

Revista de Saúde Pública
ISSN 0034-8910 versão impressa

Rev Saúde Pública 2003; 37(5)

Comparison between different mortality risk scores in a neonatal intensive care unit

Mariani Schlabendorff Zardo and Renato S Procianoy

Serviço de Neonatologia. Hospital de Clínicas de Porto Alegre. Porto Alegre, RS, Brasil

ABSTRACT

Objective

To evaluate and compare birthweight and risk scores as predictors of neonatal mortality in a Neonatal Intensive Care Unit (NICU).

Methods

The survey included 494 newborns admitted to the neonatal intensive care unit (NICU) of a general hospital in Porto Alegre, southern Brazil, immediately after delivery, between March 1997 and June 1998. Birthweight and scores were evaluated in terms of variable "death while in NICU". Exclusion criteria were: discharge or death less than 24 hours after admission, admission not immediately following delivery, incomplete study protocol, and congenital malformations incompatible with survival. For CRIB (Clinical Risk Index for Babies) evaluation purposes, only patients born weighing up to 1,500 g were considered. ROC (Receiver Operating Characteristic) curves were calculated for SNAP

(Score for Neonatal Acute Physiology), SNAP-PE (Score for Neonatal Acute Physiology – Perinatal Extension), SNAP II, SNAP-PE II, and CRIB scores, as well as for birthweight.

Results

Of the 494 patients studied, 44 died (8.9% mortality). Of the 102 patients born weighing up to 1,500 g, 32 (31.3%) died. The area below the ROC curves ranged from 0.81 to 0.94. There were no statistically significant differences between the areas obtained for all scores evaluated. All mortality risk scores evaluated performed better than birthweight, especially on newborns with birthweight =1,500 g.

Conclusions

All neonatal mortality scores had better performance and were superior to birthweight as measures of in-hospital mortality risk for newborns admitted to NICU.

Keywords

Neonatal mortality (public health). Birth weight. Inpatients. Infant, newborn. Risk assessment. Intensive care units, neonatal. Score.

INTRODUCTION

The study of illness severity and mortality risk measurement among newborns admitted to neonatal intensive care units (NICUs) is attaining an increasing level of importance.¹¹ In order to compare mortality levels of different NICUs, even after making adjustments for factors such as gender, birthweight, gestational age, and ethnicity, it is still necessary for subject disease severity to be similar.¹⁵

In 1993, three scores were described for measuring illness severity and neonatal mortality among newborns admitted to intensive care units: SNAP (*Score for Neonatal Acute Physiology*), SNAP-PE (*Score for Neonatal Acute Physiology - Perinatal Extension*), and CRIB (*Clinical Risk Index for Babies*). The SNAP score is based on 34 variables, evaluated during the worst moment of the first 24 hours after admission;^{3,8,10-12} SNAP-PE adds to SNAP^{3,8,11} birthweight, small size for gestational age (SSGA), and low Apgar score at 5 minutes after delivery. The CRIB score evaluates six factors during the first 12 hours of life, but is appropriate only for newborns with gestational age 31 weeks or less and/or birthweight up to 1,500 g.^{2,3,5,6,12,13}

In 2001, Richardson et al¹³ developed and validated SNAP II – reducing to six the number of items evaluated, in order to render the system more feasible – and SNAP-PE II, increasing the scores attributed to perinatal variables so as to level their weight to that of physiological variables in the final score. Variables are collected during the first 12 hours after delivery, in order to minimize treatment interference.

The aim of the present study is to evaluate birthweight and SNAP, SNAP-PE, SNAP II, SNAP-PE II, and CRIB scores as predictors of neonatal in-hospital mortality among newborns admitted to NICUs, comparing their results.

METHODS

The study included all newborn babies in a general hospital in the city of Porto Alegre, southern Brazil, admitted to a NICU immediately after delivery between 1 March 1997 and 30 June 1998. Studied outcome was death while in the NICU. SNAP and SNAP-PE evaluation was done through a cohort study, whereas CRIB, SNAP II, and SNAP-PE II were evaluated retrospectively through a cross-sectional study.

Patients dead or discharged to normal newborn care less than 24 hours after admitted, admitted into NICU not immediately after delivery, whose study protocol was incomplete due to missing data in patient files, or with congenital malformations incompatible with survival were excluded.

Physiological variables and scores applied in order to generate SNAP, SNAP-PE, SNAP II, and SNAP-PE II scores were extracted from studies by Richardson et al.^{10,11,13} The CRIB score was applied according to Tarnow-Mordi et al.¹⁵

SNAP and SNAP-PE were evaluated prospectively, since they were routinely applied upon admission to the NICU, for use in a previously published study.¹⁴ Variables were registered after the patient's first 24 hours in the NICU, and the worst moment within this period was considered. SSGA classification was obtained based on the birthweight/gestational age curve, as described by Alexander et al.¹ Infants were considered as SSGA whose birthweight was below percentile 5, according to SNAP-PE recommendations.

CRIB, SNAP II, and SNAP-PE II were executed retrospectively, through an examination of patient files. The SNAP II and SNAP-PE II systems were described in 2001, therefore after data collection, and thus were analyzed retrospectively.¹³ CRIB scoring was not a routine procedure in the ICU studied, and was evaluated retrospectively in order to be compared to the remaining scores. Physiological variables required for score application, considering the first 12 hours after delivery, were collected. SSGA classification was obtained based on the birthweight/gestational age curve used by Richardson et al,¹⁰ which considers as of SSGA infants with birthweight under percentile three. CRIB was applied only for patients born weighing 1,500 g or less.

Gestational age was estimated using the New Ballard method for newborns with gestational-obstetric ages up to 34 weeks, and the Capurro method for the remainder.

F_{IO₂} was considered as adequate when able to maintain hemoglobin oxygen saturation at 90-95% – as measured by a wrist oxymeter –, a level which allows for good oxygenation along with a reduced risk of oxygen toxicity.

A simple descriptive analysis was used for study groups and subgroups – mean, median, standard deviation, and interquartile range (p25-p75). Due to the asymmetrical distribution found in the investigated variables, the Mann-Whitney U-Wilcoxon test was used for comparing median scores of dead and surviving patients. A ROC (Receiver Operating Characteristic) curve was constructed in order to compare between different scores. The Roc curve is a plot of sensitivity (correct prediction of death) against 1 – specificity (correct prediction of survival), calculated for each value of each test studied. The area below the ROC curve, Az, is a parameter for score discriminating performance. A 1.0 Az corresponds to perfect prediction, and a 0.5 Az to prediction entirely by chance. The area below the ROC curve was used to represent prediction precision. Areas were statistically compared

according to the Hanley & McNeil test.⁴ Software used were SPSS (Statistical Package for the Social Sciences) – version 10 and Epi Info. P-values less than 0.05 were considered as statistically significant.

The research project was approved by the Ethics Committee for Research on Human Beings (*Comité de Ética em Pesquisa em Seres Humanos*) of the *Hospital de Clínicas de Porto Alegre*.

RESULTS

During the fieldwork stage, 789 babies were admitted to the NICU. Of these, 243 were excluded for not being admitted immediately after birth (30.7%), 21 for dying less than 24 hours after delivery (2.6%), 22 due to incomplete patient files (2.7%), and nine due to congenital malformations incompatible with survival (1%). Of the 494 patients studied, 44 died, representing an 8.9% mortality. Patients born weighing 1,500 g or less were selected for CRIB assessment, yielding a total 102 patients. Mortality among this group was 31.3% (32 deaths) and accounted for 73% of total mortality.

When all patients were evaluated, mean birthweight was 2,354 g and mean gestational age, 36 weeks. Of the 494 patients, 256 were male and 81 had Apgar scores <7 at the fifth minute. Considering only patients with birthweight up to 1,500 g, mean weight was 1,050 g and mean gestational age, 31 weeks. Of the 102 patients, 64 were female, and 30 had Apgar scores <7 at the fifth minute.

The median scores of surviving patients scores were compared to those of patients who died. All differences were statistically significant, with $p < 0,0001$ (Table 1).

Table 1 – Median score comparison for all patients studied.

Score	Survivals*	Deaths*	Significance*
All subjects			
BW (g)	2,418(1,775-3,070)	860(643-1,690)	< 0.0001
SNAP	4(2-7)	14(9-24)	< 0.0001
SNAP PE	6(3-11)	34(21-53)	< 0.0001
SNAP II	0(0-9)	25(13-42)	< 0.0001
SNAP PE II	5(0-18)	55(33-74)	< 0.0001
Subjects with birthweight =1,500 g			
BW (g)	1,195(1,038-1,326)	773(576-1,006)	< 0.0001
CRIB	2(1-5)	12(8-16)	< 0.0001
SNAP PE	10(8-18)	43(26-55)	< 0.0001
SNAP PE II	20(12-25)	60(41-78)	< 0.0001

*Median (p25-p75)

BW – Birthweight

SNAP (*Score for neonatal acute physiology*)

SNAP-PE (*Score for neonatal acute physiology extension*)

Figures 1 and 2 present the area under the ROC curve for all disease severity scores. Performance levels are located between good and excellent. Birthweight achieved a 0.81 area for all subjects. SNAP, SNAP-PE, SNAP II, and SNAP-PE II areas for the same group were 0.85, 0.90, 0.88, and 0.91, respectively. In the birthweight =1,500 g group, Az equaled 0.82 for birthweight, 0.93 for SNAP-PE, 0.94 for SNAP-PE II, and 0.91 for CRIB.

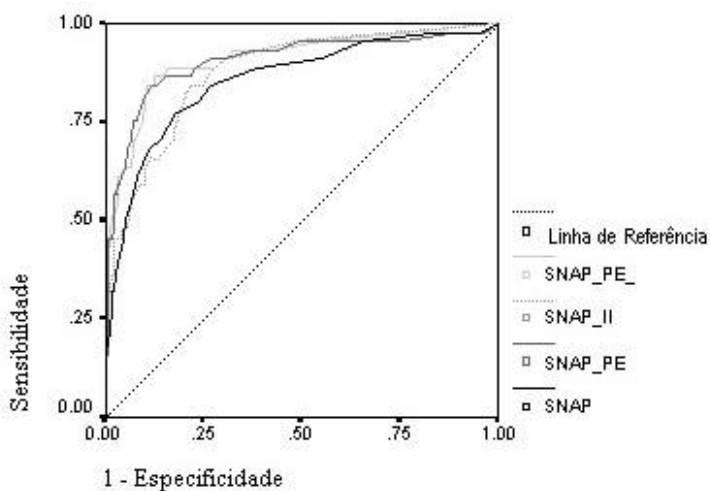
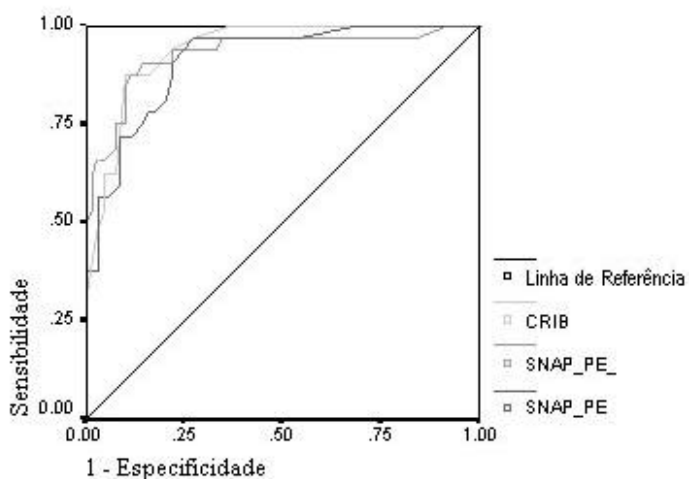


Figure 1 - Area under ROC (receiver operating characteristic curve) of all study patients' scores



Sensitivity
 Specificity
 Baseline [linha de referencia]

[sensibilidade]
 [especificidade]

Figure 2 - Areas under ROC of scores of patients with 1,500 g birth weight

SNAP-PE II achieved the greatest area among studied scores in both groups. When considering all subjects, SNAP-PE was statistically more accurate than birthweight alone in predicting neonatal mortality and SNAP-PE II achieved a borderline result (Table 2). When considering only subjects with birthweight =1,500 g, all scores were statistically more accurate than birthweight (Table 3). However, there was no statistically significant difference in Az between the various scores (Tables 2 and 3).

Table 2 – P-values for the areas below the ROC curves of the various scores for all subjects.

	SNA P	SNAP PE	SNAP II	SNAP PE II
Birthwei ght	0.40 4	0.030	0.188	0.055
SNAP		0.089	0.428	0.074
SNAP PE			0.401	0.791
SNAP II				0.056

ROC – Receiver operating characteristic curve.

Table 3 - P-values for the areas below the ROC curves of the various scores for subjects with birthweight =1.500 g.

	SNAP PE	SNAP PE II	CRI B
Birthweight	0.024	0.039	0.01 4
SNAP PE		0.547	0.25 7
SNAP PE II			0.67 4

CRIB (Clinical risk index for babies)

DISCUSSION

Birthweight has been used, for many years, as a mortality risk indicator for newborn babies. Recent studies, however, have emphasized prognosis disparities between babies born with similar weight in different NICUs.^{5,10,12} Such differences are related to variations in population risk and technology employed,¹¹ which would justify the importance of severity indicators in the comparison of different NICUs.

The areas below the ROC curves calculated in the present study are similar to those reported by the score validation studies.^{5,10,12} Birthweight was found to be the indicator least able to predict neonatal mortality. Results obtained were similar to those verified by other authors, for whom the Az for

birthweight varied from 0.74 to 0.82, values very close to those found in the present survey (0.81 ± 0.46).^{5,7,11} This finding confirms the importance of physiological evaluation in determining disease severity and evaluating mortality risk upon admission to neonatal intensive therapy. The use of physiological variables alone, however, has been considered as insufficient for predicting neonatal mortality in very low birthweight preterm babies.^{8,11,14}

Despite the lack of statistically significant differences in terms of Az, SNAP-PE II was found to be superior to birthweight, SNAP, and SNAP II when considering all patients. This demonstrates the importance of the perinatal extension included in SNAP-PE II, which takes into account – in addition to physiological variables – birthweight, Apgar score at fifth minute, and SSGA classification as factors that contribute to an increase in mortality risk.¹²

Recently, Richardson et al¹³ found a 0.78 ± 0.01 Az for birthweight alone as a predictor of mortality for newborns admitted to NICUs. When considering the perinatal extension – that is, birthweight, SSGA, and Apgar <7 at fifth minute, Az equaled 0.84. However, when considering both physiological variables and perinatal extension, Az increased to 0.91 ± 0.01 , which indicates the importance of both sets of information as predictors of mortality.

The validation study for CRIB, conducted by Tarnow-Mordi et al,¹⁵ found a 0.90 Az, a value similar to that encountered in the present survey (Az CRIB=0.91).⁵ There is a consensus in the literature – which also applies to the present study – as to the superiority of the CRIB score in relation to birthweight alone in predicting neonatal mortality.⁷ Rautonen et al⁹ concluded, after comparing SNAP, SNAP-PE, and CRIB scores, that the latter had a better performance with statistically significant differences in comparison to SNAP and SNAP-PE. In the present survey, SNAP-PE II had the best performance among patients born weighing up to 1,500 g. However, a comparison between SNAP-PE II and CRIB revealed no statistically significant differences. Both tests consider the first 12 hours after delivery, thus being less influenced by treatment.^{2,5,13} CRIB has the disadvantage of being restricted to patients with birthweight =1,500 g.

The scores studied are tools for measuring in-hospital mortality risk among very ill newborns admitted to NICUs. They should not be used for orienting individual decisions related to any individual patient.

The present survey presented the results of six disease severity measures in relation to neonatal mortality in the NICU studied. All evaluated scores had results considered as good, with no statistically significant differences. As to their applicability, CRIB, SNAP II, and SNAP-PE II are faster and easier to apply, since the number of variables considered in these tests is smaller. The present results show that all mortality risk scores investigated had a better performance than birthweight alone, particularly among newborns weighing up to 1,500 g.

ACKNOWLEDGEMENTS

To statistician Vânia Naomi Hirakata, of the *Universidade Federal do Rio Grande do Sul*, for her suggestions and orientation concerning statistical analysis; and Dr. Rita C. Silveira, of the Neonatal Intensive Care Unit of the *Hospital de Clínicas de Porto Alegre* for her suggestions and support in carrying out the project.

REFERENCES

1. Alexander GR, Himes JH, Kaufman RB, Mor J, Kogan M. A United States National reference for fetal growth. *Obstet Gynecol* 1996;87:163-8.
 2. Courcy-Wheeler RHB, Wolfe CDA, Fitzgerald A, Spencer M, Goodman JDS, Gamsu HR. Use of the CRIB (Clinical Risk Index for Babies) score in prediction of neonatal mortality and morbidity. *Arch Dis Child* 1995;73:F32-6.
 3. Fleisher BE, Murthy L, Lee S, Constantinou JC, Benitz WE, Stevenson DK. Neonatal severity of illness scoring systems: a comparison. *Clin Pediatr (Phila)* 1997;4:223-7.
- Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating curves derived from the same cases. *Radiology* 1983;148:839-43.
- The International Neonatal Network. The CRIB (Clinical Risk Index for Babies) score: a tool for assessing initial neonatal risk and comparing performance of neonatal intensive care units. *Lancet* 1993;342:193-8
6. Kaaresen PI, Døhlen G, Fundingsrud HP, Dahl LB. The use of CRIB (clinical risk index for babies) score in auditing the performance of one neonatal intensive care unit. *Acta Paediatr* 1998;87:195-200.
 7. Matsuoka OT, Sadeck LSR, Haber JFS, Proença RSM, Mataloun MMG, Ramos JLA et al. Valor preditivo do "Clinical Risk Index for Babies" para o risco de mortalidade neonatal. *Rev Saúde Pública* 1998;32:550-5.
 8. Pollack MM, Koch MA, Bartel DA, Rapoport I, Dhanireddy R, El-Mohandes AAE et al. A comparison of neonatal mortality risk prediction models in very low birth weight infants. *Pediatrics* 2000;105:1051-7.
 9. Rautonen J, Mäkelä A, Boyd H, Apajasalo M, Pohjavuori M. CRIB and SNAP: assessing the risk of death for preterm neonates. *Lancet* 1994;343:1272-3.

Address to correspondence

Mariani Schlabendorff Zardo
Rua Barão de Ubá,194/502
90450-090 Porto Alegre, RS, Brasil
E-mail: mzardo@telenova.net

Received on 29/7/2002. Reviewed on 15/5/2003. Approved on 6/6/2003.

© 2003 *Faculdade de Saúde Pública da Universidade de São Paulo*

**Avenida Dr. Arnaldo, 715
01246-904 São Paulo SP Brazil
Tel./Fax: +55 11 3068-0539**

revsp@org.usp.br