

Mark T. Keegan¹, Marcio Soares²

O que todo intensivista deveria saber sobre os sistemas de escore prognóstico e mortalidade ajustada ao risco

What every intensivist should know about prognostic scoring systems and risk-adjusted mortality

1. Department of Anesthesiology, Division of Critical Care, Mayo Clinic - Rochester, MN, USA.
2. Departamento de Terapia Intensiva, Instituto D'Or de Pesquisa e Ensino - Rio de Janeiro (RJ), Brasil.

Na prática clínica, muitos escores e sistemas de prognóstico foram desenvolvidos para quantificar a gravidade da doença, avaliar seu prognóstico e direcionar as intervenções terapêuticas. Os sistemas Escala de Coma de Glasgow, Modelo para Doença Hepática Terminal (MELD - *Model for End Stage Liver Disease*) e *American Society of Anesthesiologists Physical Status Classification* são alguns exemplos. A heterogeneidade na prática clínica da terapia intensiva, o elevado custo dos cuidados, a chance muito real de óbito nas unidades de terapia intensiva (UTI) e o desejo de fazer comparações entre UTI desencadearam o desenvolvimento e o refinamento de sistemas prognósticos específicos para a UTI.⁽¹⁻⁷⁾

Os sistemas de escore podem ser genéricos ou específicos para determinadas enfermidades e ser utilizados para análise de uma coorte ou avaliação de pacientes individuais; podem se basear em desarranjos fisiológicos ou alocação de recursos, e ser simples ou complexos. Existem, na prática clínica em terapia intensiva, duas principais categorias de sistemas de escore. Os escores de disfunção orgânica (por exemplo, o *Sequential Organ Failure Assessment*, SOFA) descrevem os desarranjos fisiológicos por sistema de órgãos para fornecer uma avaliação objetiva da extensão e da gravidade da disfunção dos órgãos. A outra categoria principal é o modelo prognóstico de gravidade da doença, uma discussão que ocupa a maior parte deste comentário. Estes sistemas (por exemplo, o *Acute Physiology and Chronic Health Evaluation*, APACHE) utilizam dados fisiológicos, condições mórbidas prévias e informações relativas à natureza da doença atual para prever a probabilidade de óbito.

***Sequential Organ Failure Assessment*: um escore de disfunção de órgãos**

A síndrome de disfunção múltipla de órgãos é uma importante causa de morbidade e mortalidade na UTI. A extensão e a gravidade da disfunção dos órgãos podem ser quantificadas em uma série de escores de disfunção de órgãos, sendo que o mais proeminente deles é o SOFA.⁽⁸⁾ Originalmente delineado para uso em pacientes com sepse, o SOFA é atualmente utilizado em todos os grupos de pacientes. Os escores diários podem ser calculados e utilizados para descrever o grau de disfunção de órgãos durante a permanência de um paciente na UTI - em contraste com sistemas prognósticos genéricos, delineados para dar uma previsão exclusivamente com base no primeiro dia na UTI. Dependendo do grau de comprometimento, são atribuídos escores entre zero e 4 a cada um dos sistemas: cardiovascular, respiratório, hepático, hematológico, neurológico e renal. Estes escores são somados para proporcionar um escore SOFA total e não foram originalmente propostos para prever a mortalidade, porém escores absolutos altos

Conflito de interesse: O Dr. Soares é fundador e acionista da Epimed Solutions®, empresa que comercializa o software Epimed Monitor®, que se destina à gestão da qualidade e à comparação de unidades de terapia intensiva.

Submetido em 30 de abril de 2016

Aceito em 4 de maio de 2016

Autor correspondente:

Mark T. Keegan
Department of Anesthesiology of the Mayo Clinic
200 First Street SW
Charlton 1145
Rochester
MN 55905
E-mail: keegan.mark@mayo.edu

Editor responsável: Jorge Ibrain de Figueira Salluh

DOI: 10.5935/0103-507X.20160052

e aumento de um escore durante as primeiras 96 horas na UTI se associam com maior risco de óbito.⁽⁹⁾

Sistemas de escore prognóstico para gravidade da doença na unidade de terapia intensiva (por exemplo APACHE, SAPS, MPM)

Há três principais sistemas de escore prognóstico genérico na UTI: o APACHE, o *Simplified Acute Physiology Score* (SAPS) e o *Mortality Probability Model* (MPM).⁽¹⁰⁻¹³⁾ Após sua primeira descrição nos anos 1980, com o tempo os modelos foram atualizados e suas versões mais recentes são o APACHE IV, SAPS 3 e MPM III. A tabela 1 fornece detalhes das versões mais recentes. Os sistemas utilizam dados de fisiologia aguda, diagnóstico agudo, condições crônicas de saúde e características do índice quando da internação à UTI, para preverem a mortalidade hospitalar. O SAPS e o MPM₀ são calculados a partir dos dados disponíveis dentro de 1 hora da internação à UTI e refletem a gravidade da doença quando da internação. Os escores APACHE e MPM₂₄ são calculados a partir dos dados disponíveis dentro das primeiras 24 horas da internação à UTI. O APACHE é mais complicado - e mais preciso - do que o SAPS, que, por sua vez é mais complicado e mais preciso do que o MPM.⁽¹⁴⁾ Diversos outros sistemas prognósticos foram desenvolvidos em áreas geográficas específicas (por exemplo, *Intensive Care National Audit & Research Centre* - ICNARC, no Reino Unido, e *Australian and New Zealand Risk of Death* - ANZROD, na Austrália e Nova Zelândia).^(15,16)

Metodologia para o desenvolvimento de sistemas de escore prognóstico

O desfecho de interesse (a variável dependente) geralmente selecionada é a mortalidade - mortalidade na UTI, mortalidade hospitalar ou mortalidade após 28 dias, por exemplo.^(17,18) A natureza binária (ou dicotômica) desta variável permite o uso de técnicas de regressão logística no desenvolvimento do modelo, embora outras técnicas também possam ser utilizadas, inclusive a análise Bayesiana, a regressão binária de Cox e redes neurais. As variáveis preditoras (independentes), que devem estar rotineiramente disponíveis, existir independentemente da intervenção e ser confiáveis, são escolhidas com base em técnicas analíticas ou (menos comumente) por um consenso de especialistas. Tais variáveis incluem, por exemplo, idade, pH e presença ou ausência de cirrose. Em uma análise de regressão logística, as variáveis preditoras (denotadas como X) são ligadas à probabilidade de óbito segundo uma série de coeficientes (denotados como β) como segue:

$$\text{Logaritmo natural (ln) das chances de óbito} = \beta_0 + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3) + \dots + \beta_n(X_n)$$

A razão de chances (*odds ratio*) de óbito para cada aumento unitário na variável X_n é e^{β_n} . Os coeficientes da função logística são determinados por análise estatística, em que os sinais e magnitudes proporcionam indicações das direções e potências das associações, respectivamente. Para prever variáveis não dicotômicas (por exemplo, tempo de permanência na UTI), utiliza-se um modelo de regressão linear. A inclusão de um número maior de variáveis no modelo geralmente proporciona uma melhor precisão preditiva, porém ao custo de um aumento da carga de trabalho para a coleta de dados. No modelo, deve haver pelo menos dez “eventos” (óbitos) para cada variável preditor. Idealmente deseja-se um modelo altamente preciso, porém parcimonioso, com poucas variáveis preditoras. Preferem-se os modelos desenvolvidos em grandes bases de dados de múltiplas instituições e validados em conjuntos de dados em separado.

Avaliação do desempenho do modelo

O desempenho do modelo deve ser avaliado por meio da avaliação da discriminação e calibração. Discriminação quantifica a acurácia de uma dada predição. Por exemplo, se um sistema prognóstico prevê uma mortalidade de 80% para uma coorte de pacientes com um determinado escore APACHE IV, a discriminação é perfeita se morrerem 80% dos pacientes dessa coorte. A área sob a curva (AUC) *receiver operating characteristic* (ROC) é utilizada para medir a discriminação.⁽¹⁹⁾ Uma discriminação perfeita dá uma AUC de 1; uma AUC de 0,5 significa que a predição do modelo não é melhor do que o acaso.

Calibração é a mensuração de quão bom é o desempenho de um determinado modelo em uma variedade de mortalidades previstas. Esta avaliação é feita com utilização do teste estatístico Hosmer-Lemeshow *goodness of fit*, que é calculado pelo agrupamento de uma coorte de pacientes em decis de risco previsto e comparação das mortalidades observada e prevista ao longo dos decis para fornecer uma estatística de qui quadrado. Um valor de p superior a 0,05 (isto é, não significativo) implica boa calibração. O teste de Hosmer-Lemeshow é afetado pelo tamanho da amostra.⁽²⁰⁾ Alguns investigadores propuseram o cinto de calibração como um método alternativo para avaliar a calibração do modelo.⁽²¹⁾ O cinto de calibração (*calibration belt*) utiliza uma função logística polinomial generalizada entre o desfecho e a transformação logística da probabilidade estimada, o que fornece informações a respeito da direção,

Tabela 1 - Características do estudo, desempenho e variáveis das versões mais recentes dos mais proeminentes modelos prognósticos

	APACHE IV	SAPS 3	MPM ₀ III
População do estudo	110.558	16.784	124.855
Período do estudo	1º de janeiro de 2002 a 31 de dezembro 2003	14 de outubro a 15 de dezembro de 2002	Outubro de 2001 a março de 2004
Regiões geográficas	Estados Unidos	35 países, 5 continentes	Estados Unidos
Número de UTI	104	303	135
Número de hospitais	45	281	98
Variáveis no modelo	142	20	16
Ocasão da coleta dos dados	Primeiras 24 horas da internação à UTI	± 1 hora após a internação à UTI	Dentro de 1 hora da internação à UTI
AUC	0,88	0,848	0,823
Estatística C de HL	16,9	14,29	11,62 (abs 10,94)
SMR	0,997	1,0	1,018
Variáveis do preditor			
Idade	Sim	Sim	Sim
Tempo de permanência no hospital antes da internação à UTI	Sim	Sim	Não
Tipo de internação à UTI	Sim	Sim	Sim
Origem da internação à UTI	8	3	Não
Comorbidades crônicas	7	6	3
Condição de ressuscitação	Não	Não	Sim
RCP antes da internação à UTI	Não	Não	Sim
Razões para internação à UTI/diagnóstico agudo	116	10	5
Situação cirúrgica quando da internação à UTI	Sim	Sim	Não
Região anatômica da cirurgia	Não	5	Não
Infecção aguda na internação à UTI	Não	Sim	Não
Tratamento com fármacos vasoativos	Sim	Sim	Não
Ventilação mecânica	Sim	Sim	Sim
Variáveis fisiológicas clínicas	6	4	3
Variáveis fisiológicas laboratoriais	10	6	0

APACHE - *Acute Physiology and Chronic Health Evaluation*; SAPS - *Simplified Acute Physiology Score*; MPM - *Mortality Probability Model*; UTI - unidade de terapia intensiva; AUC - área sob a curva; HL - Hosmer-Lemeshow; SMR - *standardized mortality ratio*; RCP - ressuscitação cardiopulmonar. Fonte: Afessa B, Gajic O, Keegan MT. Severity of illness and organ failure assessment in adult intensive care units. *Crit Care Clin.* 2007;23:639-58.

extensão e classes de risco afetadas pelos desvios entre a mortalidade observada e a prevista (Figura 1).

Precisão refere-se à diferença entre os desfechos previstos e observados no nível dos pacientes individuais, e pode ser medido pelo escore de Brier, que avalia o desvio quadrado médio entre as probabilidades previstas para um conjunto de eventos e seus desfechos.

Customização

As mudanças na prática clínica e na mescla de casos (*case-mix*), com o tempo, levam a uma deterioração do desempenho prognóstico.⁽²²⁾ Além disso, os modelos podem ter um desempenho aquém do ideal em certas regiões geográficas e populações de pacientes. A customização é

o processo pelo qual um modelo é modificado para melhorar sua precisão, tanto por meio da modificação dos coeficientes na equação (customização de primeiro nível) ou modificação das variáveis (customização de segundo nível). Por exemplo, a descrição original do SAPS 3 proporciona equações customizadas para diferentes regiões geográficas (por exemplo: América do Sul, Europa Oriental) para melhorar seu desempenho nessas regiões. A capacidade de comparar diferentes UTI, instituições ou países fica comprometida se forem utilizados em cada localização modelos customizados, já que a customização limita a validade externa do modelo. Apesar disso, a técnica é uma alternativa atraente quando comparada ao oneroso processo de desenvolver e validar um novo modelo prognóstico.

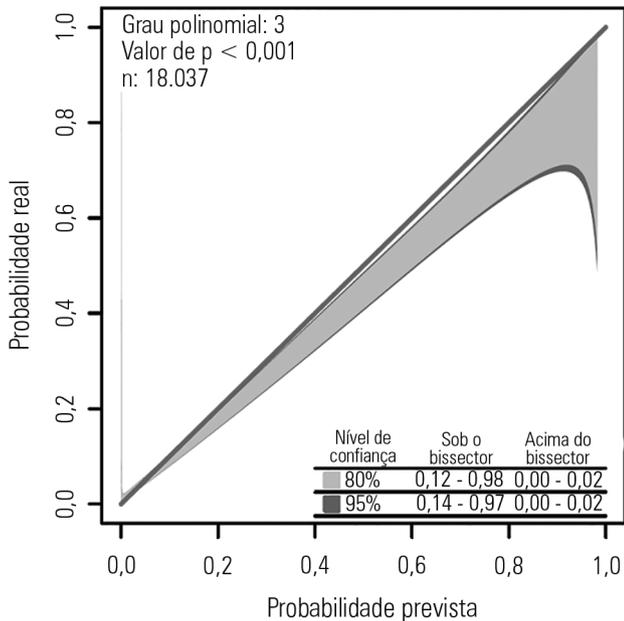


Figura 1 - Cintão de calibração para o *Simplified Acute Physiology Score 3*. Taxas de mortalidade com seus intervalos de confiança de 95%. O grau polinomial, resultados da estatística de Wald e o número de pacientes são fornecidos no quadrante superior esquerdo. No quadrante inferior direito, são relatados os momentos em que o cintão de calibração se desvia significativamente do bissektor, utilizando níveis de confiança de 80% (área cinza claro) e 95% (área cinza escuro). Cortesia: Dr. Pedro Brasil.

Razões padronizadas

A razão da mortalidade observada em relação à mortalidade prevista pelo sistema de escore prognóstico de uma coorte de pacientes é a razão padronizada de mortalidade (SMR - *standardized mortality ratio*) e deve ser relatada com um intervalo de confiança de 95% (IC95%). A SMR é amplamente utilizada para avaliar o desempenho, uma vez que a mortalidade é a medida de desfecho mais objetiva e não é propensa a erro. Se o IC95% não incluir 1, SMR abaixo de 1 e acima de 1 são consideradas, respectivamente, como indicadores de bom ou mau desempenho. No entanto, a SMR não é uma medida perfeita; além das diferenças em termos da qualidade do cuidado, ela pode ser influenciada pela precisão do modelo prognóstico, artefatos relacionados à coleta ou análise de dados, mescla de casos, tendências ligadas ao tempo inicial e confiabilidade entre diferentes avaliadores. Outras medidas de desfecho (por exemplo, duração da ventilação mecânica e tempo de permanência na UTI) também são apropriadas para cálculo das razões entre valores observados sobre previstos. Especificamente para o uso de recursos (utilizando tempo de permanência na UTI como substituto), alguns investigadores utilizaram abordagens diferentes para avaliar o

uso de recursos padronizado ajustado à gravidade (SRU - *severity-adjusted resource use*) para cada UTI individual. Neste caso, o SRU estima a quantidade média de recursos utilizados por paciente sobrevivente em uma UTI específica.⁽²³⁾

Usos de sistemas de escore prognóstico

As taxas de mortalidade, ajustadas com base nas predições de mortalidade fornecidas por sistemas de escore prognóstico, vêm sendo cada vez mais utilizadas para comparar a qualidade do cuidado fornecido por diferentes UTI e hospitais. Estas “taxas de mortalidade ajustadas pela gravidade” podem ser utilizadas para uma “comparação” com instituições similares ou reconhecidas como de alto desempenho, com a finalidade de identificar as deficiências da instituição no desfecho clínico e enfatizar áreas passíveis de melhoria. Fontes pagadoras podem utilizar as taxas de mortalidade ajustadas pela gravidade como um critério para escolha dos provedores de saúde, e os dados de desempenho podem facilitar o processo de acreditação por agências externas. Dentro da mesma organização, podem ser feitas comparações entre o cuidado fornecido por diferentes UTI, assim como o desempenho de uma mesma UTI ao longo do tempo pode ser avaliado para enfatizar as alterações na qualidade de cuidado. Os sistemas prognósticos podem funcionar como ferramentas para avaliar o impacto de novas opções terapêuticas ou modificações organizacionais, como parte de iniciativas de melhoria da qualidade. Quando se considera o ponto de vista de recursos, esses sistemas podem ajudar a identificar uma coorte de pacientes de UTI com baixo risco de mortalidade que poderiam ser cuidados em ambiente diferente da UTI, como uma progressão para uma unidade de cuidados semi-intensivos ou intermediária e pode ainda ajudar nas discussões sobre tomadas de decisão referentes ao final de vida.^(24,25) Os modelos prognósticos podem ajudar a responder questões de pesquisa de resultados referentes a UTI e na estratificação do risco de pacientes para entrada em estudos clínicos, embora essa última abordagem seja controversa em razão da complexidade do cálculo, tempo oportuno e variabilidade entre observadores.⁽²⁶⁾ Estudos concluídos podem ser submetidos a análise *post-hoc* com uso de estratificação do risco dos subgrupos, levando à geração de novas hipóteses.

Limitações do uso de modelos prognósticos para dar suporte à decisão clínica

Apesar da existência de numerosos exemplos de uso de modelos prognósticos para tomada de decisões no caso de

pacientes individuais (por exemplo, utilização do escore MELD para alocação de órgãos para transplante hepático), seu uso não é isento de problemas.⁽²⁷⁾ Os sistemas de escore prognóstico têm melhor desempenho quando utilizados em coortes. Por exemplo, em uma coorte de 1.000 pacientes com mortalidade prevista de 90%, em média 100 pacientes sobrevivem - apesar de uma mortalidade prevista de 90% para qualquer paciente individual. Estes pacientes não prejudicam a validade do modelo, antes a confirmam. Além da incerteza inerente ao que se refere à predição em pacientes individuais, até mesmo os melhores modelos clinicamente úteis têm AUC de não mais do que 0,9, o que implica imperfeição até mesmo para predição para uma coorte. Mais ainda, o desempenho do modelo pode ser prejudicado pela não disponibilidade de todos os dados necessários para o cálculo do escore - os dados faltantes são contabilizados como normais - e pelos erros na coleta e inserção dos dados, assim como pelas preferências do paciente quanto ao suporte à vida.^(14,28) As barreiras à aceitação disseminada dos modelos prognósticos incluem o custo da infraestrutura de tecnologia da informação, que é necessária para aquisição dos dados para modelos complexos, resistência dos médicos em razão de sua percepção

de superioridade de suas próprias estimativas da sobrevivência do paciente, ou sua desatenção quanto à relevância do modelo para seus pacientes, e o foco em predição de mortalidade, em vez de um desfecho funcional, como anos com qualidade de vida.

O futuro

Esperam-se versões atualizadas dos principais sistemas prognósticos, que são bem-vindas. No entanto, seriam potencialmente mais utilizáveis modelos inovadores que possam ser derivados de avanços tornados possíveis pela era dos *big data*, que incluem algoritmos de aprendizado da máquina (*machine learning*) e reavaliações dinâmicas das predições de desfecho.^(29,30) A implantação disseminada de registros médicos em formato eletrônico, juntamente de técnicas pioneiras de análise de *big data* nas vendas e setor bancário, pode, em última análise, permitir que se obtenham prognósticos confiáveis, bem apresentados e em nível do paciente dos desfechos funcionais. Além do mais, podemos estar próximos de um cenário no qual os médicos devem crer nessas previsões e aceitar estratégias geradas por computador de mitigação do risco ou “correção do curso”.

REFERÊNCIAS

- Afessa B, Gajic O, Keegan MT. Severity of illness and organ failure assessment in adult intensive care units. *Crit Care Clin.* 2007;23(3):639-58.
- Keegan MT, Gajic O, Afessa B. Severity of illness scoring systems in the intensive care unit. *Crit Care Med.* 2011;39(1):163-9.
- Zimmerman JE, Kramer AA. A history of outcome prediction in the ICU. *Curr Opin Crit Care.* 2014;20(5):550-6.
- Vincent JL, Moreno R. Clinical review: scoring systems in the critically ill. *Crit Care.* 2010;14(2):207.
- Breslow MJ, Badawi O. Severity scoring in the critically ill: part 1--interpretation and accuracy of outcome prediction scoring systems. *Chest.* 2012;141(1):245-52.
- Breslow MJ, Badawi O. Severity scoring in the critically ill: part 2: maximizing value from outcome prediction scoring systems. *Chest.* 2012;141(2):518-27.
- Salluh JI, Soares M. ICU severity of illness scores: APACHE, SAPS and MPM. *Curr Opin Crit Care.* 2014;20(5):557-65.
- Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med.* 1996;22(7):707-10.
- Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. Serial evaluation of the SOFA score to predict outcome in critically ill patients. *JAMA.* 2001;286(14):1754-8.
- Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit Care Med.* 2006;34(5):1297-310.
- Metnitz PG, Moreno RP, Almeida E, Jordan B, Bauer P, Campos RA, Iapichino G, Edbrooke D, Capuzzo M, Le Gall JR; SAPS 3 Investigators. SAPS 3--From evaluation of the patient to evaluation of the intensive care unit. Part 1: Objectives, methods and cohort description. *Intensive Care Med.* 2005;31(10):1336-44.
- Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, Iapichino G, Edbrooke D, Capuzzo M, Le Gall JR; SAPS 3 Investigators. 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(10):1345-55. Erratum in: *Intensive Care Med.* 2006;32(5):796.
- Higgins T, Teres D, Nathanson B, Copes W, Stark M, Kramer A. Updated Mortality Probability Model - MPM0-III. *Chest.* 2005;128(4 Suppl):348S.
- Keegan MT, Gajic O, Afessa B. Comparison of APACHE III, APACHE IV, SAPS 3, and MPM0III and influence of resuscitation status on model performance. *Chest.* 2012;142(4):851-8.
- Harrison DA, Parry GJ, Carpenter JR, Short A, Rowan K. A new risk prediction model for critical care: the Intensive Care National Audit & Research Centre (ICNARC) model. *Crit Care Med.* 2007;35(4):1091-8.
- Pilcher D, Paul E, Bailey M, Huckson S. The Australian and New Zealand Risk of Death (ANZROD) model: getting mortality prediction right for intensive care units. *Crit Care Resusc.* 2014;16(1):3-4.
- Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ.* 2009;338:b604.
- Ohno-Machado L, Resnic FS, Matheny ME. Prognosis in critical care. *Annu Rev Biomed Eng.* 2006;8:567-99.
- Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143(1):29-36.

20. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. *Crit Care Med.* 2007;35(9):2052-6.
21. Finazzi S, Poole D, Luciani D, Cogo PE, Bertolini G. Calibration belt for quality-of-care assessment based on dichotomous outcomes. *PLoS One.* 2011;6(2):e16110.
22. Teres D, Lemeshow S. When to customize a severity model. *Intensive Care Med.* 1999;25(2):140-2.
23. Rothen HU, Stricker K, Einfalt J, Bauer P, Metnitz PG, Moreno RP, et al. Variability in outcome and resource use in intensive care units. *Intensive Care Med.* 2007;33(8):1329-36.
24. Zimmerman JE, Wagner DP, Knaus WA, Williams JF, Kolakowski D, Draper EA. The use of risk predictions to identify candidates for intermediate care units. Implications for intensive care utilization and cost. *Chest.* 1995;108(2):490-9.
25. Barnato AE, Angus DC. Value and role of intensive care unit outcome prediction models in end-of-life decision making. *Crit Care Clin.* 2004;20(3):345-62, vii-viii.
26. Vincent JL, Opal SM, Marshall JC. Ten reasons why we should NOT use severity scores as entry criteria for clinical trials or in our treatment decisions. *Crit Care Med.* 2010;38(1):283-7.
27. Teres D, Lemeshow S. Why severity models should be used with caution. *Crit Care Clin.* 1994;10(1):93-110; discussion 111-5.
28. Afessa B, Keegan MT, Gajic O, Hubmayr RD, Peters SG. The influence of missing components of the Acute Physiology Score of APACHE III on the measurement of ICU performance. *Intensive Care Med.* 2005;31(11):1537-43.
29. Chen LM, Kennedy EH, Sales A, Hofer TP. Use of health IT for higher-value critical care. *N Engl J Med.* 2013;368(7):594-7.
30. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst.* 2014;2:3.