporte renal (pontos)	7 (5 – 8)
Classificação RIFLE	
Risco	132 (28%)
Lesão	111 (24%)
Falência	224 (48%)
Disfunção renal crônica	151 (32%)
Suporte renal no D1 de internação na UTI	227 (49%)
Durante a internação na UTI	
Sepse	354 (76%)
Ventilação mecânica	379 (81%)
Aminas vasoativas	352 (75%)
Duração da internação na UTI (dias)	18 (8 – 36)
Duração da internação hospitalar (dias)	23 (12 – 48)
Decisão para limitar o tratamento	83 (18%)
Letalidade na UTI	333 (71%)
Letalidade hospitalar	351 (75%)

DP = desvio-padrão, IIQ = intervalo interquartil; UTI = unidade de terapia intensiva; APACHE = *Acute Physiology and Chronic Health Evaluation*; SAPS = *Simplified Acute Physiology Score*; LOD = *Logistic Organ Dysfunction*.

Tabela 2 – Principais Causas de Internação na UTI (n = 467).

Variáveis	n (%)
Pacientes de cirurgias eletivas	46 (10%)
Gastrointestinal	27
Cardíaca	7
Outras	12
Pacientes de cirurgias de urgência	39 (8%)
Perfuração ou obstrução gastrointestinal	11
Complicações de cirurgias prévias	10
Cardíaca / Vasculares	9
Outras	9
Pacientes clínicos	382 (82%)
Sepse	238
Complicações cardiovasculares	38
Disfunção renal aguda	23
Complicações neurológicas	15
Hemorragia digestiva	15
Trauma	10
Outras / miscelânea	43

DESEMPENHO DE SEIS MODELOS DE PREDIÇÃO PROGNÓSTICA EM PACIENTES CRÍTICOS QUE RECEBERAM SUPORTE RENAL EXTRACORPÓREO

Tabela 3 – Área sobre a Curva ROC, Estatística “C” do Teste do *Goodness-of-fit* de Hosmer-Lemeshow e Taxas de Letalidades Padronizadas para os Escores SAPS II, APACHE II, LODS, ODIN, Liaño e Mehta (n = 467; Letalidade Hospitalar = 75,2%)

Escore de Prognóstico	Curva ROC	<i>Goodness-of-fit</i>		Letalidade Predita (Média ± DP)	TLP
	AUROC (IC 95%)	$\chi^2$	p-valor		
SAPS II	0,72 (0,66-0,77)	9,221	0,324	43,2±21,5	1,74 (1,55-1,95)
APACHE II	0,61 (0,55-0,67)	0,632	0,999	63,2±18,5	1,19 (1,09-1,30)
LODS	0,63 (0,58-0,69)	0,250	0,998	35,1±16,9	2,14 (1,87-2,45)
ODIN	0,60 (0,54-0,66)	4,176	0,383	53,1±13,5	1,41 (1,28-1,56)
Liaño	0,62 (0,56-0,67)	12,200	0,142	67,6±5,1	1,11 (1,02-1,21)
Mehta	0,72 (0,67-0,78)	26,260	0,001	78,4±16,5	0,96 (0,89-1,03)

AUROC = área sobre a curva ROC; IC = intervalo de confiança, DP = desvio-padrão; TLP = taxa de letalidade padronizada,  $\chi^2$  = Qui-quadrado, SAPS = *Simplified Acute Physiology Score*; APACHE = *Acute Physiology and Chronic Health Evaluation*; LOD = *Logisitic Organ Dysfunction*; ODIN = *Organ Dysfunction and Infection*.

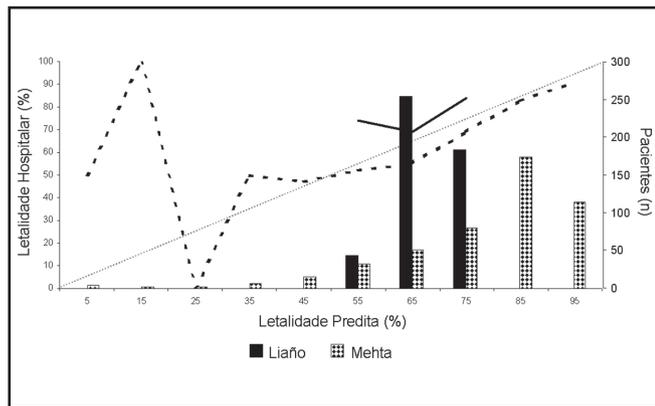


Figura 1a - Modelos Específicos para Pacientes com Disfunção Renal.

Curvas de calibração para os escores de Mehta (linha tracejada preta) e de Liaño (linha sólida preta). A linha tracejada diagonal representa a linha de predição ideal. As colunas representam o número de pacientes em cada decil de letalidade predita.

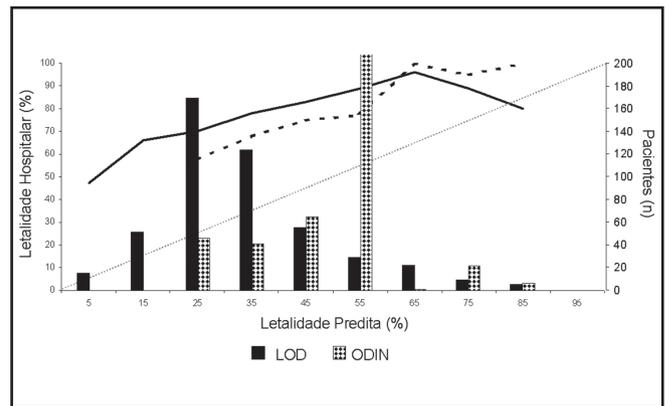


Figura 1c – Escores de Disfunção Orgânica.

Curvas de calibração para o ODIN (linha tracejada preta) e para o LOD (linha sólida preta). A linha tracejada diagonal representa a linha de predição ideal. As colunas representam o número de pacientes em cada decil de letalidade predita.

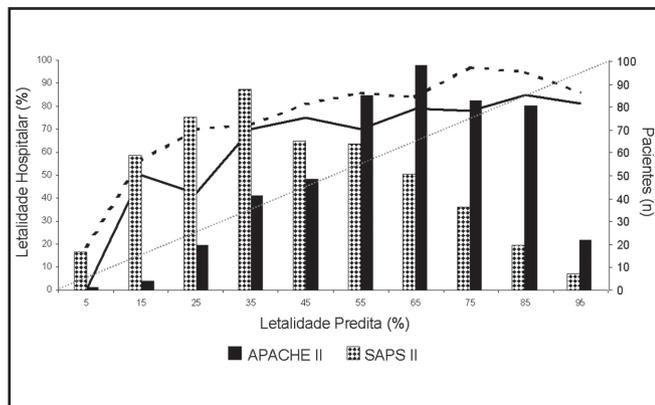


Figura 1b – Escores de Gravidade de Doença (modelos gerais). Curvas de calibração para o SAPS II (linha tracejada preta) e para o APACHE II (linha sólida preta). A linha tracejada diagonal representa a linha de predição ideal. As colunas representam o número de pacientes em cada decil de letalidade predita.

## DISCUSSÃO

Até a presente data, este é o maior estudo multicêntrico prospectivo realizado no Brasil sobre a avaliação do desempenho de modelos de predição prognóstica em pacientes com DRA que foram tratados com suporte renal durante a internação na UTI. Outro ponto de destaque para este estudo foi a avaliação simultânea de seis modelos nesta população, dentre os quais modelos prognósticos gerais (SAPS II e APACHE II), modelos de avaliação de disfunções orgânicas (LOD e ODIN) e de modelos específicos para pacientes com DRA (Mehta e Liaño). Todos os modelos estudados apresentaram desempenhos variados, mas de uma forma geral nenhum deles foi adequado para o uso nesta amostra de pacientes.

Ao longo das últimas décadas, vários escores de prognóstico foram desenvolvidos e têm sido propostos para a utilização em pacientes internados em UTI. Entretanto, estes modelos devem ser adequadamente validados antes da sua utilização rotineira em determinada população de pacientes. De modo geral, os modelos de prognóstico desenvolvidos a partir de uma população geral de pacientes de terapia intensiva têm desempenhos insatisfatórios quando testados em populações mais específicas<sup>23,24</sup>, incluindo os pacientes com DRA<sup>9-12,25-27</sup>. Entretanto, informações a respeito da aplicabilidade de tais modelos nos pacientes com DRA e necessidade de suporte renal ainda são limitadas.

O maior estudo sobre o tema foi realizado por Uchino e col., com a utilização da base de dados do Estudo BEST (*Beginning and Ending Supportive Therapy for the Kidney*)<sup>9</sup>. Neste estudo, os desempenhos de quatro modelos específicos para DRA (Mehta, Paganini, Chertow e Liaño), além do escore SAPS II e do SOFA, foram estudados em 1742 pacientes provenientes de 54 centros em 23 países. A discriminação foi uniformemente ruim para todos os modelos com AUROC variando de 0,610 para o escore de Chertow e 0,698 para o escore de Liaño. Com exceção do escore de Liaño ( $p = 0,36$ ), o ajuste da calibração foi inapropriado para os demais modelos ( $p < 0,001$ ).

No presente estudo, os escores APACHE II e SAPS II apresentaram discriminação regular e subestimaram a letalidade hospitalar. Estes resultados são semelhantes aos já descritos na literatura<sup>9-12,25-27</sup>. Uma das justificativas para estas observações reside no fato de é possível que tenha havido uma subrepresentação dos pacientes com DRA nos estudos que geraram os escores gerais como o APACHE II e SAPS II. Por conseguinte, modelos específicos para pacientes com DRA foram propostos<sup>3-6</sup>. Os modelos de Mehta e de Liaño foram escolhidos neste estudo por serem os modelos mais amplamente estudados e por gerarem probabilidades de óbito<sup>3,6</sup>. Os resultados das avaliações dos desempenhos destes escores foram muito variáveis. O escore de Mehta apresentou uma discriminação um pouco superior e, dentre todos os escores, apresentou a TLP [0,96 (IC 95%, 0,89-1,03)] mais próxima da unidade. Entretanto, a sua calibração foi ruim. Por outro lado, embora a calibração do escore de Liaño tenha sido adequa-

da, houve nítida subestimação da letalidade. Esta frustração com a utilização dos modelos específicos para DRA foram também evidenciadas em outros estudos<sup>6,9,26</sup>. Por fim dentre os modelos de disfunção orgânica, o escore SOFA é certamente o mais estudado<sup>6,9,25,28</sup>. Entretanto, o escore SOFA não gera probabilidades de óbito, somente uma contagem de pontos, razão pela qual optamos pelos escores LOD e ODIN<sup>13,14</sup>. À semelhança destes resultados, o LOD apresentou uma boa calibração em um estudo realizado em quatro centros americanos, mas a discriminação foi ruim<sup>6</sup>. No presente estudo foi realizada a primeira validação do ODIN em pacientes com DRA. Contudo, este escore não apresentou vantagens em relação ao LOD e aos demais modelos, com discriminação ruim e tendência a subestimação da letalidade.

O desempenho dos escores de prognóstico em pacientes com DRA em pacientes brasileiros foi avaliado em apenas três estudos distintos realizados em centro únicos<sup>2,26,27</sup>. Batista e col.<sup>27</sup> avaliaram o APACHE II e o modelo de Liaño em 76 pacientes de um único centro. A discriminação foi razoável para os dois modelos (AUROC de 0,76 e 0,78, respectivamente), mas a calibração foi boa apenas para o modelo específico. D'Avila e col. realizaram um estudo um pouco maior com 280 pacientes com DRA e necessidade de diálise<sup>26</sup>. Neste estudo também foram avaliados os escores APACHE II e Liaño. Embora a discriminação tenha sido melhor para o escore específico (AUROC = 0,81 *versus* AUROC = 0,65), o ajuste foi ruim para os dois modelos. Lima e col. Investigaram o desempenho dos escores APACHE II, SAPS II, LOD e Liaño em 324 pacientes com DRA<sup>12</sup>. De modo semelhante, todos os modelos se mostraram descalibrados e subestimaram a letalidade naquela população.

Na avaliação do desempenho de um modelo, são utilizadas basicamente duas medidas: a discriminação, que é avaliada pelo cálculo da área abaixo da curva ROC e a calibração, estimada pela estatística GOF de Hosmer e Lemeshow<sup>1,2,21,22</sup>. Entretanto, o número de pacientes avaliados em um estudo pode alterar os resultados. Desta forma, algumas considerações devem ser feitas em função do tamanho da amostra relativamente pequeno da maioria dos estudos. Em estudo muito elegante, Zhu e col. avaliaram a influência do tamanho da amostra na análise do desempenho do *Mortality Probability Models* (MPM) II por meio de simula-

ções por computador<sup>29</sup>. Eles demonstraram que quanto menor o tamanho da amostra, melhor a calibração do modelo. Por outro lado, a avaliação da discriminação do modelo não foi afetada.

Entretanto, o presente estudo também apresenta algumas limitações. Embora se trate de um estudo multicêntrico realizado em UTI com características diversas, todos os três hospitais são provenientes de mesma região geográfica e o suporte renal foi fornecido por equipes de nefrologistas com rotinas de cuidado semelhantes, incluindo aspectos relacionados ao procedimento de suporte renal. Desta forma, é necessário cautela com a extrapolação dos resultados aqui apresentados para outros serviços. A influência de efeitos regionais no desempenho dos escores de prognóstico foi bem demonstrada no estudo que gerou SAPS 3<sup>30</sup>. Neste estudo, além da equação padrão para o SAPS 3, foram desenvolvidas equações específicas para as diferentes regiões geográficas do mundo que participaram do estudo. Em estudo com pacientes com câncer internados em UTI brasileira, a equação específica para países das Américas do Sul e Central foi mais ajustada do que a equação geral do SAPS 3<sup>31</sup>. Além disso, também é necessário levar em consideração a ocorrência de possíveis vieses de seleção relacionados aos critérios de internação e alta a UTI. Diferenças nos critérios de admissão

e alta, diferenças nas doenças agudas e/ou crônicas dos pacientes e diferenças nos critérios para ordens de não reanimação e para a limitação de tratamento podem alterar o prognóstico. Por fim, a taxa de letalidade encontrada nestes pacientes foi relativamente elevada quando comparada com outros estudos internacionais<sup>6,10,32,33</sup>, mas ela foi semelhante às de outros resultados observados em populações brasileiras de pacientes com DRA<sup>12,26,27,33</sup>. Ademais, esta taxa de letalidade pode ser atribuída ao fato de ter sido estudada uma população mais idosa e certamente bastante grave em função das elevadas freqüências de comorbidades, sepse, comprometimento de capacidade funcional e necessidade de ventilação mecânica e aminas vasoativas.

Para concluir, todos os seis modelos estudados foram inadequados na predição prognóstica de pacientes gravemente enfermos com DRA e necessidade de suporte renal, incluindo os modelos específicos para pacientes renais. É necessário re-enfatizar que embora os escores possam ser úteis para enriquecer as discussões sobre prognóstico e para caracterizar a gravidade dos pacientes em estudos clínicos, nenhum modelo deve ser utilizado como parâmetro isolado para a indicação de condutas e tratamentos, incluindo a internação e alta da UTI, ou de início e término do suporte renal.

**APÊNDICE**

**Descrição dos escores de prognóstico específicos para DRA: \***

**Escore de Mehta (6):**

Log odds da letalidade = (0,0170 x idade (anos)) + (0,8605 x sexo masculino) + (0,0144 x nitrogênio uréico sérico) x (0,3398 x creatinina sérica) + (1,2242 x disfunção hematológica) + (1,1183 x disfunção hepática) + (0,9637 x insuficiência respiratória) + (0,0119 x frequência cardíaca) - (0,4432 x log (débito urinário)) - 0,7207

**Escore de Liaño (3):**

Probabilidade de óbito: (0,032 x idade (décadas)) - (0,086 x sexo masculino) + (0,109 x DRA nefrotóxica) + (0,109 x oligúria) + (0,116 x hipotensão) + (0,122 x icterícia) + (0,150 x coma) - (0,154 x redução da consciência) + (0,182 x ventilação mecânica) + 0,210

**Descrição dos escores de disfunção orgânica: \***

**Escore LOD (13)**

Medidas dos Sistemas Orgânicos	5	3	1	0	1	3	5
Neurológico (escala de coma de Glasgow)	3-5	6-8	9-13	14-15			
Cardiovascular							
FC (bpm)	< 30			30-139		≥ 140	
ou	ou			e		ou	
PAS (mmHg)	< 40	40-69	70-89	90-239	240-269	≥ 270	
Renal							
Uréia sérica (g/L)				< 0,36	0,36-0,59	0,60-1,19	≥ 120
ou							
Nitrogênio ureico (mg/dL)				< 17	17- < 28	28- < 56	≥ 56
e				e	ou	Ou	
Creatinina sérica (mg/dL)				< 1,20	120-1,59	≥ 1,60	
				e		ou	
Diurese (L/dia)	< 0,5	0,5-0,74		0,75-9,99		≥ 10	
Respiratório							
PaO <sub>2</sub> /FiO <sub>2</sub> em VM ou CPAP		< 150	≥ 150	sem ventilação; sem CPAP; sem IPAP			
Hematológico							
Nº de leucócitos (mm <sup>3</sup> ) x 10 <sup>3</sup>		< 1,0	1,0-2,4	2,5-49,9	≥ 50		
			ou	e			
Nº de plaquetas (mm <sup>3</sup> ) x 10 <sup>3</sup>			> 50	≥ 50			
Hepático				< 2,0	≥ 2,0		
Bilirrubina (mg/dL)				e	ou		
Tempo de protrombina (s e %)			(< 25%)	(< 3 > 25%)	≥ 3		

VM = ventilação mecânica; CPAP = pressão positiva contínua nas vias aéreas; IPAP = pressão intermitente positiva nas vias aéreas; FC = frequência cardíaca; PAS = pressão arterial sistólica

$$\text{Prob} = \frac{e^{\text{logit}}}{1 + e^{\text{logit}}} \quad \text{logit} = -3,4043 + 0,4173 * \text{Escore LOD}$$

$$\text{Probabilidade de óbito} = e^{\text{logit}} / (1 + e^{\text{logit}})$$

$$\text{Logit} = -3,4043 + 0,4173 * \text{Escore LOD}$$

**Escore ODIN (14)**

$$\text{Probabilidade de óbito} = e^{\text{logit}} / (1 + e^{\text{logit}})$$

$$\text{Logit} = -3,59 + (1,09 \times \text{respiratória}) + (1,19 \times \text{cardiovascular}) + (1,18 \times \text{renal}) + (0,86 \times \text{hematológica}) + (0,57 \times \text{hepática}) + (0,99 \times \text{neurológica}) + (0,53 \times \text{infecção})$$

Nº ODIN	0	1	2	3	4	5	6	7
Óbito	2,6%	9,7%	16,7%	32,3%	64,9%	75,9%	94,4%	100%

\*As definições utilizadas em cada escore podem ser obtidas nas suas publicações originais.

## REFERÊNCIAS

01. Rocco JR - Escores Prognósticos em Terapia Intensiva, em: David CM, Goldwasser RS, Nacul FE - Medicina Intensiva: Diagnóstico e Tratamento. Rio de Janeiro, Revinter, 1997;308-336.
02. Rocco JR - Escores Prognósticos em Medicina Intensiva, em: Nacul FE - Medicina Intensiva: Abordagem Prática. Rio de Janeiro, Revinter, 2004;555-562.
03. Liaño F, Gallego A, Pascual J, et al. Prognosis of acute tubular necrosis: an extended prospectively contrasted study. *Nephron*, 1993;63:21-31.
04. Paganini EP, Halstenberg WK, Goormastic M - Risk modeling in acute renal failure requiring dialysis: the introduction of a new model. *Clin Nephrol*, 1996;46:206-211.
05. Chertow GM, Lazarus JM, Paganini EP, et al. Predictors of mortality and the provision of dialysis in patients with acute tubular necrosis. The Auriculin Anaritide Acute Renal Failure Study Group. *J Am Soc Nephrol*, 1998;9:692-698.
06. Mehta RL, Pascual MT, Gruta CG, et al. Refining predictive models in critically ill patients with acute renal failure. *J Am Soc Nephrol*, 2002;13:1350-1357.
07. Knaus WA, Draper EA, Wagner DP, et al. APACHE II: a severity of disease classification system. *Crit Care Med*, 1985;13:818-829.
08. Le Gall JR, Lemeshow S, Saulnier F - A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA*, 1993; 270:2957-2963.
09. Uchino S, Kellum JA, Bellomo R, et al. Acute renal failure in critically ill patients: a multinational, multicenter study. *JAMA*, 2005;294:813-818.
10. Metnitz PG, Krenn CG, Steltzer H, et al. Effect of acute renal failure requiring renal replacement therapy on outcome in critically ill patients. *Crit Care Med*, 2002;30:2051-2058.
11. Halstenberg WK, Goormastic M, Paganini EP - Validity of four models for predicting outcome in critically ill acute renal failure patients. *Clin Nephrol*, 1997;47:81-86.
12. Lima EQ, Dirce MT, Castro I, et al. Mortality risk factors and validation of severity scoring systems in critically ill patients with acute renal failure. *Ren Fail*, 2005;27:547-556.
13. Le Gall JR, Klar J, Lemeshow S, et al. The Logistic Organ Dysfunction system. A new way to assess organ dysfunction in the intensive care unit. ICU Scoring Group. *JAMA*, 1996;276:802-810.
14. Fagon JY, Chastre J, Novara A, et al. Characterization of intensive care unit patients using a model based on the presence or absence of organ dysfunctions and/or infection: the ODIN model. *Intensive Care Med*, 1993;19:137-144.
15. National Kidney Foundation - K/DOQI Clinical Practice Guidelines for Chronic Kidney Disease; Evaluation, Classification and Stratification. *Am J Kidney Dis*, 2002;39:S46-S75.
16. National Kidney Foundation - K/DOQI Clinical Practice Guidelines for Chronic Kidney Disease; Evaluation of laboratory measurements for clinical assessment of kidney disease. *Am J Kidney Dis*, 2002; 39:S7-6-S110.
17. Bellomo R, Ronco C, Kellum JA, et al. Acute renal failure - definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. *Crit Care*, 2004;8:R204-R212
18. Maccariello E, Soares M, Valente C, et al. RIFLE classification in patients with acute kidney injury in need of renal replacement therapy. *Intensive Care Med*, 2007;33:597-605.
19. Bone RC, Balk RA, Cerra FB, et al. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine. *Chest*, 1992;101:1644-1655.
20. Garner JS, Jarvis WR, Emori TG, et al. CDC definitions for nosocomial infections, 1988. *Am J Infect Control*, 1998;16:128-140.
21. Hanley JA, McNeil BJ - The meaning and use of the area under receiver operating characteristic (ROC) curve. *Radiology*, 1982;143:29-36.
22. Hosmer DW, Lemeshow S - Applied Logistic Regression. 2<sup>nd</sup>, New York, Wiley-Interscience, 2000.
23. Soares M, Fontes F, Dantas J, et al. Performance of six severity-of-illness scores in cancer patients requiring admission to the intensive care unit: a prospective observational study. *Crit Care*, 2004;8:R194-R203.
24. Chatzicostas C, Roussomoustakaki M, Vlachonikolis IG, et al. Comparison of Ranson, APACHE II and APACHE III scoring systems in acute pancreatitis. *Pancreas*. 2002;25:331-335.
25. Ahlstrom A, Kuitunen A, Peltonen S, et al. Comparison of 2 acute renal failure severity scores to general scoring systems in the critically ill. *Am J Kidney Dis*, 2006;48:262-268.
26. d'Avila DO, Cendoroglo Neto M, dos Santos OF, et al. Acute renal failure needing dialysis in the intensive care unit and prognostic scores. *Ren Fail*, 2004;26:59-68.
27. Batista PB, Cendoroglo Neto M, dos Santos OF, et al. Evaluation of prognostic indexes in critical acute renal failure patients. *Ren Fail*, 2004;26:545-552.
28. de Mendonca A, Vincent JL, Suter PM, et al. Acute renal failure in the ICU: risk factors and outcome evaluated by the SOFA score. *Intensive Care Med*, 2000;26:915-921.
29. Zhu BP, Lemeshow S, Hosmer DW, et al. Factors affecting the performance of the models in the Mortality Probability Model II system and strategies of customization: a simulation study. *Crit Care Med*, 1996;24:57-63.
30. Moreno RP, Metnitz PG, Almeida E, et al. SAPS 3--From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. *Intensive Care Med*, 2005;31:1345-1355.
31. Soares M, Salluh JI - Validation of the SAPS 3 admission prognostic model in patients with cancer in need of intensive care. *Intensive Care Med*, 2006;32:1839-1844.
32. Bagshaw SM, Mortis G, Doig CJ, et al. One-year mortality in critically ill patients by severity of kidney dysfunction: a population-based assessment. *Am J Kidney Dis*, 2006;48:402-409.
33. Uchino S, Kellum JA, Bellomo R, et al. Acute renal failure in critically ill patients: a multinational, multicenter study. *JAMA*, 2005;294:813-818.