

Escalas Preditivas de Mortalidade em Cuidados Intensivos Pediátricos

PAULO OOM

*Clínica Universitária de Pediatria
Hospital de Santa Maria*

Resumo

As escalas preditivas de mortalidade têm sido utilizadas com uma frequência crescente na avaliação da qualidade da assistência prestada a crianças internadas em unidades de cuidados intensivos, na estratificação de doentes em projectos de investigação clínica e no apoio a algumas decisões clínicas individuais. Existem actualmente dois modelos preditivos de mortalidade para utilização em cuidados intensivos pediátricos a *Pediatric Risk of Mortality* (PRISM) e a *Pediatric Index of Mortality* (PIM). Ambos demonstraram uma calibração e poder discriminativo excelentes nas unidades onde foram desenvolvidos mas está ainda por esclarecer o seu desempenho relativo quando aplicados em simultâneo numa mesma população de crianças internadas em cuidados intensivos, independente das que foram utilizadas na sua construção. Este estudo decorreu de forma prospectiva e incluiu 580 internamentos consecutivos na unidade de cuidados intensivos de Pediatria do Hospital de Santa Maria. De cada criança foram obtidos dados demográficos e todas as variáveis necessárias para o cálculo das diferentes escalas de gravidade de acordo com as instruções dos seus autores originais. Ambas as escalas preditivas de mortalidade apresentaram um desempenho adequado. A colheita das variáveis da escala PRISM é mais complexa de que resultou uma maior percentagem de variáveis não colhidas. A escala PIM é de utilização gratuita, mais simples e mais recente pelo que deveria ser utilizada por rotina nas unidades de cuidados intensivos pediátricos.

Palavras-Chave: Cuidados intensivos pediátricos, escalas preditivas, mortalidade, PRISM, PIM

Summary

Scoring Systems for Mortality Risk in Pediatric Intensive Caring

Severity scoring models have frequently been used as a tool in the assessment of quality of paediatric intensive care units, in clinical trials and as support for difficult medical decisions. The Paediatric Risk of Mortality and the Paediatric Index of Mortality are two available scores to be used in paediatric intensive care. Both have demonstrated an excellent performance in the units where they were developed but there are still doubts about their relative performance on an independent database. We conducted a prospective study of the performance of both models in the Paediatric Intensive Care Unit at Santa Maria Hospital in Lisbon, Portugal. Both demonstrated a good performance. The collection of PRISM variables is more complex which resulted in a high percentage of missing values. PIM is freely available, easiest to calculate and was developed from data collected more recently. It should be used by routine in paediatric intensive care.

Key-Words: Paediatric intensive care, scores, mortality, PRISM, PIM

Introdução

A determinação da gravidade da doença constitui um aspecto fundamental na avaliação da qualidade dos cuidados prestados nas unidades de cuidados intensivos pediátricos permitindo uma melhor gestão dos seus recursos humanos, materiais e financeiros. É igualmente imprescindível na interpretação dos resultados de trabalhos de investigação clínica e no apoio à decisão médica em situações como a triagem de doentes, protocolos de actuação terapêutica e em decisões mais difíceis como as que envolvem a suspensão de medidas terapêuticas em doentes com prognóstico grave.

Existem actualmente disponíveis duas escalas prediti-

Correspondência: Paulo Oom

Unidade de Cuidados Intensivos
Serviço de Pediatria
Hospital de Santa Maria
Av. Prof. Egas Moniz 1649-035 Lisboa

vas de mortalidade para utilização em cuidados intensivos pediátricos: a *Pediatric Risk of Mortality* (PRISM) e a *Pediatric Index of Mortality* (PIM). Ambas foram construídas de forma prospectiva com base nos registos de um grande número de crianças e utilizando a regressão logística para o cálculo da equação preditiva de mortalidade.^{1,2}

Qualquer destas escalas foi já utilizada num grande número de unidades de cuidados intensivos pediátricos (UCIP) em diferentes partes do mundo abrangendo um enorme número de doentes com grande diversidade de grupos etários e de diagnósticos³⁻²⁸ mas a grande heterogeneidade nas populações estudadas, aliada à diversidade na forma de apresentação dos resultados torna muito difícil uma comparação dos resultados dos diferentes trabalhos. Apenas dois estudos avaliaram de forma prospectiva as escalas PRISM e PIM numa mesma população independente de crianças^{17,20} e permanece por definir qual a escala que melhor se adapta à realidade dos cuidados intensivos pediátricos. Foram objectivos deste trabalho avaliar, de forma prospectiva e em simultâneo, numa amostra significativa de crianças internadas numa unidade de cuidados intensivos, o desempenho relativo das escalas preditivas de mortalidade (PRISM e PIM).

Doentes e Métodos

O estudo foi realizado na Unidade de Cuidados Intensivos de Pediatria do Hospital de Santa Maria. Foram incluídas todas as crianças internadas na unidade durante um período de 21 meses, excepto as que apresentavam algum critério para exclusão: 1 – admissão não se justificava, por se tratar de situações que não exigiam cuidados intensivos e que, por rotina, seriam tratadas em outros locais do Serviço de Pediatria, 2 – tempo de internamento inferior a 2 horas, 3 – admissão durante manobras de reanimação cardiorespiratória e não tendo sido obtidos sinais vitais estáveis durante um período de, pelo menos, 2 horas. Foram obtidos de forma prospectiva e de cada criança os seguintes dados: data de nascimento, data de internamento, idade, sexo, diagnóstico de admissão de acordo com a classificação ICD-9,²⁹ tipo de admissão (electiva, emergência), origem (enfermaria, outro hospital, urgência), destino (enfermaria, outro hospital, alta) e data de alta. Foram consideradas em pós-operatório as crianças internadas até 24 após o procedimento cirúrgico. Para o cálculo da escala PIM foi considerado para cada variável o primeiro valor obtido até uma hora após a admissão.² Para o cálculo da escala PRISM foi considerado o pior valor de cada variável durante as primeiras 24 horas após a admissão.¹ Nos doentes falecidos nas primeiras 24 horas de internamento não foram considerados os valores de qualquer variável obtidos no período pré-terminal (últimas duas horas de vida). Para avaliação do grau de coma foi utilizada a vari-

ante pediátrica da escala de coma de Glasgow,³⁰ desenvolvida por Reilly.³¹ Apenas foram obtidas as variáveis que o clínico considerou úteis para a abordagem diagnóstica e terapêutica de cada criança. Valores não obtidos foram considerados normais.

A mortalidade foi entendida como mortalidade durante o internamento na UCIP. As mortes intra-operatórias foram incluídas se o procedimento cirúrgico ocorreu durante o internamento na unidade e consistiu numa medida terapêutica dirigida à causa que motivou a admissão na unidade.

O número de crianças a estudar foi determinado tendo em conta a necessidade de incluir pelo menos 20 falecimentos o que corresponde ao limite geralmente aceite de interesse clínico no estudo da mortalidade em cuidados intensivos pediátricos.²³ Para uma mortalidade estimada de 3,5% foi calculada a necessidade de incluir 571 internamentos.

Este trabalho teve o parecer favorável da Comissão de Ética da Faculdade de Medicina da Universidade Clássica de Lisboa e da Comissão de Ética para a Saúde do Hospital de Santa Maria.

Todos os dados obtidos foram introduzidos no programa Excel versão 5.0 para Windows (Microsoft Corporation, Seattle, WA). A folha de cálculo foi programada de forma a permitir o cálculo automático das estimativas de probabilidade de morte. Para a validação das escalas PRISM e PIM foram determinadas a sua calibração e poder discriminativo. A calibração foi testada pelo cálculo da razão de mortalidade padronizada (*standardized mortality rate*, SMR) e pela aplicação do teste de χ^2 *Goodness-of-fit* de Hosmer e Lemeshow.^{32,33} A discriminação foi testada pelo cálculo da área abaixo da curva ROC (*area under curve*, AUC) como proposto por Hanley e McNeil.³⁴ A análise estatística foi efectuada utilizando o programa Stata versão 7.0 para Windows (Stata Corporation, College Station, TX). Foram considerados significativos valores de $p < 0,05$.

Resultados

O estudo decorreu entre 1 de Fevereiro de 2001 e 31 de Outubro de 2002. Durante este período foram internadas na UCIP 625 crianças. Destas, apresentavam um ou mais critérios de exclusão 45 pelo que foram analisados os dados respeitantes a 580 internamentos. As características desta população estão discriminadas na tabela I. As causas de internamento, encontram-se discriminadas na tabela II. A idade variou entre os 2 dias de vida e os 19 anos com uma mediana de 2,6 anos. O tempo de internamento variou entre 1 e 113 dias, com uma mediana de 2 dias. Faleceram 31 crianças correspondendo a 5,3% dos internamentos.

As estimativas da probabilidade de morte determinadas pela escala PRISM variaram entre 0 e 0,98 com uma

mediana de 0,01 e os percentis 25 e 75 de, respectivamente, 0,01 e 0,04. A distribuição de valores está representada na figura 1.

Tabela I
Características da amostra (n=580)

Sexo	Masculino	338 (58,3)
	Feminino	242 (41,7)
Admissão	Emergência	386 (66,6)
	Electiva	194 (33,4)
	Readmissões	110 (19,0)
Doentes médicos Doentes cirúrgicos		371 (64,0)
		209 (36,0)
Origem	Enfermaria	306 (52,8)
	Outro hospital	182 (31,4)
	Urgência	92 (15,8)
Destino	Enfermaria	469 (80,9)
	Outro hospital	51 (8,8)
	Alta	29 (5,0)
	Falecido	31 (5,3)

Resultados expressos em número absoluto (percentagem).
Os doentes cirúrgicos representam as crianças admitidas provenientes do bloco operatório

Tabela II
Grupos de diagnóstico na admissão

Grupo	
Respiratório	140 (24,1)
Sistema nervoso	96 (16,6)
Oncologia	83 (14,3)
Trauma	59 (10,2)
Infecçiológica	54 (9,3)
Musculo-esquelético	42 (7,2)
Circulatório	30 (5,2)
Digestivo	22 (3,8)
Endocrinologia / metabolismo	12 (2,1)
Acidentes	9 (1,6)
Genitourinário	9 (1,6)
Hematologia	5 (0,9)
Condições mal definidas	19 (3,3)

Resultados expressos em número absoluto (percentagem)

A escala PRISM mostrou uma boa calibração com uma razão de mortalidade padronizada (*standardized mortality ratio*, SMR) de 0,90 com intervalo de confiança de 95% (IC 95%) entre 0,61 e 1,27. O teste de Hosmer-Lemeshow mostrou um valor de $\chi^2(8)=6,23$ ($p=0,62$). O poder discriminativo da escala PRISM, avaliado através da determinação da área sob a curva ROC (*area under curve*, AUC) foi de 0,91 (IC 95% 0,86 - 0,96) como se observa na figura 2.

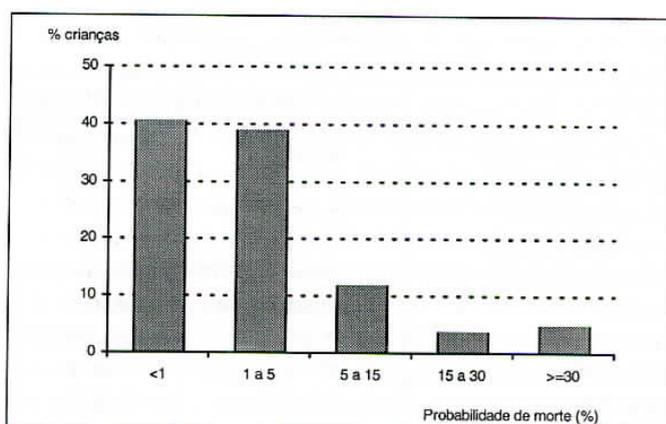


Fig. 1 – Distribuição das estimativas de probabilidade de morte de acordo com a escala PRISM

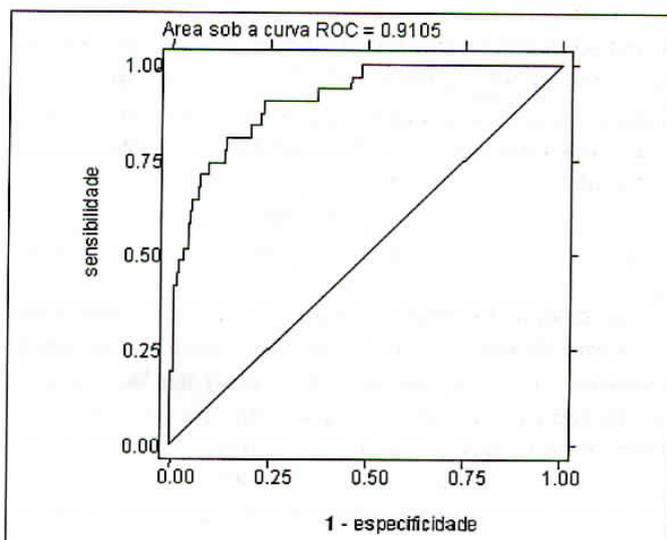


Fig. 2 – Poder discriminativo da escala PRISM

Nem todos os componentes da escala PRISM foram obtidos em todos os doentes e, de acordo com as instruções originais, nestes casos foi considerado não estarem alterados. Os componentes da escala que foram omitidos mais frequentemente foram, por ordem descendente de frequência a gasimetria arterial, bilirrubinemia, calcemia e provas de coagulação.

Também o número de alterações observadas em cada parâmetro foi muito variável. Em todos os parâmetros o valor mais frequentemente atribuído foi zero. Este facto ocorreu em mais de 90% das crianças em relação à pressão arterial diastólica, gasimetria arterial, escala de coma, avaliação pupilar, bilirrubinemia e calcemia conforme discriminado na tabela III.

Na aplicação da escala PIM as estimativas de probabilidade de morte variaram entre 0 e 0,94 com uma medi-

anças. Desta forma ambas as escalas apresentaram uma SMR próxima da unidade indicando que a mortalidade prevista pelas escalas se aproximou da mortalidade real. O teste de Hosmer-Lemeshow revelou em ambas valores de $p > 0,05$ mostrando que ao longo dos diferentes intervalos de risco a mortalidade observada não foi significativamente diferente da mortalidade prevista pelas escalas.

Tabela V
Calibração e poder discriminativo da escala PRISM em diferentes trabalhos

	Tipo	UCIP	n	MO (%)	SMR (IC 95%)	χ^2 (p)	AUC (IC 95%)
Pollack, 1988 ¹	p	6	1.227	8,6	1,01	>0,95	0,92
Goddard, 1992 ⁷	p	1	380	5,3	0,61	<0,05	-
Balakrishnan, 1992 ¹¹	p	1	270	10,7	0,94	0,12	-
Gemke, 1994 ⁴	p	1	593	8,4	1,04	0,35	0,92
Estrada, 1994 ¹²	p	1	1.131	5,9	1,09	-	-
Gemke, 1995 ¹³	p	10	1.063	7,1	0,99	0,19	-
DiCarlo, 1996 ¹⁴	p	4	583	9,1	1,32	0,02	-
Wells, 1996 ⁴	p	1	1.528	32,0	1,10	<0,01	0,73
Radovan, 1996 ¹⁵	p	1	100	11,0	0,85	0,03	-
Earle, 1997 ⁷	p	6	1.058	19,0	1,72	<0,01	-
Martinot, 1997 ³	p	9	712	13,0	1,18	0,01	0,87
Shann, 1997 ²	p	1	1.182	6,6	-	<0,01	0,87
Estrada, 1997 ¹⁴	r	1	1.120	5,0	0,91	0,10	0,96
Bertolini, 1998 ¹⁶	p	26	1.533	11,4	1,04	<0,01	0,85 (0,83-0,87)
Tilford, 1998 ⁴	p	16	10.833	4,5	-	0,01	0,90
Estrada, 2001 ⁴	r	1	2.944	5,8	1,01	0,83	0,92
Leclerc, 2001 ¹⁷	p	1	802	10,0	-	-	0,86
Cunha, 2002 ¹⁸	r	1	1.153	14,1	0,91	0,06	0,92
Tibby, 2002 ¹⁹	p	2	928	7,8	0,71 (0,53-0,89)	0,08	0,86 (0,81-0,92)

UCIP-número de unidades de cuidados intensivos pediátricas envolvidas no estudo; MO-mortalidade observada; SMR-*standardized mortality ratio*; χ^2 -teste de Hosmer-Lemeshow; AUC-área sob a curva ROC; IC 95%-intervalos de confiança de 95%; p-prospectivo; r-retrospectivo

Existem diversos trabalhos onde é feita a avaliação do desempenho destas escalas em diferentes UCIP mas a grande heterogeneidade das populações estudadas aliada à diversidade na forma de apresentação dos resultados torna muito difícil uma comparação dos resultados. Para além disso, muitos estudos são incompletos apresentando apenas resultados parciais (tabelas V e VI).

Tabela VI
Calibração e poder discriminativo da escala PIM em diferentes trabalhos

	Tipo	UCIP	n	MO (%)	SMR (IC 95%)	χ^2 (p)	AUC (IC 95%)
Shann, 1997 ²	p	8	5695	4,9	1,00	0,37	0,90
Jones, 1999 ²⁰	p	1	337	5,6	1,17	0,08	0,84 (0,74-0,95)
Pearson, 2001 ²¹	p	5	7253	8,0	0,87 (0,81-0,94)	-	0,84 (0,82-0,85)
Leclerc, 2001 ¹⁷	p	1	802	10,0	-	-	0,83
Gemke, 2002 ¹³	p	1	303	6,6	0,88 (0,55-1,20)	0,77	0,74 (0,63-0,85)
Shann, 2002 ²²	p	1	1023	4,0	0,69	-	-
Tibby, 2002 ¹⁹	p	2	928	7,8	0,57 (0,44-0,70)	<0,01	0,87 (0,82-0,92)
Sanz, 2002 ²³	p	1	500	7,2	-	-	0,81 (0,74-0,89)
Davis, 2003 ²⁴	p	1	1163	1,5	0,49 (0,26-0,72)	-	-
Hatherill, 2003 ²⁵	p	1	46	35,0	1,03 (0,71-1,35)	-	0,71 (0,53-0,88)

UCIP-número de unidades de cuidados intensivos pediátricas envolvidas no estudo; MO-mortalidade observada; SMR-*standardized mortality ratio*; χ^2 -teste de Hosmer-Lemeshow; AUC-área sob a curva ROC; IC 95%-intervalos de confiança de 95%; p-prospectivo; r-retrospectivo

Como esperado, a calibração de qualquer uma das escalas, apesar de excelente na população que serve de base para a construção da escala^{1,2} torna-se deficiente em muitas situações quando aplicada numa outra população. Alguns cuidados devem ser postos na avaliação destes resultados de forma a evitar interpretações abusivas. A calibração de um modelo, quando aplicado numa UCIP diferente da que foi utilizada na sua construção pode estar alterada por o desempenho (e a qualidade de cuidados) não ser idêntica à da unidade de referência ou porque os seus doentes são diferentes, no tipo de patologias na intensidade das lesões ou na forma como estão doentes. A interpretação pode ser difícil e nem sempre é objectiva. Uma boa calibração, traduzida por SMR cujos intervalos de confiança de 95% incluem a unidade e valores de p superiores a 0,05 no teste de Hosmer-Lemeshow significa habitualmente que o modelo está bem calibrado e que o desempenho da UCIP é semelhante ao das unidades onde as escalas foram elaboradas, na altura em que foram elaboradas. O contrário nem sempre é verdade e na interpretação de uma calibração deficiente o tipo de doentes (idades e patologias) e a sua distribuição por intervalos de risco deve ser cuidadosamente analisada e só se for semelhante à da população original se podem retirar algumas conclusões sobre a qualidade dos cuidados aí prestados.

No nosso trabalho ambas as escalas mostraram uma boa capacidade discriminativa, traduzida por áreas sob a curva ROC superiores a 0,90. Nos trabalhos originais ambas as escalas mostraram um excelente poder discriminativo com AUC de 0,92 e 0,90 para PRISM e PIM respectivamente. Na generalidade das unidades de cuidados intensivos onde estes modelos foram posteriormente aplicados (tabelas V e VI), as escalas revelaram valores superiores a 0,70, limite considerado aceitável quando avaliado o poder discriminativo de um modelo num contexto diferente daquele em que foi elaborado.^{25,35} Com poucas excepções aquele valor é mesmo geralmente superior a 0,80 mostrando que ambas as escalas são robustas e mantêm o seu poder discriminativo mesmo quando aplicadas em outras UCIP.

O nosso trabalho representa a primeiro estudo prospectivo de validação das escalas PRISM e PIM numa população de crianças internadas numa UCIP portuguesa. Anteriormente apenas três trabalhos foram publicados sobre este tema^{9,13,18} mas são todos de natureza retrospectiva, apenas avaliam a escala PRISM e dois deles referem-se à mesma UCIP analisada em diferentes ocasiões.^{9,18} No seu conjunto, estes trabalhos permitiram validar a escala PRISM para utilização nessas unidades ao mostrar que aquela encontra-se bem calibrada e com excelente poder discriminativo (AUC \geq 0,92), o que está de acordo com os nossos resultados.

Apenas dois outros trabalhos avaliaram de forma

prospectiva as escalas PRISM e PIM numa mesma população independente de crianças internadas em cuidados intensivos. Leclerc¹⁷ descreve um poder discriminativo semelhante com AUC de 0,86 e 0,83 para PRISM e PIM, respectivamente. Num trabalho mais recente, Tibby²⁰ mostra que as escalas sobrestimam a mortalidade traduzindo-se por valores de SMR cujos intervalos de confiança não incluem a unidade. No entanto, ambas mostraram uma boa capacidade discriminativa (AUC de 0,86 e 0,87 para PRISM e PIM, respectivamente). Num outro trabalho,²² abrangendo apenas crianças submetidas a cirurgia cardíaca, as escalas apresentaram um poder discriminativo semelhante mas com a escala PIM melhor calibrada (PRISM: $\chi^2_{(8)}=21,4$ $p=0,006$, AUC=0,92 e PIM: $\chi^2_{(8)}=14,1$ $p=0,08$, AUC=0,84). A comparação entre as escalas PIM e PRISM III, apesar da muito maior complexidade desta última,²⁴ não mostrou diferenças significativas na calibração ou poder discriminativo: PRISM III $\chi^2_{(8)}=13,3$ ($p=0,21$) AUC=0,78 e PIM $\chi^2_{(8)}=4,92$ ($p=0,77$) AUC=0,74.

A escolha de um modelo preditivo de mortalidade para uso por rotina na UCIP deve ser tomada tendo em conta múltiplos aspectos.^{20,24,36-38} É fundamental analisar o seu desempenho (calibração e discriminação) no novo contexto e só um modelo bem validado deve ser utilizado. As escalas PRISM e PIM mostraram uma calibração e discriminação excelentes nas populações e UCIP onde foram desenvolvidas.¹² Na nossa unidade apresentam um desempenho semelhante com uma calibração e poder discriminativo bons.

O modelo deve ser simples, fácil de utilizar no dia-a-dia atarefado de uma unidade de cuidados intensivos. Uma maior complexidade torna um modelo mais robusto e capaz de detectar pequenas variações de gravidade ou manter o interesse na presença de poucos dados. Mas a complexidade, ao tornar o modelo mais fiel à realidade, torna-o também mais difícil de utilizar e contraria o seu principal objectivo: generalizar e abstrair. A escala PRISM é complexa, exigindo a colheita do pior valor de cada uma de 14 variáveis, a que correspondem 33 intervalos predefinidos, nas primeiras 24 horas de internamento.¹ A sua última versão (PRISM III) apresenta ainda maior complexidade pois exige a obtenção do pior valor de 17 variáveis para além de oito variáveis adicionais.³⁹ Pelo contrário, a escala PIM é de mais simples aplicação pois apenas necessita da colheita de 8 variáveis, sem intervalos predefinidos, na primeira hora após a admissão.²

Nenhum modelo demonstrou ser estável ao longo do tempo. O aparecimento de novas doenças, por um lado, e o desenvolvimento de meios de diagnóstico e terapêutica novos e mais sofisticados, por outro, tem permitido uma maior sobrevida dos doentes internados na UCIP. Será por isso de esperar que, ao longo do tempo, se observe uma redução da relação mortalidade observada/mortalidade

prevista, pelo menos para os graus intermédios de gravidade.^{5,20,40} Por este motivo, qualquer modelo deve ser periodicamente validado e, se necessário, actualizada a sua equação matemática. O modelo PRISM foi desenvolvido a partir de uma base de dados de doentes colhida entre 1980-1985. Pelo contrário, o modelo PIM é mais recente, desenvolvido com base em dados colhidos entre 1994-1996 e deverá estar melhor adaptado aos progressos entretanto verificados em medicina intensiva.

Alguma controvérsia tem sido criada em relação ao tempo em que são colhidas as diferentes variáveis. Um modelo como o PRISM, que utiliza o pior valor registado nas primeiras 24 horas após a admissão parece sempre mais preciso do que é na realidade. Ao aguardar 24 horas para prever a mortalidade, e sabendo que um grande número de mortes nas UCIP ocorre no primeiro dia de internamento,^{18,20,23,26,41,42} o modelo pode estar a diagnosticar a morte mais do que a prevê-la. Argumentam os autores que a escala foi calculada para ser utilizada nas UCIP ao mesmo tempo que as medidas terapêuticas estão a ser aplicadas, e o objectivo é a medição da gravidade da doença enquanto é instituída a terapêutica (presume-se óptima).⁴³ Ao estender a colheita por 24 horas, a escala pode também atenuar diferenças que existam entre UCIP: se uma criança é tratada incorrectamente nas primeiras 24 horas após a sua admissão a sua pontuação será pior e a mortalidade atribuída ao facto de ser uma criança com uma doença mais grave. E assim a maior mortalidade de uma unidade de menor qualidade é atribuída ao facto de ter (supostamente) doentes mais graves.^{2,41,44} Segundo os autores, esta preocupação é atenuada pelo facto de uma eventual má qualidade de cuidados se estender para lá do dia de admissão pelo que a referida unidade irá sempre apresentar uma SMR elevada.⁴⁵ Pelo contrário, um modelo que utiliza valores obtidos no primeiro contacto com a criança (como o PIM) não apresenta estes inconvenientes, sendo mais independente das atitudes terapêuticas iniciais (deletérias ou não) na avaliação da gravidade da doença.² Mas está mais sujeito a ser influenciado por todos os tratamentos que foram efectuados naquela criança imediatamente antes da sua entrada na UCIP.⁴⁵ Ou seja, as variáveis colhidas à entrada na UCIP podem não corresponder à verdadeira gravidade da situação dado que alguns desequilíbrios, nomeadamente hidroelectrolíticos e metabólicos, poderão ter sido já parcialmente corrigidos durante a permanência da criança no serviço de urgência ou durante o transporte para a UCIP. Argumentam os autores da escala que estas possíveis correcções alteram também o prognóstico daquele doente e a escala não faz mais do que medir, à entrada na unidade, a gravidade da doença.² Mas é um facto que, se tais correcções tiverem efeitos transitórios, a gravidade da doença pode ser subestimada quando avaliada por esta escala.⁴⁶

Qualquer modelo deve apenas requerer dados objec-

tivos obtidos por rotina na monitorização da criança e na avaliação da sua resposta aos tratamentos efectuados. Em ambas as escalas é assumido que as variáveis não obtidas são consideradas normais. Mas nem sempre este pressuposto é verdade e a omissão pode ocorrer por falha técnica (como seja por impossibilidade de colheita de sangue, extravio dos produtos antes da chegada ou no laboratório) ou pela impressão do clínico da inutilidade em obter determinado valor que, teoricamente, já sabe. Este facto é especialmente frequente em doentes crónicos em que determinados parâmetros laboratoriais se sabe estarem alterados ou na avaliação neurológica de doentes sedados em ventilação mecânica.^{38,47} Isto significa que a um doente, mesmo grave, no qual falhe a colheita de um número significativo de variáveis é atribuída uma menor estimativa de probabilidade de morte. Um modelo deve ser suficientemente robusto para permitir que algumas falhas não influenciem significativamente o seu desempenho mas falhas que ocorrem de forma sistemática podem levar o modelo a perder a sua calibração. No nosso trabalho 4 variáveis da escala PRISM não foram obtidas em mais de 50% das crianças (gases arteriais, coagulação, bilirrubina e cálcio séricos). Todas as variáveis obtidas foram consideradas normais na maioria dos doentes mas algumas foram normais em mais de 90% dos casos, incluindo as já referidas em conjunto com a pressão arterial diastólica, escala de coma de Glasgow e reacção pupilar à luz. Estes dados estão de acordo com os referidos por Tibby que encontrou falhas na colheita de variáveis em 40% dos doentes sendo as mais frequentes as mesmas por nós referidas.²⁰ Também Balakrishan refere que, numa amostra de 270 crianças nenhuma pontuou na variável pressão arterial diastólica e apenas uma (correspondendo a 0,4% dos casos) na bilirrubina sérica.¹¹ Constatou igualmente a necessidade de revisão dos limiares propostos para a pressão arterial sistólica, frequência cardíaca, coagulação e potássio. No mesmo trabalho verificou que apenas cinco variáveis estavam relacionadas de forma independente com a mortalidade: pressão arterial sistólica, escala de coma de Glasgow, pressão arterial de dióxido de carbono, bicarbonato e cálcio.¹¹ Estes dados podem significar a necessidade de revisão de algumas variáveis ou dos limiares a elas atribuídos. Apesar de os autores terem demonstrado que a ausência de medição de algumas variáveis não altera o desempenho do modelo ⁴³ é um facto que, na sua última versão (PRISM III), algumas variáveis foram suprimidas como a pressão arterial diastólica, bilirrubina e cálcio.⁴⁸ Em relação à escala PIM no nosso trabalho apenas os gases arteriais não foram obtidos em mais de 50% das crianças. Tal como na escala PRISM também aqui a reacção pupilar à luz foi considerada normal em mais de 90% dos casos. Tibby encontrou falhas em 12% dos doentes, com destaque para os gases arteriais e o excesso de bases, o que está de acordo com os nossos resultados.²⁰ Os nossos dados, em

conjunto com os dos outros autores, vêm reforçar a ideia de que a uma maior complexidade associa-se um maior número de falhas na colheita das variáveis, e que o número de falhas será tanto maior quanto maior o número de variáveis e o tempo necessário para a sua obtenção, como acontece com a escala PRISM. A eventual reavaliação do número e peso atribuído a cada variável tem de obedecer, no entanto, a critérios matemáticos precisos pois mesmo uma variável pontuada raramente (como a reacção pupilar à luz) pode ter um significado suficientemente grave para justificar a sua inclusão no modelo preditivo de mortalidade.

A escala PRISM, por ser a pioneira, é a utilizada pela maioria das unidades de cuidados intensivos que avaliam regularmente a qualidade dos cuidados prestados. Este é um factor relevante quando se pretende comparar o desempenho de uma unidade de cuidados intensivos com outras unidades vizinhas. Mas a facilidade de obtenção do modelo é igualmente um factor a ter em conta. A equação preditiva da última versão da escala PRISM (PRISM III) ⁴⁸ é propriedade privada e é necessário o pagamento de uma anuidade para a ela ter acesso. Os autores justificam este facto argumentando que a actualização periódica da escala só será possível com um adequado suporte financeiro.^{45,49} Mas é também verdade que tem sido este facto um dos principais factores a impedir a difusão da escala e a causa de alguma controvérsia.⁵⁰

Em resumo, ambas as escalas PRISM e PIM utilizam apenas dados objectivos e apresentam um desempenho semelhante com uma calibração e poder discriminativo bons. A única vantagem da escala PRISM é a sua maior difusão entre diferentes UCIP permitindo desta forma uma comparação mais fácil entre unidades. Por outro lado, a escala PIM é de utilização gratuita, muito simples e de mais fácil aplicação, onde apenas num pequeno número de doentes falha a colheita de alguma variável. Para além disso a sua equação preditiva baseia-se em dados colhidos mais recentemente. Por fim, a estimativa gerada pela escala não se altera pelo eventual mau desempenho da UCIP dado considerar apenas o primeiro valor de cada variável após a admissão. Tomados no seu conjunto estes dados parecem justificar a utilização da escala PIM como rotina na UCIP. Seria interessante a realização de um estudo multicêntrico em UCIP portuguesas para avaliar o desempenho da sua última versão.⁵¹

Agradecimentos

O autor agradece ao Prof. Doutor J. Gomes Pedro, Prof. Doutor A Gouveia de Oliveira e Dra. Manuela Correia a revisão crítica deste trabalho. Da mesma forma agradece à Dra Filipa Negreiro o cuidado posto na análise estatística dos resultados.

Bibliografia

1. Pollack M, Ruttimann U, Getson P: Pediatric Risk of Mortality (PRISM) score. *Crit Care Med* 1988; 16: 1110-6
2. 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
3. Earle M, Natera O, Zaslavsky A, Quinones E, Carrillo H, Gonzalez E *et al*: Outcome of pediatric intensive care at six centers in Mexico and Ecuador. *Crit Care Med* 1997; 25: 1462-7
4. Wells M, Riera-Fanego J, Luyt D, 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
5. Martinot A, Grandbastien B, Leclerc F: Performance of the Pediatric Risk of Mortality score. *Crit Care Med* 1997; 25: 716-7
6. Tilford J, Roberson P, Lensing S, Fiser D: Differences in pediatric ICU mortality risk over time. *Crit Care Med* 1998; 26: 1737-43
7. Goddard J: Pediatric risk of mortality scoring overestimates severity of illness in infants. *Crit Care Med* 1992; 20: 1662-5
8. Gemke R, Bonsel G, Vugth A: Effectiveness and efficiency of a Dutch pediatric intensive care unit: Validity and application of the Pediatric Risk of Mortality score. *Crit Care Med* 1994; 22: 1477-84
9. Estrada J, Vale M, Ventura L, Santos M, Vasconcelos C: Mortalidade numa unidade de cuidados intensivos pediátricos. *Acta Pediatr Port* 2001; 2: 71-6
10. Ripamonti D, Favero A, Polidori G, Albonico L, Leonetti C, Cattaneo A *et al*: Assessment of mortality risk in children: validation of the PRISM score in a Italian sample. Results from GIViTi. *Crit Care Med* 1996; 24(Suppl): 32
11. Balakrishnan G, Aitchison T, Hallworth D, Morton N: Prospective evaluation of the pediatric risk of mortality (PRISM) score. *Arch Dis Child* 1992; 67: 196-200
12. Gemke R, Bonsel G: Comparative assessment of pediatric intensive care: a national multicenter study. *Crit Care Med* 1995; 23: 238-45
13. Cunha F, Mota T, Pinto A, Aparício J, Ribeiro A, Carvalho J *et al*: Avaliação das características da população de uma unidade de cuidados intensivos pediátricos (1995-2000). *Acta Pediatr Port* 2002; 2: 73-9
14. Schnitzler E: Pediatric intensive care in Argentina. *Crit Care Med* 1993; 21 (Suppl): 403-4
15. DiCarlo J, Zaitseva T, Khodateleva T, Belayeva I, Stroganov D, Korobko L: Comparative assessment of pediatric intensive care in Moscow, the Russian federation: a prospective, multicenter study. *Crit Care Med* 1996; 24: 1403-7
16. Bertolini G, Ripamonti D, Cattaneo A, Apolone G: Pediatric risk of mortality: an assessment of its performance in a sample of 26 italian intensive care units. *Crit Care Med* 1998; 26: 1427-32
17. Leclerc F, Leteurtre S, Noizet O, Dorkenoo A, Wirth J, Cremer R *et al*: Can pediatric generic scores calculated four hours after admission be used as inclusion criterion for clinical trials? *Crit Care Med* 2001; 29(Suppl): 177
18. Estrada J, Carmo Vale M, Ramos J, Santos M, Nóbrega S *et al*: Falências de órgão em Pediatria. *Acta Pediatr Port* 1997; 1: 27-33
19. Tibby S, 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
20. Maulen-Radovan I, Castrellon P, Rodriguez R, Natera O: PRISM score evaluation to predict outcome in pediatric patients on admission at an emergency department. *Arch Med Res* 1996; 27: 553-8
21. Jones G, Thorburn K, Tigg A, Murdoch I: Preliminary data: PIM vs PRISM in infants and children post cardiac surgery in a UK PICU. *Intensive Care Med* 2000; 26: 145
22. Pearson G, Stickley J, Shann F: Calibration of the paediatric index of mortality in UK paediatric intensive care units. *Arch Dis Child* 2001; 84: 125-8
23. Gemke R, Vugth A: Scoring systems in pediatric intensive care: PRISM III versus PIM. *Intensive Care Med* 2002; 28: 204-7
24. Shann F: Are we doing a good job: PRISM, PIM and all that. *Intensive Care Med* 2002; 28: 105-7
25. Sanz C, Lucas M, Cid J, Escribano D, Cantarero G: Valor prognóstico de la puntuación PIM (índice pediátrico de mortalidad) y del ácido láctico en niños críticamente enfermos. *An Esp Pediatr* 2002; 57:3 94-400
26. Hatherill M, Waggie Z, Purves L, Reynolds L, Argent A: Mortality and the nature of metabolic acidosis in children with shock. *Intensive Care Med* 2003; 29: 286-91
27. Davis S, Harrison M, Connor J: PIM poorly predicts mortality in a US PICU. *Crit Care Med* 2003; 30(Suppl): 137
28. International Classification of Diseases 9th revision – Clinical Modification 5th edition. Los Angeles: Practice Management Information Corporation; 1998
29. Teasdale G, Jennett B: Assessment of coma and impaired consciousness. A practical scale. *Lancet* 1974; ii: 81-4
30. Reilly P, Simpson D, Sprod R, Thomas L: Assessing the conscious level in infants and young children: a paediatric version of the Glasgow Coma Scale. *Child's Nerv Syst* 1988; 4: 30-3
31. Lemeshow S, Hosmer D: A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982; 115: 92-106
32. Hosmer D, Lemeshow S. Applied Logistic Regression. New York: Wiley and Sons, 1989
33. Hanley J, McNeil B: The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143: 29-36
34. Randolph A, Guyatt G, Calvin J, Doig G, Richardson W: Understanding articles describing clinical prediction tools. *Crit Care Med* 1998; 26: 1603-12
35. European Society of Intensive Care Medicine consensus conference: Predicting outcome in ICU patients. *Intensive Care Med* 1994; 20: 390-7
36. Selker H: Systems for comparing actual and predicted mortality rates: characteristics to promote cooperation in improving hospital care. *Ann Intern Med* 1993; 118: 820-2
37. Hunt J, Meyer A: Predicting survival in the intensive care unit. *Curr Probl Surg* 1997; 34: 527-99
38. Teres D, Lemeshow S: When to customize a severity model. *Intensive Care Med* 1999; 25: 140-2
39. Pearson G: Audit and performance. In: Pearson G editor. Handbook of Paediatric intensive care. Philadelphia: Saunders, 2002: 31-49
40. Estrada J, Ventura L, Vale M, Ramos J, Macedo A, Marques A *et al*: Movimento assistencial numa unidade de cuidados intensivos pediátricos. Análise de dois anos. *Rev Port Pediatr* 1994; 25: 331-6
41. Pollack M, Patel K, Ruttimann U, Cuerdon T: Frequency of variable measurement in 16 pediatric intensive care units: influence on accuracy and potential for bias in severity of illness assessment. *Crit Care Med* 1996; 24: 74-7
42. Boyd O, Grounds R: Physiological scoring systems and audit. *Lancet* 1993; 341: 1573-4
43. Marcin J, Pollack M: Review of the methodologies and applications of scoring systems in neonatal and pediatric intensive care. *Pediatr Crit Care Med* 2000; 1: 20-7
44. Randolph A: Paediatric Index of Mortality (PIM): do we need another paediatric mortality prediction score? *Intensive Care Med* 1997; 23: 141-2
45. Mackenzie S, Kendrick S, Howie J: From severity scores to health gain – a difficult road but one worth travelling. *Curr Opin Crit Care* 2000; 6: 181-6
46. Pollack M, Patel K, Ruttimann U: PRISM III: an updated Pediatric Risk of Mortality score. *Crit Care Med* 1996; 24: 743-52
47. Richardson D, Tarnow-Mordi W, Pollack M: Proprietary products

- and research: the proponents view. *Pediatrics* 1998; 102: 401-3
50. Kanter R: Research publications involving proprietary products: science, news or advertising? *Pediatrics* 1998; 101: 468-70
51. Slater A, Shann F, Pearson G: PIM2 a revised version of the Paediatric Index of Mortality. *Intensive Care Med* 2003; 29: 278-85