

# Evaluación del desempeño del APACHE II y SAPS III, en una unidad de cuidados intensivos

## Performance evaluation of APACHE II and SAPS III in an intensive care unit

Sandra Alvear-Vega y Jorge Canteros-Gatica

Recibido 6 septiembre 2016 / Enviado para modificación 02 septiembre 2017 / Aceptado 18 febrero 2018

### RESUMEN

**Objetivo** El objetivo principal de este trabajo, fue validar y comparar la capacidad predictiva de mortalidad de los indicadores de gravedad APACHE II (Acute Physiology and Chronic Health Evaluation, Score II) y SAPS III (Simplified Acute Physiology, Score III) en una muestra de pacientes admitidos en la Unidad de Cuidados Intensivos Adulto (UCI) del Hospital de Curicó, entre los años 2011 y 2013.

**Materiales y Métodos** Estudio analítico, observacional de cohorte histórica de casos consecutivos desde la admisión a la UCI hasta el egreso hospitalario. Para el análisis, se usó el Modelo de Regresión Logística Binaria. De un total de 1 042 pacientes ingresados a la UCI, se incluyó a 793 pacientes sobrevivientes, y a 249 pacientes fallecidos, que representaban el 76,1% y 23,9% respectivamente, del total.

**Resultados** El SAPS III presenta mejor capacidad predictiva que el APACHE II, según el área bajo la curva de características operativas del receptor 0,81 y 0,80 respectivamente. La sensibilidad para el modelo SAPS III es 0,95 y para APACHE II es 0,93. El índice de especificidad es 0,3 para el SAPS III y 0,4 para el APACHE II, con probabilidad superior a 0,5.

**Conclusión** Los indicadores de predicción de mortalidad en UCI; APACHE II y SAPS III tienen una buena capacidad predictiva general, pero ambos indicadores presentan una baja especificidad.

**Palabras Clave:** Cuidados críticos; unidad de cuidados intensivos; monitoreo fisiológico (*fuentes: DeCS, BIREME*).

### ABSTRACT

**Objective** The main objective of this work was to validate and compare the predictive capacity of mortality of the severity score systems APACHE II (Acute Physiology and Chronic Health Evaluation, Score II) and SAPS III (Simplified Acute Physiology, Score III) in a sample of patients admitted to the Adult Intensive Care Unit (ICU) of the Hospital de Curicó between 2011 and 2013.

**Materials and Methods** Analytical, observational, retrospective cohort study of consecutive cases since admission to the ICU until hospital discharge. A binary logistic regression model was used for the analysis. Out of 1 042 patients admitted to the ICU, 793 surviving patients and 249 deceased patients were included, representing 76.1% and 23.9%, respectively, of the total sample.

**Results** The SAPS III score has a better predictive capacity than the APACHE II, according to the area under the curve and the receiver operating characteristic curve: 0.81 and 0.80, respectively. Sensitivity for the SAPS III model was 0.95 and for APACHE II was 0.93. The specificity index was 0.3 for SAPS III and 0.4 for APACHE II, with a probability above 0.5.

**Conclusion** APACHE II and SAPS III, as ICU mortality prediction indicators, have a good predictive power but low specificity.

**Key Words:** Critical care; intensive care units; patient monitoring (*source: MeSH, NML*).

SA: Contador Público y Auditor. Ph. D. Ciencias de l'Education. Universidad de Talca. Talca, Chile. [salvear@utalca.cl](mailto:salvear@utalca.cl)  
JC: MD. Internista. Hospital de Curicó. Curicó, Chile. [jcanteros@ssmaule.cl](mailto:jcanteros@ssmaule.cl)

La mortalidad en las Unidades de Cuidados Intensivos (UCIs), es una preocupación constante de los sistemas sanitarios, por diversas razones, entre ellas; la variedad de patologías que comprometen la vida de los pacientes, la aplicación de técnicas terapéuticas y de monitoreo invasivas, recursos limitados y el alto costo de los procesos que allí ocurren.

Por ello, en dichas unidades, es importante disponer de información que facilite la toma de decisiones, priorizando el ingreso de pacientes que se puedan ver beneficiado de las competencias y recursos disponibles (1).

La medición de la gravedad de los pacientes que ingresan a las Unidades de Cuidados Intensivos es una práctica que permite estratificar pacientes asignándoles un puntaje proporcional a la gravedad de su enfermedad, y con ello facilitar la toma de decisiones al interior de la UCI y comparar, entre ellas, su desempeños (2-3).

Existen diversos sistemas que permiten medir la gravedad de un paciente y con ello predecir la probabilidad de mortalidad de los pacientes admitidos en una Unidad de Cuidados Intensivo, entre ellos; Acute Physiology and Chronic Health Evaluation (APACHE), en sus versiones I, II, III y IV (4-7). También, Simplified Acute Physiology Score (SAPS), en sus versiones I, II y III, todos construidos sobre la base del análisis y comparación de las variables fisiológicas (8-10).

En Chile, los más usados son el APACHE II y SAPS III. Sin embargo, existe un único estudio realizado en UCI chilena, publicado el año 1993, que evalúa la capacidad predictiva de uno de ellos. El objetivo de dicho estudio fue comparar el valor predictivo del APACHE II máximo y el determinado en las primeras 24 horas. Dicho estudio muestra que los pacientes fallecidos se agrupan en los intervalos superiores del APACHE, es decir, en los valores máximos, que coinciden con una mayor capacidad predictiva de mortalidad (0,886;  $p < 0,001$ ) (11).

Por ello, se considera relevante realizar estudios complementarios a los realizados y publicados, tendientes a evaluar el desempeño de los scores APACHE II y SAPS III, contribuyendo de esta forma con los procesos de toma de decisiones en las UCI(s) chilenas.

Con el fin de describir y analizar la tasa de sobrevivencia atribuible a la UCI(s) Adultos, del Hospital Público de la ciudad de Curicó; se realizó un estudio descriptivo transversal de los pacientes que acudieron a dicho servicio entre los años 2011-2013, inclusive, clasificados según las variables APACHE II (Acute Physiology and Chronic Health Evaluation Score II) y SAPS III (Simplified Acute Physiology Score III), sobre la base de los siguientes objetivos:

- Validar los puntajes APACHE II y SAPS III como predictores de mortalidad en una muestra de pacientes

admitidos en la UCI del Hospital de Curicó, entre los años 2011 y 2013.

- Comparar la capacidad predictiva de estos dos en cuanto a especificidad, sensibilidad, discriminación, calibración.

## MATERIAL Y MÉTODOS

### Diseño del estudio

El presente trabajo es un estudio analítico, observacional de cohorte histórica de casos consecutivos desde la admisión a la UCI hasta el egreso hospitalario. Contexto: Los datos fueron recopilados entre el 01 de enero 2011 y el 31 de diciembre 2013 en la Unidad de Cuidados Intensivos de Adultos del Hospital de Curicó. Esta unidad tiene una dotación de 10 camas, cuenta con un equipo de salud conformado por 43 funcionarios: 7 médicos, 13 enfermeras, 13 técnicos paramédicos, 4 auxiliares de servicio, 5 kinesiólogos y una secretaria.

### Participantes

Los pacientes críticamente enfermos admitidos a UCI entre 1° enero 2011 y 31 diciembre 2013. Como criterio de inclusión se consideró los pacientes mayores de 18 años, con una estancia en UCI mayor a 24 horas.

### Fuentes de datos

Los datos se recolectaron retrospectivamente de la historia clínica de los pacientes de UCI del Hospital de Curicó. Fueron registrados por personal médico, usando el formulario electrónico (Base de datos Access-Microsoft Office) diseñado de acuerdo con las reglas generales y la definición de los predictores APACHE II y SAPS III. El APACHE II fue medido a las 24 horas de ingreso y el SAPS III en las primeras horas del ingreso.

### Sesgo

Para disminuir el sesgo, el formulario electrónico, permite prevenir valores inexactos. Los puntajes APACHE II y SAPS III, son revisados por el Médico Jefe de la UCI. El puntaje APACHE II y SAPS III, promedio, se mantiene estable para los tres períodos en estudio, tanto para los pacientes fallecidos como sobrevivientes.

### Datos

Los datos analizados en esta investigación son de fuente primaria, recogidos mediante técnicas de análisis documental, la observación directa no participativa y el análisis de los procesos. No se tuvo acceso a la identidad de los pacientes, por ello no fue necesario solicitar consentimiento informado.

Los datos fueron recopilados entre el 01 de enero 2011 y el 31 de diciembre 2013 en la Unidad de Cuidados In-

tensivos de Adultos del Hospital de Curicó. De un total de 1.042 pacientes ingresados a la Unidad de Cuidados Intensivos (UCI) del Hospital de Curicó, entre los años 2011 y 2013, se incluye 793 pacientes sobrevivientes y a 249 pacientes fallecidos que representan el 76,1% y 23,9% respectivamente, del total.

**Tamaño muestral**

De un total de 1 042 pacientes ingresados a la Unidad de Cuidados Intensivos (UCI) del Hospital de Curicó, entre los años 2011 y 2013, se incluyeron 941 pacientes, que cumplieron criterios de inclusión; pacientes mayores de 18 años, con una estancia en UCI mayor a 24 horas. La muestra representa el 90,3% del total.

**Variables**

Variable dependiente: Sobrevivir = 0; Fallecer = 1  
 Variables independientes: APACHE II para el modelo estadístico 1 y SAPS III para el modelo estadístico 2.

**Método estadístico**

El análisis estadístico de los datos se realizó con el apoyo del Software SPSS 15.0 y se complementó con Minitab 16. Se usó el Modelo de Regresión Logística Binaria. En donde se plantea la hipótesis de distribución binomial de la variable respuesta (12):

$$y_i = \frac{e^{\sum_{p=0}^{34} \beta_p X_{pi}}}{1 + e^{\sum_{p=0}^{34} \beta_p X_{pi}}} + \varepsilon_i$$

**Tabla 1. Resultados Comparativos APACHE II y SAPS III**

	Punto de Corte	Sensibilidad	Especificidad	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
APACHE II	0,5	93%	41,7%	0,2	0,4
SAPS III	0,5	95%	36,8%	0,2	0,3

**Bondad del ajuste**

El primer modelo, Logit1, donde  $\beta_1$ =puntaje APACHE II, es significativo, explica entre el 0,2 y el 0,4 de la variable dependiente, y clasifica correctamente el 80,9% de los casos (Tabla 1).

$$y_i = F(\beta_0 + \beta_1 \text{Puntaje APACHE}) + \varepsilon_i$$

El segundo modelo, Logit2 donde  $\beta_1$ =SAPS III es significativo, explica entre el 0,2 y el 0,3 de la variable dependiente, y clasifica correctamente el 81,2% de los casos (Tabla 1).

$$y_i = F(\beta_0 + \beta_1 \text{Puntaje SAPSIII}) + \varepsilon_i$$

Bondad de ajuste en regresión logística:

- Test de Bondad de Ajuste Hosmer – Lemeshow (13).
- Test de Cox y Snell
- Test de Nagelkerke

El desempeño de los puntajes de gravedad para los modelos estadísticos APACHE II y SAPS III, se evaluará en función de dos variables: discriminación y calibración. La discriminación es la capacidad de diferenciar entre individuos que sobreviven de los que mueren. La calibración hace referencia a qué tan cercana está, desde el punto de vista de la ocurrencia del evento en cada paciente, la probabilidad de mortalidad estimada por el modelo, de la observada a lo largo del rango de probabilidades (14).

**RESULTADOS**

Del total de 941 pacientes incluidos en el estudio, el 54,5% son hombres y el 45,5% son mujeres, con edad promedio de 57 y 59 años respectivamente. En promedio 8 días cama para el caso de los hombres y 6 días para las mujeres.

No existen diferencias en el promedio de puntaje APACHE II (19,8 ± 9,7) y SAPS III (55,4 ± 18,7) entre hombres y mujeres. Dichos puntajes se mantiene estable para los tres períodos en estudio.

La mortalidad global fue de 23,7% y las principales causas de ingreso a la UCI, son Sepsis respiratoria 12,9% (N=121); Trauma 8,6% (N=81); Sepsis digestiva 8,2% (N= 77); Cardiovascular – Paro cardio respiratorio 7,2% (N=68).

El SAPS III, presenta mejor discriminación o capacidad predictiva que el APACHE II, según el áreas bajo la curva de características operativas del receptor 0,81 y 0,80 respectivamente, tomado como referencia que valores cercanos a 1 se entienden como perfecta discriminación y cercanos a 0,5 como ninguna discriminación (15). (Figura 1).

La sensibilidad para ambos modelos estadísticos, también es buena. Del total de pacientes pronosticados como fallecidos por ambos modelos, el SAPS III clasificó correctamente el 95% y el APACHE II el 93%, con probabilidad superior a 0,5. No obstante, la especificidad para ambos modelo no es buena. Del total de pacientes pronosticados como sobreviviente por ambos modelos estadísticos, el SAPS III clasificó correctamente el 36,8% y el APACHE II el 41,7% (Tabla1).

Figura 1. Curva ROC, APACHE II v/s SAPS III

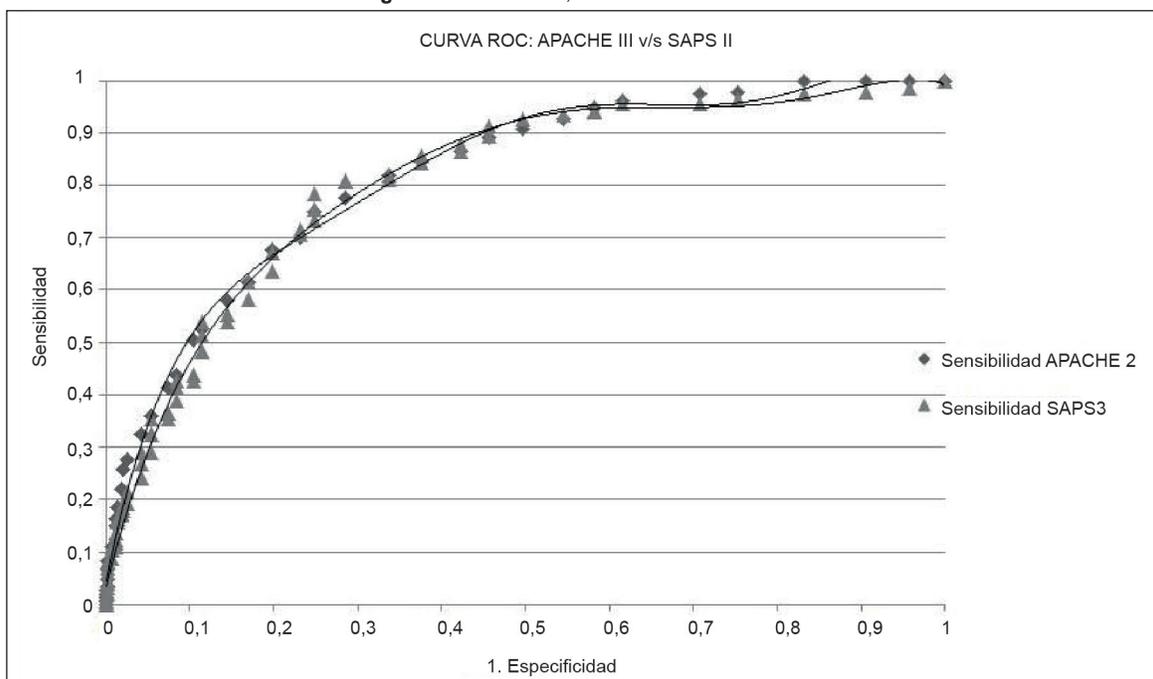


Tabla 2. Prueba de Hosmer y Lemeshow, APACHE II  
Chi2 = 3,7 (8); p – valor = 0,9

Paso	Destino = 0		Destino = 1		Total
	Observado	Esperado	Observado	Esperado	
1	88	86,2	0	1,8	88
2	88	88,6	4	3,4	92
3	100	98,1	4	5,9	104
4	86	88,4	11	8,6	97
5	89	89,4	14	13,6	103
6	65	65,3	15	14,7	80
7	65	66,3	23	21,7	88
8	59	60,5	33	31,5	92
9	50	48,6	46	47,4	96
10	28	26,6	73	74,4	101

(\*) = Destino: 0 = Sobrevivir, 1 = Fallecer

Tabla 3. Prueba de Hosmer y Lemeshow, SAPS III  
Chi2 = 10,2 (8); p-valor = 0,3

Paso	Destino (*) = 0		Destino(*) = 1		Total
	Observado	Esperado	Observado	Esperado	
1	84	87,2	5	1,8	89
2	95	94,3	4	4,7	99
3	95	91,8	4	7,1	99
4	91	90,4	10	10,6	101
5	90	86,1	11	14,9	101
6	77	77,2	20	19,8	97
7	66	70,0	31	26,9	97
8	58	60,1	39	36,9	97
9	45	45,7	51	50,3	96
10	17	15,0	48	49,9	65

(\*) = Destino: 0 = Sobrevivir, 1 = Fallecer

Con respecto a la calibración, ambos modelos estadísticos, tal como observamos en las Tablas 2 y 3 presentan buena calibración. Para cada uno de los intervalos existe proximidad entre los valores observados (real) y los esperados (teórico). Sin embargo, según el p-valor el modelo que mejor se ajusta a la realidad es el APACHE II.

No obstante, ambos modelos estadísticos no se pueden considerar aceptables dado que presentan especificidad inferior a un 75%.

## DISCUSIÓN

Para ambos modelos estadísticos, APACHE II (Logit1) y SAPS III (Logit2), se observa en general una buena capacidad de discriminación y calibración, no obstante, la especificidad para ambos modelos es baja. Estos resultados no son muy distintos a lo demostrado en otros estudios de similares características metodológicas.

En Colombia, para una cohorte 2006-2011, en una muestra de 2 523 pacientes admitidos en UCI, se concluyó que el SAPS III, tiene mejor capacidad discriminatoria que el APACHE II (0,8; 0,7) (14), v/s (0,81; 0,80) del presente estudio.

En Brasil, en un estudio realizado en dos UCIs especializadas en pacientes quirúrgicos, de hospitales diferentes el SAPS III tiene un buen desempeño en sensibilidad y especificidad (0,8; 0,9) (16). Dichos resultados son congruentes con los resultados obtenidos en sensibilidad (0,95),

pero, los resultados se alejan para la especificidad (0,4). No obstante, también en dicho país existen estudios que no muestran un buen desempeño del SAPS III, se trataba de una población en UCI de patologías de casos mixtos, con una menor proporción de casos quirúrgicos, contexto similar de los pacientes del presente estudio (17).

En Corea, en un estudio realizado entre los años 2011 y 2012, con una muestra de 1 314 pacientes, se concluyó que tanto el APACHE II como SAPS III tienen buena capacidad de discriminación (0,85; 0,86). No obstante, ambos modelos no presentan buena calibración según el Test de Hosmer y Lemeshow (18).

En España, en una UCI mixta, se obtuvieron resultados en la misma dirección del presente estudio. En una cohorte de 864 pacientes, la discriminación fue buena para ambos modelos SAPS III (0,9) y APACHE II (0,8). Sin embargo, la calibración solo resultó apropiada para el SAPS III (19).

Paralelamente, en Australia y Nueva Zelanda, actualmente, existen estudios que muestran que los sistemas APACHE II y SAPS III, obtienen mayor calidad de calibración y discriminación sobre la base de sistemas adaptados, es decir, los sistemas originales se adaptan a las características propias de la población. También, se propone un nuevo predictor de mortalidad, tomando como base ambos indicadores, es decir, APACHE II y SAPS III (20-21).

Finalmente, en concordancia con otros estudios internacionales, se muestra que los indicadores originales APACHE II y SAPS III, son un buen predictor en términos generales, no obstante, existe una tendencia internacional a adaptar dichos indicadores. Por ello, se recomienda realizar estudios analíticos enfocados a comprender e identificar las variables que explican la mortalidad de los pacientes admitidos en UCIs, sobre la base de las características propias de una población, es decir, teniendo presente la complejidad y diversidad de los pacientes que ingresan a una unidad de cuidados intensivos.

Los indicadores de predicción de mortalidad en UCI; APACHE II y SAPS III tienen una buena capacidad predictiva general, pero ambos indicadores presentan una baja especificidad. Por ello, es recomendable inferir un indicador de predicción de mortalidad en pacientes de UCI más exacto, reproducible y aplicable al contexto nacional ♠

## REFERENCIAS

1. Abdel P, Amaury Ch, Armando P, Teddy T, Rosa J. Evaluación del pronóstico de mortalidad por los índices de gravedad APACHE II y IV. *Revista Cubana de Medicina Intensiva y Emergencia*. 2013; 12 (1): 2334 -57.
2. Sakallaris B, Jastremski C, Von Rueden K. Clinical decision support systems for outcome measurement and management. *Critical Care*. 2000; 11(3):351-62.
3. Steyerberg EW. *Clinical Prediction Models. A Practical Approach to Development, Validation, and Updating*. 2009. Disponible en: <http://www.springer.com/us/book/9780387772431#reviews>. Consultado en marzo 2016.

4. Knaus W, Zimmerman J, Wagner D, Draper E, Lawrence D. APACHE-acute physiology and chronic health evaluation: a physiologically based classification system. *Crit Care Med*. 1981; 9 (8):591-7.
5. Knaus W, Draper E, Wagner D, Zimmerman J. APACHE II: a severity of disease classification system. *Crit Care Med*. 1985; 13(10):818-29.
6. Wagner D, Draper E, Knaus W. APACHE III study design: analytic plan for evaluation of severity and outcome in intensive care unit patients. Development of APACHE III. *Crit Care Med*. 1989; 17(12 Pt 2): S199-203.
7. Zimmerman J, Kramer A, McNair D, Malila F. 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.
8. Le Gall J, Loirat P, Alperovitch A, Glaser P, Granthil C, Mathieu D, Mercier P, Thomas R, Villers D. A simplified acute physiology score for ICU patients. *Crit Care Med*. 1984; 12(11):975-7.
9. Le Gall J, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA*. 1993 22-29; 270(24):2957-63. Erratum in: *JAMA* 1994 4; 271(17):1321.
10. Rui P, Moreno P, Metnitz E, Almeida B, Jordan P, Bauer R, Abizanda C, Gaetano I, David E, Maurizia C, Le Gall J. on behalf of the 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 Medicine*. 2005; 31, (10):1345-55.
11. Dougnac A, Andresen H, Rabagliati B, Landerretche S, Francois P, Villarreal L, Prado J. Mortalidad en Unidades de Tratamientos Intensivos: Valor predictivo de puntaje de gravedad APACHE II versus APACHE Máximo. *Rev. Méd. Chile*. 1993; 121:52-58.
12. Collett D. *Modelling Binary Data*. Chapman and Hall. London; 1991.
13. Lemeshow S., Hosmer DW. A review of goodness of fit statistics for use in the development of logistic regression models. *American Journal of Epidemiology*. 115:92–106, 1982.
14. Nelson G, Juan Manuel T, Carlos C, Felipe Z, Fabián J. Desempeño del APACHE II y el SAPS 3. Adaptación regional en una población de pacientes críticos de Colombia. *Acta Med Colomb*. 2014; 39 (2):148- 158.
15. Hanley J, McNeil B. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve1. *Diagnostic Radiology*. 1982; 143 (1): 29 -36.
16. Silva J, Malbouisson L, Nuevo H, Barbosa L, Marubayashi L, Teixeira I, Nassar A, Carvalho C, Ferreira S, Costa J, Rezende E. Aplicabilidade do Escore Fisiológico Agudo Simplificado (SAPS 3) em Hospitais Brasileiros. *Revista brasileira de anestesiologia*. 2010; 60(1): 20-31.
17. Nassar A, Mocelin A, Nunes A, Giannini F, Brauer L, Andrade F, Poianas F, Brauer L, Moreira F, Augusto C. Caution when using prognostic models: A prospective comparison of 3 recent prognostic models. *J Crit Care*. 2011; 27(4):1-7.
18. Hannah L, Yoon-Jung S, Hyerim K, Hyesun P, Hee-Pyoung P. Validation of the APACHE IV model and its comparison with the APACHE II, SAPS 3, and Korean SAPS 3 models for the prediction of hospital mortality in a Korean surgical intensive care unit. *Korean Journal of Anesthesiology*. 2014; 67(2):115-22.
19. Mbongo C, Monedero P, Guillen-Grima F, Yepes M, Vives M, Echarri G. Performance of SAPS3, compared with APACHE II and SOFA, to predict hospital mortality in a general ICU in Southern Europe. *Eur J Anaesthesiol* 2009; 26(11):940-5.
20. Eldho P, Bailey M, Pilcher D. Risk prediction of hospital mortality for adult patients admitted to Australian and New Zealand intensive care units: Development and validation of the Australian and New Zealand Risk of Death model. *Journal of Critical Care*. 2013; 28 (6): 935-941.
21. Schneider A, Lipcsey M, Bailey M, Pilcher D, Bellomo R. Simple translational equations to compare illness severity scores in intensive care trials. *Journal of Critical Care*. 2013; 28(5): 885.e1–885.