



Comparação entre dois escores de prognóstico (PRISM e PIM) em unidade de terapia intensiva pediátrica

Comparison of two prognostic scores (PRISM and PIM) at a pediatric intensive care unit

Vanessa Feller Martha¹, Pedro Celiny Ramos Garcia², Jefferson Pedro Piva³,
Paulo Roberto Einloft⁴, Francisco Bruno⁵, Viviane Rampon⁶

Resumo

Objetivo: Comparar o desempenho do PRISM (*Pediatric Risk of Mortality*) e o PIM (*Pediatric Index of Mortality*) em uma unidade de terapia intensiva pediátrica geral investigando a relação existente entre a mortalidade e a sobrevivência observadas com a mortalidade e sobrevivência estimadas pelos dois escores.

Métodos: Estudo de coorte contemporâneo realizado entre 1º de junho de 1999 a 31 de maio de 2000 na unidade de terapia intensiva pediátrica do Hospital São Lucas da Pontifícia Universidade Católica do Rio Grande do Sul. Os critérios de inclusão no estudo e o cálculo do PRISM e do PIM foram feitos conforme preconizado e utilizando as fórmulas de seus artigos originais. Para estatística, utilizou-se o teste z de Flora, teste de ajuste de Hosmer-Lemeshow, área sob a curva típica de um receptor operador de características (curva ROC) e teste de correlação de Spearman. O estudo foi aprovado pelo Comitê da instituição.

Resultados: Internaram na unidade de terapia intensiva pediátrica 498 pacientes, sendo 77 excluídos. Dos 421 pacientes estudados, 33 (7,83%) foram a óbito. A mortalidade estimada pelo PRISM foi de 30,84 (7,22%), com *standardized mortality rate* 1,07 (0,74-1,50), $z = -0,45$. Pelo PIM, foi de 26,13 (6,21%), com *standardized mortality rate* 1,26 (0,87-1,77), $z = -1,14$. O teste de ajuste de Hosmer-Lemeshow obteve um qui-quadrado 9,23 ($p = 0,100$) para o PRISM e 27,986 ($p < 0,001$) para o PIM. A área abaixo da curva ROC foi 0,870 (0,810-0,930) para o PRISM e 0,845 (0,769-0,920) para o PIM. Teste de Spearman $r = 0,65$ ($p < 0,001$).

Conclusão: Na análise dos testes podemos constatar que, embora o PIM apresente uma pior calibração no conjunto dos resultados, tanto o PRISM como o PIM apresentaram boa capacidade de discriminar entre sobreviventes e não sobreviventes, constituindo-se em ferramentas de desempenho comparáveis na avaliação prognóstica de pacientes pediátricos admitidos em nossa unidade.

J Pediatr (Rio J). 2005;81(3):259-64: Escore prognóstico, PRISM, PIM, mortalidade.

Abstract

Objective: To compare the performance of the PRISM (*Pediatric Risk of Mortality*) and the PIM (*Pediatric Index of Mortality*) scores at a general pediatric intensive care unit, investigating the relation between observed mortality and survival and predicted mortality and survival.

Methods: A contemporary cohort study undertaken between 1 June 1999 and 31 May 2000 at the *Pontifícia Universidade Católica do Rio Grande do Sul, Hospital São Lucas* pediatric intensive care unit. The inclusion criteria and the PRISM and PIM calculations were performed as set out in the original articles and using the formulae as published. Statistical analysis for model evaluation employed the Flora z test, Hosmer-Lemeshow goodness-of-fit test, ROC curve (receiver operating characteristic) and Spearman's correlation tests. The study was approved by the institution's Ethics Committee

Results: Four hundred and ninety-eight patients were admitted to the pediatric intensive care unit, 77 of whom presented exclusion criteria. Thirty-three (7.83%) of the 421 patients studied died and 388 patients were discharged. Estimated mortality by PRISM was 30.84 (7.22%) with a standardized mortality rate of 1.07 (0.74-1.50), $z = -0.45$ and by PIM this was 26.13 (6.21%) with a standardized mortality rate of 1.26 (0.87-1.77), $z = -1.14$. The Hosmer-Lemeshow test gave a chi-square of 9.23 ($p = 0.100$) for PRISM and 27.986 ($p < 0.001$) for PIM. The area under the ROC curve was 0.870 (0.810-0.930) for PRISM and 0.845 (0.769-0.920) for PIM. The Spearman test returned $r = 0.65$ ($p < 0.001$).

Conclusion: Analyzing the tests we can observe that, although the PIM test was less well calibrated overall, both PRISM and PIM offer a good capacity for discriminating between survivors and moribund patients. They are tools with comparable performance at the prognostic evaluation of the pediatric patients admitted to our unit.

J Pediatr (Rio J). 2005;81(3):259-64: Prognostic scores, PRISM, PIM, mortality.

1. Mestre em Pediatria. Médica, Unidade de Terapia Intensiva Pediátrica, Hospital Moinhos de Vento e Emergência do Hospital da Criança Santo Antônio, Porto Alegre, RS.
2. Doutor. Professor adjunto, Departamento de Pediatria e Curso de Pós-Graduação em Pediatria, Faculdade de Medicina, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS). Médico chefe, Serviço de Terapia Intensiva e Emergência Pediátrica, Hospital São Lucas, PUCRS, Porto Alegre, RS.
3. Doutor. Professor adjunto, Departamento de Pediatria e Curso de Pós-Graduação em Pediatria, Faculdade de Medicina, PUCRS. Médico chefe associado, Serviço de Terapia Intensiva e Emergência Pediátrica, Hospital São Lucas, PUCRS, Porto Alegre, RS.
4. Professor adjunto, Departamento de Pediatria. Médico chefe associado, Serviço de Terapia Intensiva e Emergência Pediátrica, Hospital São Lucas, PUCRS, Porto Alegre, RS.
5. Professor adjunto, Dep. de Pediatria. Médico, Serviço de Terapia Intensiva e Emergência Pediátrica, Hospital São Lucas, PUCRS, Porto Alegre, RS.
6. Médica plantonista, Serviço de Terapia Intensiva e Emergência Pediátrica, Hospital São Lucas, PUCRS, Porto Alegre, RS.

Artigo submetido em 10.05.04, aceito em 16.05.05.

Como citar este artigo: Martha VF, Garcia PC, Piva JP, Einloft PR, Bruno F, Rampon V. Comparação entre dois escores de prognóstico (PRISM e PIM) em unidade de terapia intensiva pediátrica. *J Pediatr (Rio J)*. 2005;81:259-64.

Introdução

As unidades de terapia intensiva pediátrica (UTIP) objetivam promover assistência qualificada, visando alcançar os melhores resultados e uma melhor evolução para as crianças criticamente enfermas. Essas unidades são locais de grande transferência tecnológica e uma das principais consumidoras do orçamento hospitalar. Entretanto, ao tratar pacientes com diferentes prognósticos e níveis de gravidade, o resultado final do uso dos recursos disponíveis nessas unidades é, muitas vezes, incerto. Nesse cenário, a incorporação de tecnologia nem sempre segue regras estritas de análise quanto a evidências científicas de suporte e, menos ainda, de custo-eficiência¹.

Uma forma de comparar a qualidade e efetividade de atendimento prestado em uma unidade é através da comparação com outras em situação semelhante². As UTIP comparam componentes que estão relacionados, como a severidade da doença e os recursos disponíveis em sua unidade, com o desfecho de determinados tipos de pacientes. Mortalidade e tempo de internação são exemplos de desfechos mais utilizados. Para medir a gravidade, utilizam escores de risco de mortalidade que estabelecem uma escala numérica e, dessa forma, comparam essa mortalidade estimada em porcentagem com a mortalidade realmente observada³. Esses chamados escores prognósticos podem ser utilizados para a avaliação de qualidade do atendimento médico e otimização dos recursos empregados, visando uma melhor relação custo-benefício. Por compararem a mortalidade ajustada à gravidade da doença, servem também para a classificação dos pacientes por severidade de doença, que pode ser utilizada para comparação entre estudos clínicos e planejamento de recursos tecnológicos nessa área⁴.

Os principais escores desenvolvidos para população pediátrica são o PRISM (*Pediatric Risk of Mortality*)⁵ e o PIM (*Pediatric Index of Mortality*)⁶ e suas novas versões, o PRISM III⁷ e o PIM-2⁸. Esses escores foram desenvolvidos a partir da identificação de variáveis relevantes para o risco de mortalidade e pontuados após uma posterior análise estatística multivariada de regressão logística⁹.

O PRISM, publicado em 1988 por Pollack *et al.*, apresentou uma excelente performance discriminatória e preditiva⁵. É ainda o índice mais amplamente conhecido e utilizado nas UTIP, sendo aplicado em estudos clínicos como o escore prognóstico padrão para avaliação de severidade de doença em pacientes pediátricos. Uma versão revisada do PRISM, o PRISM III, está disponível desde 1996⁶ e, segundo seus autores, fornece uma maior capacidade preditiva¹⁰. Entretanto, um pagamento considerável é cobrado para sua utilização rotineira, o que tem limitado seu uso mesmo em países desenvolvidos¹¹⁻¹³ e, por esse motivo, não foi avaliado neste estudo.

Os resultados do artigo original do PIM, publicado em 1997 por Shann *et al.*, evidenciaram que tal modelo tem uma boa capacidade de prever e classificar risco de mortalidade em grupos de crianças internadas em unidades de terapia intensiva⁷. Os autores apontam como uma vantagem do PIM em relação ao PRISM o fato de o PIM ser

baseado em apenas oito variáveis, todas coletadas no momento da internação, o que facilitaria a coleta dos dados e evitaria o efeito no resultado do escore após 24 horas de diferentes estratégias de manejo intensivo¹⁴. Diversos artigos que avaliaram o PIM mostraram que ele apresenta um bom desempenho para prever morte¹¹⁻¹⁶. Em 2003, o *PIM Study Group* publica uma versão revisada do PIM, o PIM-2^{16,17}, que, comparada à versão original¹⁸, seria mais bem calibrada, mais segura e com melhor ajuste em diferentes grupos diagnósticos. Essa nova versão ainda não foi avaliada de forma independente, carecendo de maiores informações quanto a sua performance em outras regiões e, por ter sido publicada após o início deste estudo, não está incluída neste trabalho.

Os desempenhos dos sistemas PRISM e PIM foram algumas vezes comparados pelos próprios autores que desenvolveram os escores^{15,18-20}, mas poucas vezes comparados de forma independente. Até o momento, estudos independentes não utilizaram grupos de pacientes heterogênicos próprios da UTIP, mas algumas categorias específicas de doença^{11,16,21}, novas versões do método¹⁰ ou grupos de pacientes homogêneos de alta mortalidade²². Nenhum estudo nesse sentido foi publicado na América Latina.

Neste estudo independente, nosso objetivo foi comparar o desempenho do PRISM e do PIM em uma UTIP geral investigando a relação existente entre a mortalidade e a sobrevivência observadas com a mortalidade e sobrevivência estimadas pelos dois escores.

Métodos

Estudo de coorte contemporâneo realizado entre 1º de junho de 1999 a 31 de maio de 2000 na UTIP do Hospital São Lucas da PUCRS. Os dados foram coletados prospectivamente para gerar os escores e predições para o período de tempo e com a técnica preconizada por cada um dos escores (PRISM primeiras 24 horas da internação e PIM até 1 hora após a internação)^{5,6}. Foram excluídos do estudo aqueles pacientes que evoluíram a óbito nas primeiras 8 horas ou receberam alta nas primeiras 24 horas de internação.

Estimou-se um mínimo de 253 pacientes para o presente estudo. O cálculo da amostra tomou por base uma população média de 500 pacientes internados na UTIP, considerando-se uma mortalidade de 15% e tolerando-se uma variação de mortalidade entre 10 e 20% e um limite de confiança de 99%.

O cálculo do PRISM e do PIM foi feito através da fórmula do escore disponível em seus artigos originais^{5,6}. Nenhum exame foi realizado para atender às necessidades da pesquisa, uma vez que esta considera exame não coletado como normal. Foram coletados, ainda, dados demográficos para caracterização da amostra, como idade na admissão, gênero e procedência. O desfecho avaliado foi o tempo de internação na unidade e a evolução do paciente (alta ou óbito).

Análise descritiva simples foi utilizada para os grupos e subgrupos estudados (média, mediana, desvio padrão). A

estatística z, conforme descrita por Flora²³, foi utilizada para comparar a semelhança geral na mortalidade observada e esperada pelo SMR (*standardized mortality rate* - índice padronizado de mortalidade). Para a calibração dos escores, utilizou-se o teste de ajuste de Hosmer-Lemeshow, que testou a concordância da mortalidade observada com a esperada em cinco diferentes intervalos de risco²⁴. A capacidade de discriminação entre sobreviventes e não-sobreviventes foi feita através da área sob a curva típica de um receptor operador de características (curva ROC)^{25,26}, e a correlação quantitativa entre os resultados dos escores foi analisada pelo teste de correlação de Spearman.

O estudo foi aprovado pelo Comitê de Ética em Pesquisa da Pontifícia Universidade Católica do Rio Grande do Sul e, por não trazer nenhum risco ao paciente, foi obtida dispensa do termo de consentimento pós-informado, sendo firmado compromisso em manter sigilo quanto à identidade dos pacientes.

Resultados

No período do estudo, internaram na UTIP 498 pacientes. Porém, 77 pacientes apresentaram critérios de exclusão, oito por evoluírem a óbito nas primeiras 8 horas de internação, e o restante por receber alta antes de completar 24 horas na UTIP. Nenhum paciente foi excluído por falta de dados. As características gerais da amostra estão na Tabela 1.

Tabela 1 - Características gerais da amostra no estudo

Características	Valor
Total da amostra	421 pacientes
Óbitos	33 (7,8%)
Idade (meses); média (mediana)	44,3 (26%)
Tempo Internação (dias); média (mediana)	11,6 (5,5%)
Peso (quilos); média (mediana)	14,9 (11,4%)
Cor branca	345 (82%)
Gênero masculino	212(50%)
Admissões	
Emergência	107 (25%)
Hospital	314 (75%)
Tipo de internação	
Clínica	234 (56%)
Cirúrgica	187 (44%)
Diagnósticos por disfunção orgânica	
Cardiocirculatória	126 (30%)
Respiratória	113 (27%)
Neurológica	92 (22%)
Outra disfunção	90 (21%)

Dos 421 pacientes estudados, 33 (7,83%) foram a óbito. A mortalidade estimada pelo PRISM foi de 30,84 (7,22%) e pelo PIM 26,13 (6,21%) pacientes. Isso corresponde a um

SMR (IC = 95%) de 1,07 (0,74-1,50) (z = -0,45) para o PRISM e 1,26 (0,87-1,77) (z = -1,14) para o PIM. Quando testados pelo teste z de Flora, encontram-se dentro dos valores de não-rejeição da hipótese nula (< 1,96 e > -1,96). A Tabela 2 sintetiza a performance dos modelos.

A Tabela 3 avalia as semelhanças na mortalidade observada e esperada em cinco intervalos de risco de mortalidade, através do teste de ajuste de Hosmer-Lemeshow para o PRISM e para o PIM (Tabela 3).

Na performance discriminatória do modelo, medida pela área abaixo da curva ROC, obteve-se uma área de 0,870 (0,810-0,930) para o PRISM e 0,845 (0,769-0,920) para o PIM (Figura 1).

As probabilidades de morte estimadas mostram uma correlação positiva e significativa entre PRISM e PIM, sendo o coeficiente de correlação de postos de Spearman r = 0,65 (p < 0,001).

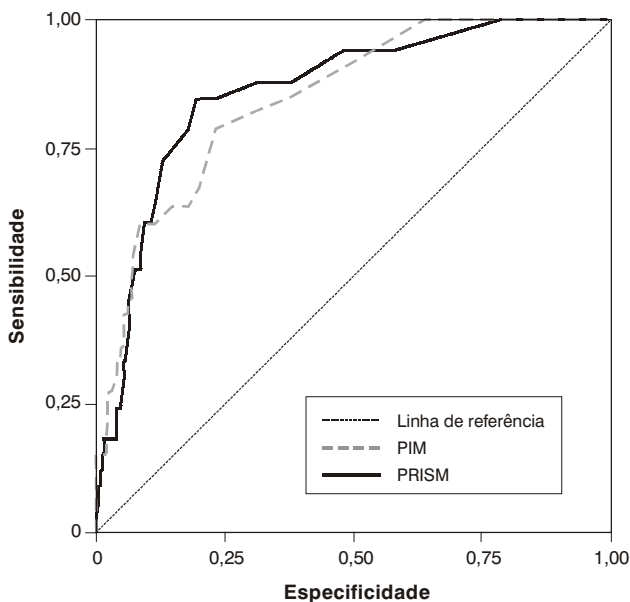


Figura 1 - Superposição de duas curvas típicas de um receptor operador de características (curva ROC). Obteve-se uma área sob a curva ROC de 0,87 (IC95% 0,78-9,2) para o PRISM e 0,85 (IC95% 0,78-0,91) para o PIM. Os resultados encontrados mostram que os testes têm um bom poder de discriminação entre sobreviventes e não-sobreviventes

Discussão

A análise individual dos resultados dos escores pelo SMR mostra que o PRISM e o PIM têm um bom desempenho na previsão da mortalidade geral de nossa população. Embora os dois modelos tenham subestimado a mortalidade (o PRISM previu 93,45% dos óbitos, e o PIM, 79,18%), os dois resultados quando testados não apresentam diferença significativa entre si ou com a mortalidade observada.

Tabela 2 - Performance dos dois escores

	PRISM	PIM
Média de risco de mortalidade; % (DP)	7,3±5,5	6,2±6,1
Mediana de risco de mortalidade; % (IQ)	4,0 (1,5-8,0)	2,0 (1,0-5,0)
Previsão de mortalidade; n	30,84 (7,22%)	26,13 (6,21%)
Índice padronizado de morte (SMR) (IC 95%)	1,07 (0,74-1,50)	1,26 (0,87-1,77)
Escore z de Flora *	-0,45	-1,14
Teste de ajuste de Hosmer-Lemeshow	$\chi^2 = 9,23$, p = 0,100	$\chi^2 = 27,98$, p < 0,001
Área abaixo da curva ROC (Az) (IC 95%)	0,87 (0,81-0,93)	0,85 (0,77-0,92)

DP = desvio padrão; IQ = intervalo interquartil; SMR = *standard mortality ratio*; IC = intervalo de confiança; ROC = receptor operator de características.

* Quando testados pelo teste z de Flora, o PRISM e o PIM encontram-se dentro dos valores de não-rejeição da hipótese nula (< 1,96 e > -1,96).

Tabela 3 - Calibração dos modelos em cinco intervalos de risco de mortalidade

	Risco em %	Número de pacientes	Sobrevivência observada	Sobrevivência esperada	Óbito observado	Óbito esperado
PRISM	0 - 1	83	83	82,95	0	0,05
	> 1-5	200	190	196,78	10	3,22
	> 5-15	84	73	77,6	11	6,4
	> 15-30	24	17	19,88	7	4,12
	> 30-100	30	25	17,66	5	12,34
	Total	421	388	394,87	33	26,13
PIM	0-1	33	33	32,95	0	0,05
	> 1-5	269	261	261,15	8	7,85
	> 5-15	67	56	60,26	11	6,74
	> 15-30	27	19	21,65	8	5,35
	> 30-100	25	19	14,15	6	10,85
	Total	421	388	390,16	33	30,84

Os resultados previstos foram semelhantes aos encontrados no caso do PRISM ($\chi^2 = 9,23$, p = 0,100) e diferentes dos encontrados para o PIM ($\chi^2 = 27,986$, p < 0,001).

Na avaliação do poder de calibração dos escores através do teste de ajuste de Hosmer-Lemeshow, observou-se que os resultados previstos foram semelhantes aos encontrados no caso do PRISM e diferentes dos encontrados para o PIM, indicando uma boa calibração do PRISM e um fraco desempenho para a calibração do PIM.

Ao avaliar a performance discriminatória dos modelos através da curva ROC, os resultados encontrados mostram que tanto o PRISM como o PIM têm um semelhante bom poder de discriminação entre sobreviventes e não-sobreviventes.

O presente estudo procurou validar os escores PRISM e PIM e comparar os resultados encontrados entre eles. Em alguns aspectos, os resultados encontrados pelo PRISM foram um pouco melhores em relação aos resultados do PIM. Seus desempenhos foram semelhantes quan-

to à capacidade de discriminação entre sobreviventes e não-sobreviventes, e eles apresentaram valores com correlação direta e positiva. Entretanto, nesta amostra o PIM apresentou uma fraca capacidade de calibração, um problema comum encontrado em estudos que avaliam escores prognósticos²⁷.

Não existe consenso sobre qual função seria mais importante em um escore prognóstico: calibrar ou discriminar. Ambas são importantes para determinar a capacidade de ajuste do modelo. A função mais importante depende do objetivo para o qual o escore prognóstico está sendo usado²⁸. Se o objetivo for, por exemplo, distinguir aqueles com maior chance de morrer daqueles com maior chance de sobreviver, então a capacidade de discriminar é a mais importante; entretanto, se o objetivo da utilização do escore for comparar a mortalidade observada com a espe-

rada em diferentes intervalos de gravidade, então a capacidade de calibrar é a mais importante. Porém, para uma avaliação global do escore, tanto a discriminação quanto a calibração devem ser consideradas.

O PIM não mostrou uma boa calibração. Uma possível causa para tais resultados seria um número pequeno de mortes em cada nível. Em artigo que avalia o PIM em unidades de terapia intensiva inglesas, o autor sugere que deve haver um cuidado especial quando as diferenças são pequenas em séries pequenas (por exemplo, menos de 20 mortes por unidade)¹⁵. Slater & Shann publicaram recentemente²⁰ um estudo comparativo da performance do PIM, PIM-2, PRISM e PRISM III em unidades da Austrália e Nova Zelândia. Eles constataram que o PIM-2 foi o mais seguro e com melhor ajuste em diferentes grupos diagnósticos. Isso nos permite especular que a nova versão possa vir a apresentar também melhores resultados no nosso meio.

Wells *et al.*²⁹ atribui a duas causas básicas a dificuldade em conseguir fazer com que um paciente com uma instabilidade clínica semelhante a outro doente, ou seja, com igual valor de escore prognóstico, tenha exatamente a mesma evolução. A primeira causa seria a diferença nas condições clínicas individuais, que não são avaliadas pelo escore e que podem influenciar na evolução do paciente, como, por exemplo, a condição nutricional ou reserva fisiológica de cada indivíduo. A segunda causa seria a diferença nas condições de trabalho e infra-estrutura de cada UTIP. Unidades com mais disponibilidade de medicamentos e aparelhos podem oferecer um tratamento mais rápido para seus doentes e, com isso, modificar sua evolução.

No nosso estudo, quase 50% dos pacientes eram provenientes do bloco cirúrgico e chegavam na unidade necessitando de cuidados intensivos. Entretanto, na maioria das vezes, tanto hemodinamicamente quanto ventilatoriamente, estavam compensados. Tais pacientes, embora tivessem um escore de gravidade baixo na admissão (e consequentemente um PIM baixo), eram pacientes com risco de vida e que poderiam deteriorar suas condições nas primeiras 24 horas (e consequentemente com um PRISM mais elevado) pelo risco de complicações pós-cirúrgicas.

Já os pacientes com disfunção respiratória que chegavam na UTIP provenientes da emergência ou de outro hospital poderiam ter uma gasometria dentro dos limites da normalidade às custas de parâmetros elevados de ventilação, com frequência respiratória normal (determinada pelo respirador ou médico que estivesse ventilando com bolsa auto-inflável) e demais condições clínicas estáveis. Esses pacientes apresentavam um escore PIM baixo, embora fossem considerados pacientes graves com risco de mortalidade elevado, pois suas doenças de base ainda não tinham sido resolvidas e poderiam progredir ultrapassando os limites possíveis de ventilação assistida e serem melhor detectados nas próximas horas pelo PRISM.

Fica claro que existem muitas variáveis não mensuradas nos escores prognósticos em estudo, o que dificulta a classificação em níveis de gravidade de diferentes pacientes em diferentes unidades de terapia intensiva e, por isso, tornam difícil encontrar um modelo de índice prognóstico

com boa capacidade de calibração²⁵. O grande desafio está em identificar quais variáveis não têm o poder de predição semelhante para a população em estudo.

A interpretação do índice de mortalidade de uma UTIP é dependente de fatores estatísticos, como tamanho da amostra, taxa de mortalidade em cada nível de gravidade e variações randômicas da população em estudo. A variável de maior poder será aquela que, além de alterar o escore, acontecerá muitas vezes durante o tempo, ou seja, será encontrada em muitos pacientes da população. Devemos, então, buscar o poder das variáveis que mais se assemelham à realidade de nossa população³⁰.

Até que tais questões sejam estabelecidas, avaliando o desempenho do PIM e do PRISM em uma UTIP brasileira, podemos constatar que, embora o PIM apresente uma pior calibração, no conjunto dos resultados ambos apresentam boa capacidade de discriminar entre sobreviventes e não-sobreviventes, constituindo ferramentas de desempenho comparáveis na avaliação prognóstica de pacientes pediátricos admitidos em nossa Unidade.

Referências

- Gemke RJ, Bonsel GJ, Bught AJ. Outcome assessment and quality assurance in pediatric intensive care. In: Tibboel D, van der Voort E, editors. *Intensive care in childhood – a challenge to future*. 2nd ed. Berlin: Springer; 1996. p. 117-32.
- Mitchell I. Nature and nurture: the future of predictor variables. *Curr Opin Crit Care*. 2000;6:166-70.
- Pollack MM, Cuerdon TT, Patel KM, Ruttimann UE, Getson PR, Levetown M. Impact of quality-of-care factors on pediatric intensive care unit mortality. *JAMA*. 1994;272:941-6.
- Seneff M, Knaus WA. Predicting patient outcome from intensive care: a guide to APACHE, MPM, SAPS, PRISM, and other prognostic scoring systems. *J Intensive Care Med*. 1990;5:33-52.
- Pollack MM, Ruttimann UE, Getson PR. The Pediatric Risk of Mortality (PRISM) score. *Crit Care Med*. 1988;16:1110-6.
- PRISM III: an updated Pediatric Risk of Mortality score. PRISM III: An updated Pediatric Risk of Mortality score. *Crit Care Med*. 1996;24:743-52.
- Shann F, Pearson G, Slater A, Wilkinson K. Paediatric index of mortality (PIM): a mortality prediction model for children in intensive care. *Intensive Care Med*. 1997;23:201-7.
- Slater A, Shann F, Pearson G. PIM2: a revised version of the Paediatric Index of Mortality. *Intensive Care Med*. 2003;29:278-85.
- Gunning K, Rowan K. ABC of intensive care outcome data and scoring systems. *BMJ*. 1999;319:241-4.
- Marcin JP, Pollack MM, Patel KM, Ruttimann UE. Combining physician's subjective and physiology-based objective mortality risk predictions. *Crit Care Med*. 2000;28:2984-90.
- Gemke RJ, van Vught J. Scoring systems in pediatric intensive care: PRISM III versus PIM. *Intensive Care Med*. 2002;28:204-7.
- Tibby SM, Taylor D, Festa M, Hanna S, Hatherill M, Jones G, *et al.* A comparison of three scoring systems for mortality risk among retrieved intensive care patients. *Arch Dis Child*. 2002;87:421-5.
- Slater A. Monitoring outcome in paediatric intensive care. *Paediatr Anaesth*. 2004;14:113-6.
- Jones GD, Thorburn K, Tigg A, Murdoch IA. Preliminary data: PIM vs PRISM in infants and children pos cardiac surgery in a UK PICU. *Intensive Care Med*. 2000;26:145.
- Pearson GA, Stickley J, Shann F. Calibration of the paediatric index of mortality in UK paediatric intensive care units. *Arch Dis Child*. 2001;84:125-8.

16. Leteurtre S, Leclerc F, Martinot A, Cremer R, Fourier C, Sadik A, et al. Can generic scores (Pediatric Risk of Mortality and Pediatric Index of Mortality) replace specific scores in predicting the outcome of presumed meningococcal septic shock in children? *Crit Care Med.* 2001;29:1239-46.
17. Slater A, Shann F, Pearson G. Paediatric Index of Mortality (PIM) Study Group. PIM2: a revised version of the Paediatric Index of Mortality. *Intensive Care Med.* 2003;29:278-85.
18. Shann F. Are we doing a good job: PRISM, PIM and all that. *Intensive Care Med.* 2002;28:105-7.
19. Marcin JP, Pollack MM. Review of the methodologies and applications of scoring systems in neonatal and pediatric intensive care. *Pediatr Crit Care Med.* 2000;1:20-7.
20. Slater A, Shann F, ANZICS Paediatric Study Group. The suitability of the Pediatric Index of Mortality (PIM), PIM2, the Pediatric Risk of Mortality (PRISM), and PRISM III for monitoring the quality of pediatric intensive care in Australia and New Zealand. *Pediatr Crit Care Med.* 2004;5:447-54.
21. Castellanos-Ortega A, Delgado-Rodriguez M, Llorca J, Sanchez Buron P, Mencia Bartolome S, Souto Rubio A, et al. A new prognostic scoring system for meningococcal septic shock in children. Comparison with three other scoring systems. *Intensive Care Med.* 2002;28:341-51.
22. Ozer EA, Kizilgunesler A, Sarioglu B, Halicioglu O, Sutcuoglu S, Yaprak I. The Comparison of PRISM and PIM Scoring Systems for Mortality Risk in Infantile Intensive Care. *J Trop Pediatr.* 2004;50:334-8.
23. Flora, JD. A method for comparing survival of burn patients to a standard survival curve. *J Trauma.* 1978;18:701-8.
24. Hosmer DW, Lemeshow S. *Applied logistic regression.* New York: John Wiley; 1989.
25. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology.* 1983;148:839-43.
26. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem.* 1993;39:561-77.
27. Moreno R, Matos R. The new scores: what problems have been fixed, and what remain? *Curr Opin Crit Care.* 2000;6:158-65.
28. Mourouga P, Goldfrad C, Rowan KM. Does it fit? Assessment of scoring systems. *Curr Opin Crit Care.* 2000;6:176-80.
29. Wells M, Riera-Fanego JF, Luyt DK, Dance M, Lipman J. Poor discriminatory performance of the Pediatric Risk of Mortality (PRISM) score in a South African intensive care unit. *Crit Care Med.* 1996;24:1507-13.
30. Rowan KM, Angus DC. Don't let perfection be the enemy of the good: it's for optimism over the role of severity scoring systems in intensive care unit performance measurement. *Curr Opin Crit Care.* 2000;6:153-4.

Correspondência:
Pedro Celiny Ramos Garcia
Rua Curupaiti, 62
CEP 90820-090 – Porto Alegre, RS
Fone: (51) 3266.5121
E-mail: celiny@terra.com.br