Use of APACHE II Score for Predicting Mortality in Cancer Patients at the National Oncological Institute of Ecuador.

Nidia Patricia Rodríguez Ormaza

Tesis de grado presentada como requisito para obtención del título de Médico General

Quito, Mayo 2008
ABSTRACT

Objective

To evaluate the implementing features and the impact of the APACHE II score and several co-variables on the out-of-hospital survival rate during a 62 month follow-up period of cancer patients admitted to the Intensive Care Unit (ICU).

Design

Prospective observational study.

Setting
Intensive Care Unit at the National Oncological Institute “Dr. Juan Tanca Marengo”. Guayaquil – Ecuador

Patients

393 patients admitted to the Intensive Care Unit during a 62 month period.

Interventions

Each patient was assigned an APACHE II score and designated to either clinical or post-surgical groups. The groups were subclassified according to their final outcome as survivors or non survivors.

Results

Cancer was diagnosed clinically in 103 patients and postoperatively in 290 patients. A total of 79 (20.16%) patients died from cancer complications. The APACHE II score mean was 17.03 ± 7.38 SD; the mean score was 14.70 ± 5.34 SD for survivors and 25.85 ± 7.40 SD for non survivors (p<0.001). The Goodness-of-Fit test provided the following results for mortality rates of patients in the ICU: X² = 12.70 df = 9, p = 0.17. The Standardized Mortality Ratio (SMR) was 0.84 (CI 0.67-1). The Receiver Operating Characteristic (ROC) curve was 0.61 (CI 0.55 - 0.67, SE 0.0031). The predictive power was 24.96% for the expected mortality rates and 20.10% for the observed mortality rates. Mortality rates for cancer groups were the following: solid tumors 40/318 (12.57%), metastatic disease 10/29 (34.48%), and hematologic-oncologic disease 29/35 (82.85%). The out-of-hospital survival rate was 52% after one year and 14.2% after 62 months.

Conclusion

We consider that the APACHE II score is not an ideal predicting index for mortality in cancer patients.
Key Words


Introduction

In general, patients with cancer have a dismal prognosis; therefore, they are not regularly admitted to the intensive care unit (ICU) (1,2). Criteria used to admit these patients are complex and often contradictory. Frequently, there is an unjustified interest to transfer these patients to the ICU, which is related to improper triage or unacceptable curative and palliative expectancies (3) and, thus, to higher morbidity and mortality rates (4). Therefore, treating patients with an advanced disease who do not respond to supportive care is a often questioned decision. This may represent a loss of economic resources and hospital supplies and may raise false expectations on patients’ families, causing emotional and economical stress due to the high cost involved in patients’ care (5,6,7,8,9,10).

The Acute Physiology and Chronic Health Evaluation (APACHE) II score has been validated in several studies performed on different groups of patients in the ICUs. However, in some cases, its use has been declined in patients with coronary heart disease and in trauma units where other scores are used (11). Some studies have evaluated the APACHE II Score’s predictive capacity in cancer patients and have found several results (12). On the other hand, this scoring system should be validated with various groups of patients in order to correct for demographic differences, specific
diagnostic work up, and critical management all of which can lead to differences in the reported results (13).

Considering the previous information, the primary end point of this study is to evaluate the implementing features and the impacts of the APACHE II score (calibration and discrimination). In cancer patients admitted to the National Oncological Institute “Dr. Juán Tanca Marengo” in Guayaquil, Ecuador. The secondary end point is to evaluate the impact of certain covariables on the out-of-hospital survival rate during a 62 month follow-up period of these patients.

Materials and Methods

Patients Characteristics

Patients with cancer diagnosis were prospectively recruited at the ICU of the National Oncological Institute “Dr. Juán Tanca Marengo” in Guayaquil, Ecuador between March 1996 and January 1999, and were followed after discharge until April 2004. The study was approved by the ethics committee of the University Of San Francisco De Quito in Cumbayá, Ecuador and by the National Oncological Institute “Dr. Juán Tanca Mareno” in Guayaquil, Ecuador.

The inclusion criteria used in this study were: age B 18 years, patients with a cytopathologically confirmed cancer diagnosis; patients admitted to the ICU without a cytopathologically confirmed cancer diagnosis (with a highly suggestive clinical, radiological, ultrasonographical or tomographical evidence, which was evaluated afterward by the oncologist and confirmed cytopathologically); patients who stayed in the ICU for > 24 hours; and patients with cancer in remission, recent cancer
diagnosis, and progression or recurrence and who required intensive care. In the case of multiple admissions, only the first one was considered.

The exclusion criteria were: patients admitted to the ICU for monitoring only and patients without clinical suspicion of cancer diagnosis, including patients with cancer diagnosis confirmed cytopathologically after admission.

Variables obtained before patients’ admission to the ICU were the following: time (in months) since cancer diagnosis, if cancer was in remission or diagnosed during the admission period, and if cancer was in progression or relapse. Variables obtained during the admission were as follows: age, sex, pathological diagnosis, and type of malignancy (solid tumors, metastatic tumor, hematologic-oncologic diseases and other classifications). Variables collected after the admission were the following: days of stay in the ICU, APACHE II Score, maximum number of dysfunctional organs during their stay in the ICU, and procedures performed during the stay in the ICU (analgesia, sedation, use of inotropic agents, barbiturates, mechanical ventilation, Cardiopulmonary resuscitation (CPR) maneuvers), and in-hospital and ICU mortality rates.

Patients were classified according to their admission diagnosis as either (clinical or surgical), final in-hospital outcome as (survivors or non-survivors), and type of malignancy (solid tumors, metastatic tumors, hematologic-oncologic cancers and other classifications).

**Statistical Analysis**

All data were expressed as mean ± standard deviation for variables with a normal distribution and as a median ranges for variables with non-normal distribution. Student’s t-test was used for continuous and categorical variables with a normal
distribution, and Chi squared ($X^2$) or Mann-Whitney-U test for variables with non-normal distribution.

To measure the predictive capability of the logistic regression model, the Lemeshow-Hosmer Goodness-of-Fit test was utilized (15,16). This method divides subjects into deciles based on predicted probabilities and then computes a $X^2$ from observed and expected frequencies. To evaluate the model adjustment, predicted results were compared with the observed results in each decile. Values from all cells of the table were added to generate the statistical $H$ test. This test was compared with the $X^2$ distribution (degrees of freedom, $df = 9$) to evaluate the Goodness-of-Fit model.

The number of predicted deaths in each decile corresponded to the addition of patients’ individual death probabilities from each decile. The number of expected deaths was calculated by subtraction of patients’ individual death probabilities from each decile. The statistical $X^2$ is the addition of all deciles’ corresponding values ($\text{Observed-Predicted})^2/\text{Predicted}$ (17,18).

Degrees of freedom were calculated using the following formula: $n-1-k$; where $n$ is the number of deciles and $K$ is the number of variables (in this case, the $n$ is equal to 10 and $K$ is equal to 1). The patients’ in-hospital final outcome was reported in terms of survival. If the Hosmer-Lemeshow Goodness-of-Fit test gave a value lower than the $X^2$ value, it indicated a better calibration of this index. A $p$ value $> 0.05$ validated the model, showing that there are no significant statistical differences between the observed and expected morality values.

The Standardized Mortality Ratio (SMR) was calculated dividing the observed mortality by the predicted mortality. The 95% Confidence Interval (CI) for the SMR was also calculated using the observed mortality as a measure of Poisson (19).
The analysis of discrimination capability compared the calculated values for the Area under the Receiver Operating Characteristic (ROC) curves (20). The highest value corresponded to a higher yield of discrimination. Sensibility, specificity, positive predictive value, negative predictive value, precision, positive likelihood ratio, and negative likelihood ratio were calculated for each of the death probability deciles.

An equation based on a multiple regression logistic model was used to convert the APACHE II Score into an in-hospital death probability (21,22).

**Out-of-Hospital Survival Analysis**

All patients were followed after discharge from the hospital for a period of at least 62 months. For those patients who stopped coming to the hospital, a phone call or home visit was conducted by a physician to inquire regarding the patients’ final outcome (survivor or nonsurvivor). The Kaplan-Meier method was used to determine the patients’ out-of-hospital survival rate (23).

The significant prognostic factors in the univariable model were included in the Cox model to determine the impact of each co-variable (APACHE II score, analgesia, sedation, use of inotropic agents, barbiturates, mechanical ventilation, CPR maneuvers) in the out-of-hospital survival rate (24,25). A p value < 0.05 represented statistical significance.

**Results**

**Sample**
A total of 393 patients were included in the study. During the 3 year recruitment period, there were 651 admissions to the ICU. The following patients were excluded: 23 for being < 18 years old, 140 because their hospital stay was < 24 hours, 85 patients without a cytopathologically confirmed cancer diagnosis or without a highly suggestive clinical, radiological, ultrasonographical or tomographical evidence at the time of admission, and 10 for not completing their hospital stay in the ICU.

**General Patients’ Characteristics**

The age mean (in years) was 55.36 ± 17.36 Standard Deviation (SD). There were 229 (41.7%) women and 164 (58.3%) men. The APACHE II Score mean was 17.03 ± 7.38 SD. 103 (26.2%) patients were admitted with a clinical cancer diagnosis and 290 (73.8%) with a postsurgical cancer diagnosis. Overall patients’ characteristics are shown in Table 1.

**Patients’ Comparisons**

Data for cancer patients diagnosed clinically and post-surgically are compared in Table 2 and data for survivors and non survivors are compared in Table 3.

**APACHE II Score Calibration**

The Goodness-of-Fit test provided the following results for mortality rates of patients in the ICU: $X^2 = 12.70; \ df = 9; \ p = 0.17$.

The predictive power was 24.04% for the expected mortality rates and 20.10% for the observed mortality rates ($p = 0.19$). The SMR was 0.84 (CI 0.67 - 1).
**APACHE II Score Discrimination**

Discrimination of the APACHE II Score expresses each patient’s death probability in relation to their survival probability. The model’s discriminative capacity was considered perfect if the ROC curve = 1, good if the AROC curve > 0.8, moderate if the ROC curve was between 0.6-0.8, and poor if the ROC curve was < 0.6. The ROC curve was 0.61 (CI 0.55 - 0.67, SE 0.0031). The Receiver Operating Characteristic (ROC) curve is shown in Figure 1.

Regarding the discriminative score power, for which the logistic regression equation proposed by Knaus et al. was used, we found low sensitivity and predictive values for death probabilities of 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, and 0.7. As death probabilities increased to 0.8, 0.9, and 1, the sensitivity, the positive predictive value, the negative predictive value, positive and negative likelihood ratios, and the discriminative power of the APACHE II score increased as well. Efficacy Values of the APACHE II Score model are shown in Table 4.

In our study, APACHE II scores > 31 showed a death predicting capacity with the following parameters: sensitivity of 100% (CI 85.86-99.69); specificity of 85.67% (CI 81.54- 89.08); positive predictive value of 36.58% (CI 26.42-48.01); negative predictive value of 100% (CI 98.47-99.97); accuracy of 86.76% (CI 82.91-89.87); and positive likelihood ratio of 6.9 (CI 5.42-8.9).

**ICU Mortality Rates**

The mortality for the clinical group in the ICU was 59.74% (58 out of 103 patients) while in the surgical group was 7.24% (21 out of 90 patients). The overall mortality rate in the ICU was 20.10% (79 out of 393 patients). Mortality rates according
to malignancy groups were as follows: solid tumors 12.57% (40 out of 318 patients), metastatic disease 34.48% (10 out of 29), and hematologic-oncologic disease 82.85% (29 out of 35 patients). APACHE II Score death probabilities are shown in Table 5.

**Out-of-Hospital Survival Rates**

Out-of-hospital survival rate was 52% after one year of follow-up and 14.2% after 2 months. These results are shown in Figure 2.

**Proportional Hazard Model**

By using the proportional hazard model (Cox’s model), it was found that the type of malignancy (Odds Ratio (OR) 1.39. CI: 0.39 – 4.8), the use of inotropic agents (OR 1.98. CI: 0.68 – 5.7), and the utilization of invasive mechanical ventilation (OR 1.17. CI: 0.37 – 3.6) were significant. The results of the Proportional Hazard Model for these variables are shown in Table 6.

**Discussion**

In evaluating the quality of care, treatment efficacy and the efficient use of resources in the ICU, it is necessary to standardize the disease severity. The APACHE II score has been chosen as a mortality rate indicator because it fits two basic requirements for a severity index: it uses objective data, and it is simple to calculate. This guarantees the feasibility to compare the results of this study with other observations (26).

To evaluate the predictive power of a mortality index, it is necessary to consider two basic aspects: calibration, meaning the accuracy of the model’s probability predictions, and discrimination, meaning the capacity of the model to discriminate survivors from non survivors. In relation to the discriminative capacity, this study showed an ROC curve of 0.61 (0.40 –0.67). Several authors have reported different
results when evaluating cancer patients. Giangiuliani et al. (27) found an ROC curve of 0.54. Sculier, (28) when studying 261 patients, found an ROC curve of 0.60. Schellongowski (12) et al, reported an ROC curve of 0.77 (CI 0.70-0.82) when studying 242 patients. Berghmans (29) evaluated 247 patients and found an ROC curve of 0.65. Benoit (30) in 124 patients with haemathologic-oncologic diseases found an ROC curve of 0.71 (Standard Error (SE): 0.043). In addition, Soares (31) reported an ROC curve of 0.88 (CI 0.86-0.90) when studying 1257 patients; and when the author excluded patients who stayed in the ICU after a previously planned surgical procedures, he reported ROC curve of 0.75 (SE: 0.02. (CI: 0.71-0.79) in 542 patients.

To increase the discriminative capacity, other variables related to oncologic diseases should be taken into account. These variables were reported by Groeger et al., who found a ROC curve value of 0.81 when evaluating the Probability of Mortality Model (PMM) in cancer patients (32). Nevertheless, Schellongowski (12) found a lower discriminative capacity of the ROC curve of 0.70 (CI: 0.63-0.76) when compared with the APACHE II Score ROC curve of 0.77 (CI: 0.70-0.82). These differences in the discriminative capacity of the scores are probably due to the following reasons: the quality of care in the ICU, the study’s sample size, the number of centers involved, and the changes in the therapeutics approaches throughout the time of the study (41,42).

According to the results provided by the Goodness-of-Fit Test and the Hosmer–Lemeshow–
Hosmer X² statistic, we observed that there is no significant difference between the predicted and the observed mortality rates. Our results for the Hosmer-Lomenshow were $X^2 = 12.70 \ df = 9, \ p = 0.17$. A number of authors have reported different results when using the Goodness – of – fit test to evaluate the APACHE II Score. Sculier (28)
analyzed 261 patients with solid tumors and hemathologic-oncologic diseases and found a $X^2$ statistic of 52.95 (df 9; $p < 0.001$). Berghmans (29) studied 247 patients with solid tumors and hemathologic-oncologic diseases and found $X^2 = 18.89$ (df 5, $p = 0.002$). Schellongowski et al. (12), found in 242 patients a Goodness-of-Fit test result of $X^2 = 15.03$ (df 8, $p = 0.066$). Benoit (30) in 124 patients with hemathologic-oncologic diseases reported a Goodness-of-Fit test result of $X^2 = 5.12$ (df 5, $p = 0.39$). Soares (31) described a Goodness–of–Fit test of 78.18 (df 8. $p < 0.001$) when studying 1257 patients. This discrepancy in the reported results may be due to the influence of the sample size since a small sample can result in an overestimation of the Goodness–of–Fit test results (33).

The expected mortality rate was 24.04% and the observed mortality rate was 20.10% ($p = 0.19$). The low mortality percentage in our group of patients could be due to the high number of post-surgical cancer patients included in this study. The mortality rate in the group of clinical cancer patients (59.74%), does not differ from the results reported by other authors (12,28,30,32,34,35).

Our results suggest an increase in the survival rate of surgical patients when compared to those admitted with clinically diagnosed cancers. This may be due to a better condition of patients admitted to elective surgeries. These results suggest that this group of patients could benefit the most if they are admitted to the ICU after undergoing surgery to treat their cancer (36).

The mortality rate also differs based on the different malignancy groups. We found a low mortality in patients with solid tumors in comparison with those with hematologic-oncologic diseases and metastatic cancers. These results were found to be similar to previous reported data (28,32).
Groeger et al. (32), in a prospective multicentric study carried out in 5 hospitals, analyzed 782 patients, and found a mortality rate of 76% associated to acute respiratory failure without significant differences between the centers. Additionally, other authors also reported poor results associated to the use of mechanical ventilation (28,30,31).

In our study, the Standardized Mortality Ratio (SMR) (observed/expected ratio) was found to be 0.84. Giangiuliani et al., when studying 152 patients with lung cancer and high or low surgical risk, obtained a SMR of 0.94 (37). Headley et al. retrospectively evaluated 52 patients with breast cancer and found a SMR of 1.27 (38). Sculier (28) reported a SMR of 1.25. Berghmans (29) reported a SMR of 0.93. Schellongowski, (12) reported a SMR of 1.05. Affesa et al reported a SMR of 1.03 (CI 0.77-1.36). Soares (31) described a SMR of 1.41 (CI 1.22- 1.62).

Similar to other authors, we also planned to evaluate the predictive capacity of the APACHE II score at a cross-sectional point. This study showed that the APACHE II score demonstrated a predicting capacity for oncologic patients with a score > 31. This result differs from other findings reported by Giangiuliani et al. (37), and Headley et al. (38) who described the presence of predicting power for APACHE II scores of > 23 and 35 respectively. Fakhry et al. showed that higher APACHE II scores correlated to a higher death probability. In this study for an APACHE II score > 20 the death probability was 78% (39). Sculier (28) observed a mortality of 86% with an APACHE II predictive mortality score > 70%. Berghmans (29) reported a mortality rate of 78% with an APACHE II score > 60.

By using the equation for multiple logistic regression proposed by Knaus et al., the
APACHE II score proved to be a strong predictor for mortality only in groups with high
death probability of 0.8, 0.9 and 1. However, in groups with low and intermediate death
probabilities, a weaker correlation with their final outcome was found.

Several other predicting factors should be taken into account when analyzing
these results. Groeger et al., found that patients with progression of malignancy and
associated severe respiratory failure, had approximately twice the probability of
dying than patients without these factors. Thus, they found that an acute disease and/or
major organ dysfunctions were accompanied by disseminated intravascular coagulation,
cardiac arrhythmias and the need of vasopressors. Patients with heart failure had twice
the probability of dying, and those with hematologic failure had four times the
probability of dying compared to those without these factors (32).

In the analysis of the proportional hazard model we found that the presence of
mechanical ventilation, the use of vasopressors, and the type of the malignancy clearly
influenced the patient’s survival. However, other variables such as multiple organ
dysfunction (40), maximum or delta physiology score after admission, and
inflammatory cytokines, should be considered in posterior studies.

We found a cumulative survival rate of 52% after a year, and 14.2% after 62
months. Staudinger (35) found 77% of mortality at 1 year, and Sculier (28) found a 23%
survival rate after the same period of time.

Numerous co-variables could be implicated in our reported survival; among
those are the presence of surgical patients and non-analyzed factors (neoadjuvant cancer
treatments and comorbidities) after discharge from the ICU (43). Future studies should
consider these factors in their analysis.

Limitations of this study were due to failure of implementing a predicting score
in another group of patients with similar characteristics. Our work included patients
with a severe clinical condition and patients with good prognosis after undergoing elective surgery. Another source of bias could come from the great number of patients transferred from elective surgery, who were mostly elderly patients (44). These considerations should be considered when extrapolating the results.

**Conclusion**

In summary, even though the APACHE II score showed concordance between expected and observed mortality rates, we consider that this is not the ideal index for predicting mortality in cancer patients. This may be related to the fact that in our series we were only able to predict mortality in the lower end and the intermediate groups of patients. Nonetheless, a cutoff point > 31 in the APACHE II score may offer relevant information. This fact could be useful in the decision making process in particular situations such as the identification of cancer patients that will not benefit from their stay in the ICU. Hence, this resolution should always go along with a solid clinical judgment, an objective interpretation of the results, as well as the assessment of the patients and their families’ previous wishes.
References


7. Zimmerman JE, Knaus WA, Steven M, Sharpe AA, Draper EA, Wagner DP. The Use


43. Soares M, Salluh JL, Ferreira C, Luiz R, Spector N, Rocco J. Impact of to different
comorbidity measures on the 6 – month mortality of critically ill cancer patients.
