OP 2-5

DEVELOPMENT AND VALIDATION OF A DEEP LEARNING MODEL FOR THE PREDICTION OF HEPATOCELLULAR CANCER RECURRENCE AFTER TRANSPLANTATION: AN INTERNATIONAL STUDY

Quirino LAI*[1], Karim HALAZUN[2], Prashant BHANGUI[3], Yuji SOEJIMA[4], Armin FINKENSTEDT[5], Shinji UEMOTO[6], Chung Mau LO[7], Chao-Long CHEN[8], Umberto CILLO[9], Jan LERUT[10]

[1]*General Surgery Department, Sapienza University, ITALY*
[2]*General Surgery Department, Columbia University New York, UNITED STATES OF AMERICA*
[3]*General Surgery Department, Medanta New Delhi, INDIA*
[4]*General Surgery Department, Kyushu University, JAPAN*
[5]*Medicine Department, Innsbruck University, AUSTRIA*
[6]*General Surgery Department, Kyoto University, JAPAN*
[7]*General Surgery Department, Hong Kong University, HONG KONG*
[8]*General Surgery Department, Kaohsiung Taiwan, TAIWAN*
[9]*General Surgery Department, Padua University, ITALY*
[10]*General Surgery Department, UCL Brussels, BELGIUM*

Background : Identifying patients at high risk for hepatocellular carcinoma (HCC) recurrence after liver transplantation (LT) represents a challenging issue. The present study aims at developing an accurate post-LT recurrence prediction calculator using the machine learning method (Time_Radiological-response_Alpha-fetoproteIN_Artificial-Intelligence, TRAIN-AI).

Methods : 3,381 patients with HCC listed for LT from 2000 to 2018 and coming from 17 centers from North America, Europe, and Asia were included in the study. The original dataset was split to generate the two main data sets used for the research. The Training Set was composed of 70% of the records of the original dataset, and the Test Set was composed by the remaining 30%. Using the Training Set data, a prognostic model for HCC recurrence was developed with a Deep Surv model, and a Cox proportional hazards deep neural network was constructed. Validation of the model was done using the Test Set. The TRAIN-AI was compared using the DeLong test with Metroticket 2.0 Score, AFP-French Model, Milan Criteria, San Francisco Criteria, Up-to-Seven Criteria, TRAIN Score, NYCA Score, and HALT-HCC Score.

Results : The developed TRAIN-AI model showed an escellent c-statistics, with an AUC=0.78 (95%CI=0.73-0.82). The TRAIN-AI always outperformed the other scores: Metroticket 2.0 Score AUC=0.66, P<0.0001; AFP-French Model AUC=0.65, P<0.0001; Milan Criteria AUC=0.63, P<0.0001; San Francisco Criteria AUC=0.61, P<0.0001; Up-to-Seven Criteria AUC=0.60, P<0.0001; TRAIN Score AUC=0.59, P<0.0001; NYCA Score AUC=0.58, P<0.0001; HALT-HCC Score AUC=0.57, P<0.0001.

Conclusions : The proposed TRAIN-AI score showed higher accuracy than other available risk scores in terms of post-LT recurrence risk. Further validation is required. A web calculator has been developed for improving the user-friendly availability of the model.

Corresponding Author : **Quirino LAI** (lai.quirino@libero.it)