

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM PEDIATRIA E SAÚDE DA CRIANÇA

**Validação do Escore Prognóstico Pediatric Index of Mortality (PIM 3) em
uma Unidade de Terapia Intensiva no Brasil**

KARLA LOYOLA DE OLIVEIRA ARANTES

PORTO ALEGRE
2016

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM PEDIATRIA E SAÚDE DA CRIANÇA

**Validação do Escore Prognóstico Pediatric Index of Mortality (PIM 3) em
uma Unidade de Terapia Intensiva no Brasil**

KARLA LOYOLA DE OLIVEIRA ARANTES

PORTO ALEGRE

2016

KARLA LOYOLA DE OLIVEIRA ARANTES

**Validação do Escore Prognóstico Pediatric Index of Mortality (PIM 3) em
uma Unidade de Terapia Intensiva no Brasil**

Dissertação de Mestrado apresentado ao Programa de Pós-Graduação em Pediatria e Saúde da Criança na Pontifícia Universidade Católica Do Rio Grande do Sul para obtenção de título de Mestre em Pediatria.

Orientador: Prof. Doutor Pedro Celiny Ramos Garcia

PORTO ALEGRE

2016

Ficha Catalográfica

A662v Arantes, Karla Loyola de Oliveira

Validação do Escore Prognóstico Pediatric Index of Mortality (PIM 3) em
uma Unidade de Terapia Intensiva no Brasil / Karla Loyola de Oliveira Arantes
. – 2016.

53 f.

Dissertação (Mestrado) – Programa de Pós-Graduação em
Medicina/Pediatria e Saúde da Criança, PUCRS.

Orientador: Prof. Dr. Pedro Celiny Ramos Garcia.

1. Escore prognóstico. 2. Índice pediátrico de mortalidade. 3. PIM. 4. Unidade
de Terapia Intensiva. 5. Criança. I. Garcia, Pedro Celiny Ramos. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da PUCRS
com os dados fornecidos pelo(a) autor(a).

RESUMO

Introdução: Escores prognósticos são ferramentas úteis na avaliação da eficácia de tratamentos, risco de mortalidade e qualidade de serviços, possibilitando comparação entre diferentes UTI (Unidade de Terapia Intensiva), além de implementação e melhorias de protocolos de tratamento e políticas de saúde pública. O PIM (Pediatric Index of Mortality) é atualmente um dos escores prognósticos mais utilizados na pediatria, tendo sido aperfeiçoado gerando o PIM 2 e mais recentemente o PIM 3, este último até o momento não validado em países em desenvolvimento.

Objetivos: Validação do PIM3 em um hospital pediátrico terciário no sudeste do Brasil, e comparação de sua performance com o PIM 2, escore atualmente utilizado.

Métodos: Estudo de coorte histórico retrospectivo, realizado entre 1º de janeiro e 31 de dezembro de 2014 na Unidade de Terapia Intensiva Pediátrica (UTIP) do HEINSG (Hospital Estadual Infantil Nossa Senhora da Glória). A caracterização da amostra foi realizada através da frequência observada, porcentagem, medidas de tendência central e de variabilidade. A calibração dos escores foi analisada pelo teste de ajuste de Hosmer-Lemeshow, enquanto a estatística Z de Flora foi utilizada para avaliar a semelhança entre a mortalidade geral e observada através do índice padronizado de mortalidade (SMR - Standardized Mortality Rate). Para o teste z de Flora, considerou-se valores críticos para não-rejeição da hipótese nula o intervalo de dois desvios padrão (DP) (ou entre $< 1,96$ e $> -1,96$).

A área sob a curva ROC (Receiver Operating Characteristic) foi utilizada para a análise da capacidade de discriminação dos modelos PIM2 e PIM3 entre os pacientes que teriam alta ou evoluiriam para o óbito, e a avaliação da semelhança entre as médias do PIM2 e PIM3 foi feita através do teste t de Student para amostras independentes. A concordância entre as medidas do PIM2 e PIM3 foi avaliada pelo gráfico de Bland & Altman. O nível alfa de significância utilizado nas análises foi de 5% e intervalo de confiança de 95%.

Os dados foram coletados em uma tabela Excel, conferidos em prontuários médicos e posteriormente transferidos para o software IBM SPSS para a realização de todas as análises.

Resultados: Foram admitidos 293 pacientes no período estudado, sendo 38 excluídos por apresentarem um dos critérios de exclusão. Dos 255 pacientes analisados, 35 (13,7%) foram a óbito. O escore médio do PIM2 foi significativamente maior que o do PIM3, e a estatística Z de flora evidenciou não haver diferença entre a mortalidade geral observada e esperada no PIM2, mas há diferença destas no PIM3. O PIM2 obteve uma discriminação excelente (AUC = 0.830) e sua sensibilidade foi de 85.7 e especificidade de 69.1. Já o PIM3 obteve uma discriminação aceitável (AUC = 0.748) e sua sensibilidade foi de 74.3 e especificidade de 67.7. A comparação entre as áreas sob a curva ROC do PIM2 e PIM3 foi significativa ($p = 0.015$), evidenciando que há diferença entre as suas áreas, com melhor desempenho do PIM2 em relação ao PIM 3 (Z de Flora 2.427). O gráfico de Bland-Altman indicou que os limites de 95% de concordância entre as 2 versões do PIM variaram de -1,2 a 2,3, indicando que as medidas são inconsistentes, havendo discordância entre as mesmas que incluem 10,6% valores acima e abaixo do limite $\pm 1,96$ DP, cerca do dobro do tolerável de 5%.

Conclusões: Neste estudo, o PIM 2 apresentou melhores resultados para discriminar aqueles pacientes que irão a óbito, se comparado ao PIM 3. Poderia se sugerir, partindo desses resultados, que fosse mantida a coleta de dados utilizando as 2 versões do escore nesta unidade, para que estes dados possam ser novamente analisados com uma amostra maior, e que esses resultados possam ser comparados com novos estudos que devem surgir em locais com populações com características semelhantes.

Palavras chave: *Escore prognóstico; Índice pediátrico de mortalidade; PIM; Unidade de Terapia Intensiva; Criança.*

ABSTRACT

Introduction: Prognostic scores are useful tools in assessing the effectiveness of treatments, mortality risk and quality of services, allowing the comparison between different Intensive Care Units as the implementation and improvement of treatment and public health policies and protocols. The PIM (Pediatric Index of Mortality) is one of the most widely used prognostic scores in pediatrics and has improved generating PIM 2 and PIM 3. The latest has not been validated in developing countries.

Objectives: Validation of PIM 3 score in a tertiary pediatric hospital in southeastern Brazil, and comparison of its performance with the PIM 2, currently used.

Methods: A contemporary cohort study undertaken between January 1 and December 31, 2014, at the Pediatric Intensive Care Unit of HEINSG (Hospital Estadual Infantil Nossa Senhora da Glória). The sample characterization was performed using the observed frequency, percentage, measures of central tendency and variability. The calibration of the scores was analyzed by the Hosmer-Lemeshow test setting, while the Z statistic Flora was used to evaluate the similarity between overall mortality and the one observed through the standardized mortality rate (SMR - Standardized Mortality Rate). For Flora z test, it is considered critical values for the non-null hypothesis rejected the two standard deviations (SD) (or between <1.96 and > -1.96). The area under the ROC curve (Receiver Operating Characteristic) was used to analyze the discrimination capacity of PIM2 and PIM3 models among patients who were discharged or died, and the assessment of the concordance between the measures of PIM 2 and PIM3 was performed using the Student t test for independent samples. The agreement between the measures of PIM 2 and PIM3 was evaluated by Bland & Altman plot. The significance alpha level used in the analyzes was 5% and 95% confidence interval. Data were collected in an Excel table, confirmed on medical records and later transferred to IBM SPSS software to perform all analyzes.

Results: 293 patients were admitted to the PICU during the studied period, 38 of whom presented exclusion criteria. 35 (13.7%) of the 255 patients studied died. The average score PIM2 was significantly higher than the PIM3, and Flora Z statistics showed no difference between the overall mortality observed and the expected one in PIM2, but this difference was found in PIM3. The PIM2 score got an excellent discrimination (AUC = 0.830) and its sensitivity was 85.7, and the specificity was 69.1. On the other hand, the PIM 3 score had an acceptable discrimination (AUC = 0.748), while its sensitivity was 74.3, and its specificity was 67.7. The comparison between the areas under the ROC curve of PIM2 and PIM3 was significant ($p = 0.015$), showing that there is a difference between their areas, with better performance for PIM2 compared to PIM 3 (Z Flora 2.427). The Bland-Altman diagramme indicated that the 95% limits of concordance between the two versions of PIM ranged from -1.2 to 2.3, indicating that the measures are inconsistent. There is discordance of 10.6% above and below the limit ± 1.96 standard deviations (SD) between the mentioned values, that is about twice the tolerable 5%.

Conclusion: In our study, the PIM 2 shows better results to discriminate those patients who will die. We suggest, based on these results, that data collection should be maintained using the 2 versions of the score in this unit. Than, these data could be reanalyzed with a larger sample, and these results could be compared with new studies conducted in locations where population have similar characteristics.

Key words: Prognostic scores, Pediatric Index of mortality, PIM, Pediatric Intensive Care Unit, children.

LISTA DE FIGURAS

CAPÍTULO III

Figura 1. Área sob a curva ROC para PIM2 e PIM3	38
Figura 2. Gráfico de Bland & Altman para PIM2 e PIM3.	39

LISTA DE TABELAS

CAPÍTULO III

Tabela 1 - Características gerais da amostra	36
Tabela 2 - Teste de Hosmer-Lemeshow por faixa de PIM2, PIM3 e SMR.	37

LISTA DE ABREVIATURAS

APACHE	Acute Physiology and Chronic Health Evaluation – Avaliação do Estado de Saúde Fisiológico Agudo e Crônico
HEINSG	Hospital Estadual Infantil Nossa Senhora da Glória
PELOD	Pediatric Logistic Organ Dysfunction – Logística de Disfunção Orgânica Pediátrica
PIM	Pediatric Index of Mortality – Índice Pediátrico de Mortalidade
PRISM	Pediatric Risk of Mortality – Risco de Mortalidade Pediátrico
PSI	Physiologic Stability Index – Índice de Estabilidade Fisiológica
ROC	Receiver Operating Characteristic – Recebedor Operador de Características
SMR	Standardized Mortality Rate – Índice Padronizado de Mortalidade
UTI	Unidade de Tratamento Intensivo
UTIP	Unidade de Tratamento Intensivo Pediátrico

SUMÁRIO

CAPÍTULO I	9
1.1 APRESENTAÇÃO	10
1.2 JUSTIFICATIVA	12
1.3 OBJETIVOS	13
1.3.1 Objetivo principal.....	13
1.3.2 Objetivos secundários	13
1.4 REFERÊNCIAS.....	14
CAPÍTULO II.....	15
2.1 ARTIGO DE REVISÃO.....	16
CAPÍTULO III	31
3.1 ARTIGO ORIGINAL	32
CAPÍTULO IV.....	43
4.1 CONCLUSÕES	44
ANEXOS	45
ANEXO A – ÍNDICE DE MORTALIDADE PEDIÁTRICA PIM2	46
ANEXO B – ÍNDICE DE MORTALIDADE PEDIÁTRICA PIM3	49
ANEXO C – APROVAÇÃO CEP	51

CAPÍTULO I

1.1 APRESENTAÇÃO

Desde a criação das Unidades de Tratamento Intensivo (UTI) gerais, e o posterior desmembramento desses serviços em outros específicos para tratamento crítico do paciente adulto, da criança e do neonato, são desenvolvidas ferramentas para uma avaliação mais objetiva da eficácia do tratamento, do risco de mortalidade, e da qualidade do serviço, de modo que esses dados possam ser comparados entre UTI de países diversos, com populações com características próprias, diferentes diagnósticos e gravidade¹.

Essas ferramentas surgiram a partir da necessidade de uma mudança de paradigma em relação à avaliação subjetiva que era feita até então dos pacientes que internavam em UTI, baseada, sobretudo, na experiência dos profissionais médicos e de enfermagem em realizar prognóstico principalmente relacionado ao desfecho óbito/alta. Fez-se necessário, em tempos de facilidade de troca de informações entre serviços e comparação de dados, o estabelecimento de critérios objetivos, que permitissem não só a comparação entre grupos de UTI, após ajuste de gravidade e diagnóstico, como a avaliação de um mesmo serviço em diferentes períodos. Dessa forma, poderiam ser obtidos dados essenciais para a verificação de desempenho e padrão de funcionamento, gerando substrato para a melhoria da qualidade da assistência e orientação à implementação e melhorias em protocolos de tratamento e políticas de saúde pública¹.

Na busca da melhor forma de obter dados objetivos e comparáveis, foram desenvolvidos diversos escores de prognóstico, de mortalidade, de desfecho e de intervenção, inicialmente de forma geral, sendo o primeiro deles o APACHE (Acute Physiology and Chronic Health Evaluation), em 1981². Posteriormente surgiram modelos específicos para diferentes grupos de pacientes, que se apresentavam de forma muito diversa em relação as taxas de mortalidade devido a características próprias do grupo, como por exemplo as crianças e os neonatos^{3,4,5,6,7,8}.

Apesar de, atualmente, serem amplamente aceitos pelos gestores, os escores prognósticos, e a avaliação de risco de óbito baseados em ferramentas estatísticas, ainda encontram resistência do médico assistente que trabalha em UTI: há médicos que ainda acreditam que a avaliação clínica experiente é superior a qualquer ferramenta estatística em

relação ao prognóstico questionando portanto a utilidade e acurácia desses modelos preditivos, e demonstrando que se mantem – até os dias atuais – a desconfiança dos médicos em relação ao uso da matemática como parte de seu arsenal.

Com a mesma resistência, que nos acompanha desde Hipócrates, em aceitar a matemática e a estatística como ferramentas prognósticas, os escores encontraram tais questionamentos à sua utilização – que substituiria de forma padronizada e com critérios objetivos o “olho clínico”, altamente dependente da experiência do médico avaliador e de sua capacidade de utilizar a memória com quadros semelhantes vistos no passado para fazer prognósticos.

Neste estudo, nos debruçamos inicialmente sobre o assunto em uma revisão narrativa que visa mostrar, sob um prisma histórico, os caminhos percorridos até a aceitação da matemática e da estatística pela medicina, e de seus usos derivados – como por exemplo os escores preditivos – hoje inseparáveis ferramentas do bom funcionamento de qualquer unidade de Terapia Intensiva.

Descreveremos ainda, de forma breve, o aparecimento dos primeiros escores prognósticos, e posteriormente, e em maior detalhe, a criação dos escores específicos para a Pediatria, culminando com a descrição do objeto deste estudo, o PIM⁵ (Pediatric Index of Mortality), desde sua primeira versão até a atual, PIM 3⁹.

Apresentaremos, por fim, por meio do artigo original, a pesquisa realizada em uma UTIP de Hospital Pediátrico terciário no Sudeste do Brasil, visando avaliar o desempenho da predição do risco de mortalidade pelo PIM3, através da comparação entre óbitos esperados e os realmente observados, considerando as atuais características epidemiológicas da população em estudo, e a comparação entre a performance da nova versão do escore com aquela do PIM 2, utilizado nesta UTIP desde 2005, avaliando possíveis diferenças entre suas médias, objetivando-se a validação da atualização do PIM no serviço estudado.

1.2 JUSTIFICATIVA

Aos serviços de saúde, sobretudo aos de Cuidados Intensivos por suas características próprias, impõe-se a necessidade do estabelecimento de prognósticos corretos e fundamentados em critérios cientificamente comprovados, de forma a basear escolhas eticamente adequadas tanto para o diagnóstico quanto para o tratamento, evitando o uso de recursos sofisticados e custosos em pacientes que poderiam ser tratados com recursos mais simples, assim como o uso de procedimentos e exames diagnósticos dolorosos em paciente sem chances de recuperação da doença, respeitando os princípios de maleficência/beneficência. Aliada a esta, está a necessidade de fornecer adequada orientação ao paciente e sua família, de modo a garantir o princípio da autonomia (que exige, sempre, informação adequada).

Por fim, embora não menos importante, está a exigência legal, estabelecida pela RDC nº 7 da ANVISA, que em sua seção IX sobre Avaliação cita, no artigo 48, que:

Devem ser monitorados e mantidos registros de avaliações do desempenho e do padrão de funcionamento global da UTI, assim como de eventos que possam indicar necessidade de melhoria da qualidade da assistência, com o objetivo de estabelecer medidas de controle ou redução dos mesmos.

§ 1º Deve ser calculado o Índice de Gravidade / Índice Prognóstico dos pacientes internados na UTI por meio de um Sistema de Classificação de Severidade de Doença recomendado por literatura científica especializada.

Justificando assim estudos que, como este, propõem a verificação da adequação dos escores prognósticos aos serviços nos quais são utilizados, assim como de suas modificações e atualizações.

1.3 OBJETIVOS

1.3.1 Objetivo principal

Avaliar a capacidade de predição da mortalidade do escore prognóstico PIM após sua última atualização, PIM3, em um hospital pediátrico terciário no sudeste do Brasil, objetivando sua validação no serviço estudado.

1.3.2 Objetivos secundários

- Comparar a capacidade de predição de mortalidade do PIM3, com o índice atualmente utilizado na UTIP em questão (PIM2);
 - Determinar a capacidade de discriminação e calibração dos índices PIM2 e PIM3 na UTIP estudada;
 - Medir os desfechos mortalidade, tempo de internação e tempo de ventilação, de acordo com as variáveis de caracterização epidemiológica.
-
-

1.4 REFERÊNCIAS

1. 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 Sep 28;272(12):941-6
 2. Knaus WA, Zimmerman JE, Wagner DP, Draper EA, Lawrence DE. APACHE-acute physiology and chronic health evaluation: a physiologically based classification system. *Crit Care Med*. 1981 Aug;9(8):591-7.
 3. Pollack MM, Ruttimann UE, Getson PR. Pediatric risk of mortality (PRISM) score. *Crit Care Med*. 1988; Nov; 16(11):1110-6.
 4. Pollack MM, Patel KM, Ruttimann EU, PRISM III: an updated Pediatric Risk of Mortality score; *Crit Care Med*, 1996 May;24(5):743-52
 5. Shann F, Pearson G, Slater A, et al. Paediatric index of mortality (PIM): a mortality prediction model for children in intensive care. *Intensive Care Med*. 1997; 23:201–207.
 6. 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; 29:278–285
 7. Leteurtre S, Martinot A, Duhamel A, Proulx F, Grandbastien B, Cotting J, et al. Validation of the paediatric logistic organ dysfunction (PELOD) score: prospective, observational, multicentre study. *Lancet*. 2003 Jul 19;362(9379):192-7.
 8. Richardson DK, Gray JE, McCormick MC, Workman K, Goldmann DA. Score for Neonatal Acute Physiology: a physiologic severity index for neonatal intensive care. *Pediatrics* 1993;91:617-23.
 9. Straney L, Clements A, Parslow RC, Pearson G, Shann F, Alexander J, Slater A. ANZICS Paediatric study group and the paediatric intensive care audit network paediatric index of mortality 3: an updated model for predicting mortality in pediatric intensive care. *Pediatr. Crit Care Med*. 2013 Sep; 14(7):673-81
-