Prediction of Mortality by Using the Standard Scoring Systems in a Medical Intensive Care Unit in Thailand

Ranistha Ratanarat, MD*,
Maneerat Thanakittiwirun, MD*, Warakarn Vilaichone, MD*,
Surat Thongyoo, MD*, Chairat Permpikul, MD*

** Department of Medicine, Faculty of Medicine, Siriraj Hospital, Mahidol University*

In order to evaluate and compare the predictive ability of the APACHE II (Acute Physiology and Chronic Health Evaluation II) and the SAPS (Simplified Acute Physiology Score) scoring systems in relation to outcome in a medical intensive care unit (ICU). The authors reviewed consecutive medical ICU admissions (n = 482) at a tertiary hospital over a 2-year period. For each patient, demographic data, diagnosis, APACHE II score, SAPS score and ICU outcome complied during the first 24 hrs of the ICU stay were obtained. The comparison of predictive ability between APACHE II and SAPS was assessed by forward stepwise logistic regression and area under the receiver operating characteristic (ROC) curves. Overall ICU mortality was 36.93%. Mean APACHE II and SAPS scores were 21.17 ± 9.35 and 14.61 ± 6.47, respectively. APACHE II and SAPS scores of nonsurvivors (26.97 ± 8.27 and 18.01 ± 5.84 respectively) were significantly higher than those of survivors (17.77 ± 8.22 and 12.62 ± 5.99 respectively) (p < 0.001). Correlation between both systems was excellent (Pearson correlation coefficient, r = 0.825: p < 0.001). The predicted risk of death calculated by using the APACHE II risk of death equation was 38.98%. The predictive ability to discriminate between survivors and nonsurvivors of APACHE II was higher than SAPS according to forward stepwise logistic regression and area under the ROC curves (APACHE II was 0.788 while SAPS was 0.746). In conclusion, the APACHE II scoring system is an efficient predictor for monitoring the hospital outcome and has more predictive ability than the SAPS in the medical ICU patients.

Keywords: *Critically ill patient, ICU, APACHE II, SAPS, Outcome*

J Med Assoc Thai 2005; 88(7): 949-55
Full text. e-Journal: *http://www.medassocthai.org/journal*

The intensive care unit (ICU) is a specialized area in which medical technology and personnel are concentrated for cases of critically ill patients in virtually every major hospital[1,2].

A reasonable goal of intensive care is to save the life of patients with reversible medical conditions and offer the dying a peaceful and dignified death. The current quest for improved effectiveness has motivated different groups in health care to develop tools that would be of use in predicting the outcome from critical illness. The concept of providing cost-effective intensive care has now generalized to all developed countries, becoming a major interest of clinicians, hospital administrations, health care managers, medical economists and governmental policy makers[3,4].

For these economic and therapeutic reasons, a number of general multipurpose severity scoring systems have been developed over the past few decades by applying linear regression models to prospectively collected data[5-11]. These scores help to standardize the estimation of treatment effectiveness, the risk of hospital death, and the performance of various ICUs.

Among these scoring systems, Acute Physiology and Chronic Health Evaluation System II (APACHE II) and Simplified Acute Physiology Score (SAPS) have been widely used because they are reliable, inexpensive and relatively easy to calculate[5]. APACHE II, introduced in 1985[7], was a simplified

Correspondence to : Ratanarat R, Division of Critical Care Medicine, Department of Medicine, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.

refinement of the original APACHE[6]. It consisted of three parts: a) Acute Physiology Score (APS) (0-4) based on 12 most easily measured variables to cover most physiologic systems, b) age points (0-6) and c) chronic health points (0-5) were assigned only for severe organ system dysfunction. APACHE II has been shown to be an accurate measurement of severity of illness and correlate strongly with patient's outcome[8,12-14]. In addition, it is the most well known, the most frequently used and cited scoring system predictive model. Over the past few decades, the ability of the APACHE II system in predicting group outcome has been validated in several countries[5,12-19]. Le Gall et al, concerned about the number of variables necessary for calculation of the original APCHE and developed SAPS in 1984[9]. This system is based on 13 physiologic variables, found in 90% of patients in the APS survey. In Thailand, there was only one published report of validation of APACHE II in postoperative ICU[20], but none in medical ICU or by using SAPS. Moreover, a comparison of medical practice in Thailand and developed countries reveals a great number of differences including availability of ICU care, funding systems, and attitudes of both physicians and patients toward illness and death. Therefore, it is important that these severity scoring systems should be validated locally.

In the present study, research designed to assess the ability of the scoring systems (APACHE II and SAPS) to predict ICU outcome in medical ICU patients and to compare the predictive ability of the APACHE II and SAPS scoring systems.

Material and Method

The Intensive-Care Setting

Data were obtained from all patients admitted to the 7-bed medical intensive care unit at the Department of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand between January 1, 1999 and December 31, 2000. The ICU served medical patients in the Department of Medicine.

Data Collection

The admission records of each patient admitted to the ICU were reviewed retrospectively. Patient characteristics including demographic variables, diagnosis, APACHE II score and SAPS score at the time of ICU admission and length of stay in the ICU were obtained. The treatment outcome was determined by survival at the time of ICU discharge.

Scoring systems

The two severity scoring systems, APACHE II and SAPS, were calculated as described in the original publications[7,9]. At ICU admission, each patient admitted to the unit was assigned to the appropriate diagnostic categories which were used by the APACHE II scoring system[7].

A prediction of the patient's mortality risk was made on the basis of the APACHE II score and the diagnostic category weight using the regression formula developed by Knaus and associates[7]. (ln (R/1-R) = -3.517 + (APACHE II score x 0.146) + (0.603, only if postemergency surgery) + diagnostic category weight)

Statistical analysis

SPSS 9.0 J software was used for statistical analysis. The clinical parameters were reported as mean $\pm$ standard deviation or percent. Student's t-test was used for analysis of continuous variables and Chi-square (c2) test was used for analysis of categorical variables. A p value of less than 0.05 was considered as statistically significant. Linear correlation between the two indexes was assessed by the Pearson correlation coefficient. The correlation between the predicted mortality derived from APACHE II and the actual mortality was calculated by linear regression analysis.

The accuracy of outcome prediction by the APACHE II and SAPS systems and the comparisons between both scoring systems for prediction ability were assessed by a) forward stepwise logistic regression and b) area under the receiver operating characteristic (ROC) curves.

ROC curves, plotted separately for each index, were intended to illustrate the relationship between the proportion of true-positive (sensitivity) and false-positive (1-specificity)[21]. ROC curve is a generalized concept to determine sensitivity and specificity of a diagnostic test. Because a single threshold value of the diagnostic criterion (here the predicted death rate) may not be representative of the entire prediction procedure, the ROC curve graphs the performance of the procedure overall, summarized by the area under the ROC curve[21].

The survival graphs for each index were constructed from the patients' severity scores and actual hospital outcomes by using Kaplan-Meier model.

Results

During the study period, data were obtained from records of 505 patients but records of 482 patients

(95.44%) were completed and analyzed. Forty-seven percent of the admissions involved male patients. The mean age was 56 ± 21 years, and the mean length of ICU stay was 7.9 days (range 1 to 64 days). The average APACHE II and SAPS score were 21.17 ± 9.35 and 14.61 ± 6.47, respectively. As shown in Fig. 1, both systems had a high correlation (Pearson correlation coefficient, 0.825; $p < 0.001$)

One hundred and seventy-eight patients (36.93%) died during their ICU admission. Table 1 summarizes patient data, APACHE II score, SAPS score and length of ICU stay according to actual ICU death. Nonsurvivors had a statistically higher APACHE II score and SAPS score (26.97 ± 8.27 and 18.01 ± 5.84, respectively) than the survivor group (17.77 ± 8.22 and 12.62 ± 5.99, respectively) ($p < 0.001$), while other variables (sex, age, baseline serum creatinine and length of ICU stay) were not different between the two groups. Table 2 illustrates the principle diagnostic categories leading to ICU admission. The most common primary admission diagnosis was sepsis, accounting for 27 percent.

On the basis of APACHE II score and the diagnostic category weight, as described previously[7], the mean predicted risk of death for the overall patient population was 38.98%, whereas the actual death rate was 36.93%. Table 3 shows the relationship between predicted risk of death by using APACHE II score and actual ICU mortality. When a calibration curve was constructed by plotting the actual ICU mortality against the predicted mortality, linear regression yielded an r^2 of 0.92.

Discrimination between the two indices was assessed by two different methods; a) The forward stepwise logistic regression revealed a significantly

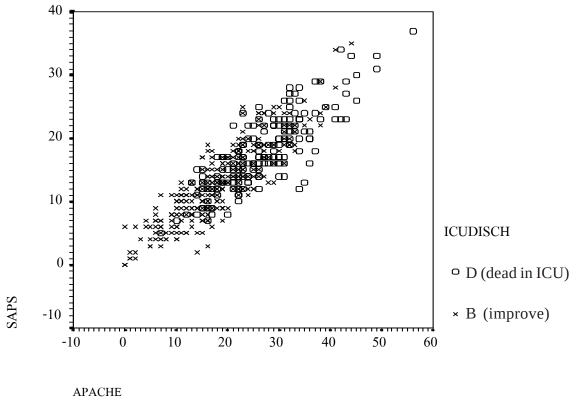


Fig. 1 Linear correlation between APACHE II and SAPS scores

better fit in favor of APACHE II, Chi-square model = 120.835; $p < 0.001$; b) Comparison of the area under the ROC curves of each severity index[22] (Fig. 2). Area under the APACHE II ROC curve was 0.788 with a standard error (SE) 0.021, larger than that found in SAPS (area under the ROC curve = 0.746, SE = 0.022). This result indicates that APACHE II had a better predictive ability than SAPS. When using the cut off point at APACHE II of 20 to predict patient outcome, the sensitivity and specificity were 80.9% and 63.2%, respectively. These values were the sensitivity 70.2% and the specificity 67.1% if the authors used the cut off point at SAPS score of 15.

Based on the Kaplan-Meier model, the estimated probability of survival in the ICU against APACHE II and SAPS scores were constructed by plotting a survival graph (Fig. 3 and 4). The survival graphs show that the median survival APACHE II score = 31 (95%CI = 30, 32) and median survival SAPS

Table1. Demographic data, APACHE II score, SAPS score, length of ICU stay according to actual outcome in ICU patients

Variable	All patients (n = 482)	Survivors (n = 304)	Nonsurvivors (n = 178)	p value
Sex				0.314
- male	227 (47.1%)	149 (49.0%)	78 (43.8%)	
- female	255 (52.9%)	155 (51.0%)	100 (56.2%)	
Age (yrs) ± SD	56.56±20.81	56.40±20.72	56.02±21.01	0.844
Creatinine (mg/dl)	2.70	2.59	2.86	0.59
APACHE II ± SD	21.17±9.35	17.77±8.22	26.97±8.27	<0.001*
SAPS ± SD	14.61±6.47	12.62±5.99	18.01±5.84	<0.001*
Length of ICU stay (days)	7.92	8.17	7.50	0.429

SD = standard deviation

Table2. Principle diagnostic categories for the 482 consecutively admitted medical ICU patients

Principle diagnostic categories	n (%)
Sepsis	130 (26.97%)
Congestive heart failure	38 (7.88%)
ARDS	37 (7.67%)
Respiratory failure from infection	28 (5.81%)
Post cardiac arrest	28 (5.81%)
Coronary artery disease	20 (4.12%)
Rhythm disturbance	15 (3.11%)
COPD	13 (2.70%)
Postoperative care	12 (2.49%)
Seizure	11 (2.28%)
Hemorrhagic/hypovolemic shock	11 (2.28%)
Respiratory failure from aspiration/poisoning/toxic	6 (1.24%)
Others	43 (8.92%)
Principle major vital organ	
Metabolic/renal	23 (4.77%)
Respiratory	30 (6.22%)
Neurology	18 (3.73%)
Cardiovascular	11 (2.28%)
Gastrointestinal	8 (1.66%)

Table 3. The actual ICU mortality in the 482 patients grouped according to mortality predicted by APACHE II

APACHE II	N	Predictedrisk of death (%)(mean)	Std. Deviation (%)	Actual mortality (%)
0-4	12	3.16	2.58	0
5-9	32	5.34	4.02	3.13
10-14	68	12.48	6.01	10.29
15-19	114	22.64	9.27	22.80
20-24	89	38.37	12.40	41.57
25-29	71	57.93	11.38	53.52
30-34	60	73.91	9.18	71.67
35-39	20	85.44	6.22	65.00
$\geq$ 40	16	92.23	5.19	81.25
Total	482	38.98	27.30	36.93

Linear regression analysis, $r^2 = 0.92$

= 21 (95%CI = 20, 22). This data analysis indicates that the estimated probability of survival in the authors' ICU patients with APACHE II score 31 or SAPS score 21 was 50%.

Discussion

In the present study, the result shows that the APACHE II and the SAPS scores of nonsurvivors were statistically higher than those of survivors in ICU, while the other variables i.e., age, sex, baseline renal function, length of ICU stay were not. Both the APACHE II and SAPS scoring systems were found to be accurate in group outcome prediction in ICU patients as assessed by ROC curves.

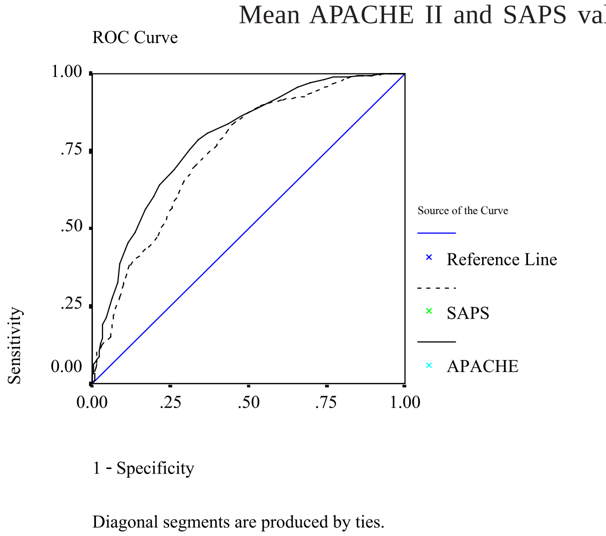
The ability of the APACHE II system to predict patients' risk of death was assessed by using the regression formula, developed by Knaus et al(7). the authors found that the APACHE II predicted risk of death was close to actual ICU mortality in all ranges of APACHE II scores (Table 3). These findings suggest that the APACHE II risk of death equation correlates well with actual mortality in patients. Based on the present results, the APACHE II and the SAPS scoring systems are useful tools to predict prognosis of these

Source of the Curve

Fig. 2 Receiver operating characteristic curve for medical ICU patients

patients and may be helpful in assisting clinical decision making. The authors proposed the survival graphs (Fig. 3, 4) for each severity system to predict mortality in an individual in the ICU, based on data from patients with the same degree of illness. However, these survival graphs should be periodically updated to make them valid for current use.

Mean APACHE II and SAPS values were



assessed
ı the two
'his is not
ıble con-
ity of the
ɔ live and
/ forward
ınder the
ability of
S system.
stem was
bles and
conclude
predictor
m in the

To determine ICU mortality as the primary outcome, while the original SAPS and APACHE II studies reported the hospital death. In the authors' opinion, there were a few good reasons to support that the ICU death is better than the hospital death to measure outcome of ICU. The first reason is that no intermediate care or step down unit. Secondly, the hospital has more critically ill patients who need ICU care than beds available in ICU. As a result, some patients who are at risk and need intermediate support have to move directly from the ICU to the medical ward. These reasons probably lead to a higher hospital death rate that is not related to the ICU performance.

The actual death rate of 36.93% compares higher than those of the other reports. Previously reported death rates varied from 16.9% in Japan[17], 18% in New Zealand[13], 19.7% in the United States[12], 24.8% in Canada[19], and 36% in Hong Kong[16]. After stratifying for severity of illness by using APACHE II score, ICU patients had an overall mean APACHE II score of 21.17 which was higher than previously reported APACHE II score, varying from 14.7 in Japan[17], 14.2 in New Zealand[13], 14.2 in the United States[12], 16.5 in Canada[19] to 20.1 in Hong Kong[16].

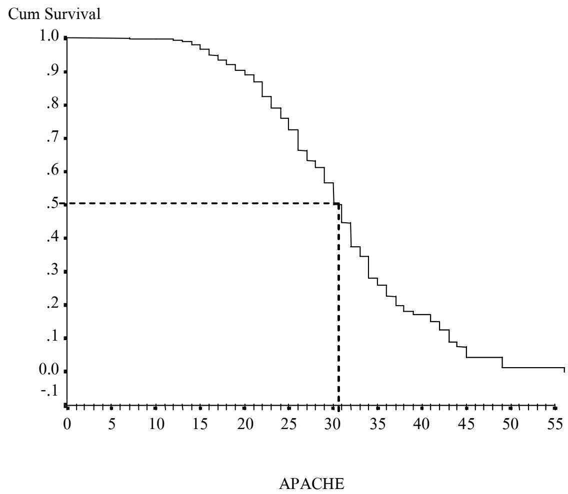


Fig. 3 Survival graph of APACHE II scoring system in ICU

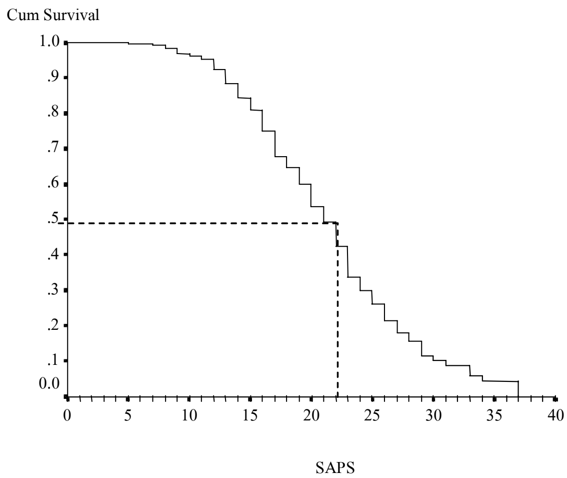


Fig. 4 Survival graph of SAPS scoring system in ICU

There were some differences between these ICU and others in developed countries. Firstly, ICU was a medical unit, while ICUs involved in other studies were mixed medical/surgical units. For patients with identical APACHE II scores, the predicted risk of death for medical patients was greater than for surgical patients. This finding is due to the higher weighting for disease categories of medical patients. Secondly, our ICU is a university teaching hospital and referral center in Thailand, and medical ICU had only 7 beds, accounting for 3% of all medical beds in the hospital. The patients, who are admitted to ICU were typically transferred from medical wards or other hospitals, have high APACHE II and SAPS scores. Moreover, the differences in the duration or aggressiveness of therapy received before ICU admission with normalization of many of the physiologic variables used in calculation of the predictive scoring model will affect the final score(23).

The authors have shown the APACHE II and SAPS scoring systems to be accurate predictors of ICU outcome in our ICU in Thailand. In addition, the present study indicates that the predictive ability of APACHE II system was higher than the SAPS system. However, the varieties of scoring system using large database and a higher number of physiologic variables have been developed. There were the recent version of APACHE III(10) (developed by Knaus and Colleagues) and SAPS II(11) (described by Le Gall and Colleague and based on a European/North American multicenter database). The APACHE III system has not been widely accepted or used because of a variety of factors but most notably that the APACHE III logistic regression coefficients and equations are proprietary information and not freely available(5). In contrast, the SAPS II scoring system is widely used and the SAPS II resulted in a significantly higher area under the ROC curve than was obtained with the original SAPS. Future studies using SAPS II system are needed to evaluate the predictive accuracy of SAPS II system and compare these two scoring systems (APACHE II AND SAPS II) in medical ICU in Thailand.

References

1. American Hospital Association. Hospital statistics 1978. Chicago: American Hospital Association, 1978.
2. Relman AS. Intensive-care units: who needs them? N Engl J Med 1980; 302: 965-6.
3. Thibault GE, Mulley AG, Barnett GO, Goldstein RL, Reder VA, Sherman EL, et al. Medical intensive care: indications, interventions, and outcomes. N Engl J Med 1980; 302: 938-42.
4. Marik PE, Kraus P, Lipman J. Intensive care utilisation: the Baragwanath experience. Anaesth Intensive Care 1993; 21: 396-9.
5. Marik PE, Varon J. Severity scoring and outcome assessment. Computerized predictive models and scoring systems. Crit Care Clin 1999; 15: 633-46.
6. 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; 9: 591-7.
7. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985; 13: 818-29.
8. Wong DT, Knaus WA. Predicting outcome in critical care: the current status of the APACHE prognostic scoring system. Can J Anaesth 1991; 38: 374-83.
9. Le Gall JR, Loirat P, Alperovitch A, Glaser P, Granthil C, Mathieu D, et al. A simplified acute physiology score for ICU patients. Crit Care Med 1984; 12: 975-7.
10. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. Chest 1991; 100: 1619-36.
11. Le Gall JR, Lemeshow S, Saulnier F. A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. JAMA 1993; 270: 2957-63.
12. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. An evaluation of outcome from intensive care in major medical centers. Ann Intern Med 1986; 104: 410-8.
13. Zimmerman JE, Knaus WA, Judson JA, Havill JH, Trubuhovich RV, Draper EA, et al. Patient selection for intensive care: a comparison of New Zealand and United States hospitals. Crit Care Med 1988; 16: 318-26.
14. Marsh HM, Krishan I, Naessens JM, Strickland RA, Gracey DR, Campion ME, et al. Assessment of prediction of mortality by using the APACHE II scoring system in intensive-care units. Mayo Clin Proc 1990; 65: 1549-57.
15. Knaus WA, Le Gall JR, Wagner DP, Draper EA, Loirat P, Campos RA, et al. A comparison of intensive care in the USA and France. Lancet 1982; 2: 642-6.

16. Oh TE, Hutchinson R, Short S, Buckley T, Lin E, Leung D. Verification of the acute physiology and chronic health evaluation scoring system in a Hong Kong intensive care unit. Crit Care Med 1993; 21: 698-705.
17. Sirio CA, Tajimi K, Tase C, Knaus WA, Wagner DP, Hirasawa H, et al. An initial comparison of intensive care in Japan and the United States. Crit Care Med 1992; 20: 1207-15.
18. Bion JF, Edlin SA, Ramsay G, McCabe S, Ledingham IM. Validation of a prognostic score in critically ill patients undergoing transport. Br Med J (Clin Res Ed) 1985; 291: 432-4.
19. Wong DT, Crofts SL, Gomez M, McGuire GP, Byrick RJ. Evaluation of predictive ability of APACHE II system and hospital outcome in Canadian intensive care unit patients. Crit Care Med 1995; 23: 1177-83.
20. Lertakyamanee J, Somprakit P, Vorakitpokaton P, Kururattapan SA, Udompunturak S, Pensuk S. APACHE II in a postoperative intensive care unit in Thailand. J Med Assoc Thai 1997; 80: 169-77.
21. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143: 29-36.
22. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983; 148: 839-43.
23. Dragsted L, Jorgensen J, Jensen NH, Bonsing E, Jacobsen E, Knaus WA, et al. Interhospital comparisons of patient outcome from intensive care: importance of lead-time bias. Crit Care Med 1989; 17: 418-22.

การทำนายอัตราการเสียชีวิตโดยใช้ระบบคะแนนมาตรฐานในหออภิบาลอายุรศาสตร์

รณิษฐา รัตนะรัต, มณีรัตน์ ธนกิติวิรุฬ, วรการ วิไลชนม์, สุรัตน์ ทองอยู่, ไชยรัตน์ เพิ่มพิกุล

เพื่อประเมินและเปรียบเทียบความสามารถของระบบคะแนน APACHE II (Acute Physiology and Chronic Health Evaluation II) และระบบคะแนน SAPS (Simplified Acute Physiology Score) ในการทำนายผลการรักษาในหออภิบาลอายุรศาสตร์ คณะผู้วิจัยทบทวนรายงานการรับผู้ป่วยทุกราย (482 ราย) ซึ่งเข้ารับการรักษาในหออภิบาลอายุรศาสตร์ โรงพยาบาลตติยภูมิ แห่งหนึ่งในระยะเวลา 2 ปี ในผู้ป่วยแต่ละรายได้รวบรวมข้อมูล demographic, การวินิจฉัย, คะแนน APACHE II, คะแนน SAPS ใน 24 ชั่วโมงแรก และผลการรักษาในหออภิบาล เปรียบเทียบความสามารถในการทำนายผลการรักษาในหออภิบาลระหว่าง APACHE II และระบบคะแนน SAPS โดย forward stepwise logistic regression และ พื้นที่ใต้ receiver operating characteristic (ROC) curves พบว่าอัตราตายในหออภิบาล (ICU mortality) เท่ากับ 36.93% โดยมีค่าเฉลี่ย APACHE II score และ SAPS score เท่ากับ 21.17 ± 9.35 และ 14.61 ± 6.47ตามลำดับ โดยค่าดังกล่าว ของผู้ป่วยซึ่งเสียชีวิต (APACHEII score 26.97 ± 8.27, SAPS score 18.01 ± 5.84) สูงกว่า ผู้ป่วยซึ่งรอดชีวิต (APACHE II score 17.77 ± 8.22, SAPS score 12.62 ± 5.99) อย่างมีนัยสำคัญทางสถิติ ($p < 0.001$) ความสัมพันธ์ระหว่างระบบทั้งสองดีมาก (ค่าสัมประสิทธิ์ความสัมพันธ์, r เท่ากับ 0.825: $p < 0.001$)การประเมินอัตราตายโดยใช้ APACHE II risk of death equation เท่ากับ 38.98% ความสามารถในการแยกผู้ป่วยซึ่งเสียชีวิตและรอดชีวิตของระบบ APACHE II ดีกว่า SAPS (พื้นที่ใต้ ROC curve ของ APACHE II เท่ากับ 0.788, SAPS เท่ากับ 0.746) โดยสรุป ระบบคะแนน APACHEII เป็นตัวทำนายที่มีประสิทธิภาพในการประเมินผลการรักษาในหออภิบาลและมีความสามารถในการทำนายผลการรักษาในหออภิบาลอายุรศาสตร์ดีกว่าระบบ SAPS