Seminars in Oncology 45 (2018) 264–274

Contents lists available at ScienceDirect

Seminars in Oncology

journal homepage: www.elsevier.com/locate/seminoncol

Risk score system for the prediction of hepatocellular carcinoma in patients with type 2 diabetes: Taiwan Diabetes Study

Tsai-Chung Li[a,b], Chia-Ing Li[c,d], Chiu-Shong Liu[c,d,e], Wen-Yuan Lin[c,e], Chih-Hsueh Lin[c,e], Sing-Yu Yang[a], Cheng-Chieh Lin[c,d,e,*]

[a] Department of Public Health, College of Public Health, China Medical University, Taichung, Taiwan
[b] Department of Healthcare Administration, College of Medical and Health Science, Asia University, Taichung, Taiwan
[c] School of Medicine, College of Medicine, China Medical University, Taichung, Taiwan
[d] Department of Medical Research, China Medical University Hospital, Taichung, Taiwan
[e] Department of Family Medicine, China Medical University Hospital, Taichung, Taiwan

ARTICLE INFO

Article history:
Received 20 November 2017
Revised 17 July 2018
Accepted 18 July 2018

Keywords:
Hepatocellular carcinoma
Type 2 diabetes
Prediction model

ABSTRACT

This study aims to develop a risk score system for hepatocellular carcinoma (HCC) in patients with type 2 diabetes using the Taiwan National Diabetes Care Management Program database. This retrospective cohort study included 31,723 Chinese patients who had type 2 diabetes, aged 30–84 years. Participants were randomly grouped into derivation and validation sets in 2:1 ratio. Cox proportional hazard regression models were used to identify the risk factors of HCC in the derivation set. Discrimination ability of the model was assessed by means of a receiver operating characteristic curve and performance was expressed as the c statistic, assessed internally on validation data sets. The average follow-up was 8.33 years with 748 HCC incident cases in the derivation set. The final HCC risk score system included age (−2 to 8 points), gender (0–2 points), smoking (0–2 points), variation in hemoglobin A1c (0–1 point), serum glutamic–pyruvic transaminase (0–6 points), liver cirrhosis (9 points), hepatitis B (4 points), hepatitis C (3 points), antidiabetes medications (0–3 points), and antihyperlipidemia medications and total/high-density lipoprotein cholesterol ratio (−4 to 2 points). The HCC risk score was the sum of these individual scores (range −6 to 40). The area under the receiver operating characteristic curve for 3-, 5-, and 10-year HCC risks was 0.81, 0.80, and 0.77 for the derivation set, respectively. This HCC risk score system has good prediction accuracy and discriminatory ability, and serves a simple tool for HCC risk prediction.

Introduction

Hepatocellular carcinoma (HCC) is the fourth leading cause of cancer mortality worldwide. Based on reports from the International Agency for Research on Cancer, in 2012 HCC was the fifth most common cancer for men (554,369 cases, 7.5% of total cases) and the ninth for women (228,082 cases, 3.4% of total cases) [1]. Eastern Asia has the highest incidence of HCC. In 2012 Mongolia, Lao People's Democratic Republic, Vietnam, Korea, Thailand, and China had the highest HCC incidence rates in the world [1] http://bmccomplementalternmed.biomedcentral.com/articles/10.1186/1472-6882-12-146—CR1. In Taiwan, HCC accounted for 17.5% of all cancer deaths and in 2016 ranked as the second most common cause of cancer deaths for men and women [2].

The worldwide prevalence of diabetes is increasing, and its prevalence and incidence rates have increased in various age, gender, and race groups. Based on the world standard population reported by the World Health Organization, the global prevalence of diabetes increased from 4.7% in 1980 to 8.5% in 2014 [3]. The epidemiologic findings of recent cohort and case–control studies have indicated a link between type 2 diabetes and cancer in several gastrointestinal organs, especially the liver [4-7]. Dysregulation of the tuberous sclerosis 1/tuberous sclerosis 2/mTOR signaling pathway by IκB kinaseβ [8-11] supports an independent contribution of diabetes as a risk factor for cancer.

Previous studies have established prediction models for HCC for the general population [12-14], for patients with type 2 diabetes [15] but without chronic viral hepatitis or alcoholic cirrhosis [16], for patients with chronic hepatitis B [17-24], for patients with chronic hepatitis C and/or cirrhosis [25-32], and for patients

* Corresponding author. China Medical University, 91 Hsueh-Shih Road, Taichung 40421, Taiwan.
E-mail address: cclin@mail.cmuh.org.tw (C.-C. Lin).

https://doi.org/10.1053/j.seminoncol.2018.07.006
0093-7754/

with chronic liver diseases [33]. These studies are summarized in Table 1. Only 2 studies have been conducted in patients with type 2 diabetes [15,16]. One is a case–control study using an artificial neural network in Chinese individuals with type 2 diabetes [15]. This study did not consider biomarkers, such as fasting plasma glucose, hemoglobin A1c (HbA1c), total cholesterol, triglyceride, high-density lipoprotein (HDL), or low-density lipoprotein (LDL), that are routinely measured for diabetes care. The authors of this study reported its discriminatory ability but did not develop a point system that can quantify the effect of each risk factor. In contrast, the other study developed a point system; however, the sample size was small, and γ-glutamyl transferase was incorporated as a biomarker, a value not measured in the routine follow-up of individuals with diabetes [16]. Thus, a prediction model of HCC risk in patients with type 2 diabetes must be developed by addressing the deficiencies of previous models. The present study aimed to develop a risk score system for HCC in patients with type 2 diabetes using the Taiwan National Diabetes Care Management Program (NDCMP) database to provide information that may help design studies assessing preventive interventions and serve as a benchmark for testing novel putative risk factors.

Methods

Data source

The NDCMP, which was established by the National Health Insurance (NHI) Program in 2001, is a nurse case management program for patients with type 1 and type 2 diabetes. A retrospective cohort study was conducted in patients with type 2 diabetes in the NDCMP from 2001 to 2004. The date of entry for NDCMP was defined as the index date. Patients were monitored for withdrawal from the NHI Program, death, development of HCC, or the end of follow-up until December 31, 2011. We used NDCMP and NHIRD databases to construct a cohort of patients with type 2 diabetes. Each person in Taiwan has a unique personal identification number (PIN). For security and privacy purposes, data on patient identities were scrambled cryptographically. Each patient has 1 PIN for NDCMP and NHIRD. All NHI datasets can be interlinked using scrambled PINs for each patient. We combined the NDCMP and NHIRD datasets, including details of ambulatory care orders of NHIRD from 2001 to 2004. This approach enabled us to acquire information on baseline characteristics, including sociodemographic factors, diabetes-related factors and biomarkers, comorbidity, types of antidiabetes medications, etc.

Study subjects

The Taiwan Diabetes Study includes all patients with type 2 diabetes in the NDCMP from 2001 to 2004. A total of 63,084 enrolled patients were initially diagnosed with diabetes based on the American Diabetes Association criteria (International Classification Disease, Ninth Revision, Clinical Modification, ICD-9-CM; code 250). The main goal of NDCMP is to increase the quality of diabetes care by increasing the frequency of monitoring, providing continuity of care, and decreasing diabetes-related complications. Patients included had to have had at least 1 year of follow-up for calculation of visit-to-visit variation in HbA1c and FPG; be free of HCC at baseline; and must not have had missing information regarding baseline characteristics, comorbidities, and biomarkers. Fig. 1 shows the flowchart of recruitment procedure. We excluded people who had type 1 diabetes (ICD-9-CM; code 250.x1/x3) and gestational diabetes ($n = 2{,}108$), those who were under 30 or above 85 years of age ($n = 1{,}025$), those who had a diagnosis of HCC at baseline ($n = 198$), and those who were diagnosed with other cancers at baseline ($n = 1{,}520$). Participants with missing data on sociodemographic factors, lifestyle behavior, blood biomarkers, and less than 1 year of follow-up ($n = 26{,}510$) were also excluded. Finally, 31,723 participants (14,826 men and 16,897 women) were included. These participants were randomly assigned to a derivation set ($n = 21{,}149$) and a validation set ($n = 10{,}574$) in a 2:1 ratio. Ethics approval was obtained from the Ethical Review Board of China Medical University Hospital.

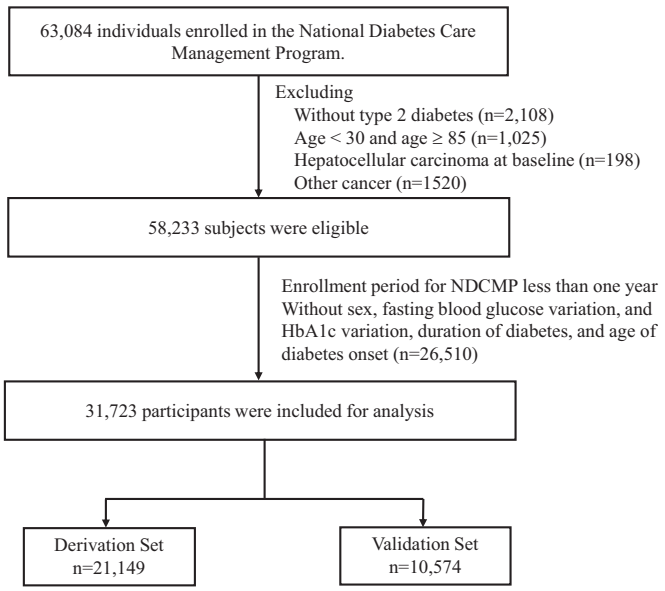

Fig. 1. Flowchart for recruitment procedures of the predictive model for HCC.

Outcome ascertainment

The primary outcome measure was HCC, determined through record linkage with ambulatory and inpatient care data in the NHIRD. The time of follow-up began with the index date and ended with a new diagnosis of HCC, death, withdrawal from insurance program, or end of follow-up in December 31, 2011. HCC incident was ascertained through ICD-9-CM. The ICD-9-CM code for HCC is 155. HCC cases had to satisfy at least one of the following criteria: at least 3 ambulatory claims or 1 inpatient care claim.

Covariates

Baseline data on other chronic medical conditions were defined based on information at the 24-month period preceding cohort entry by using outpatient and inpatient claim data. A 24-month period was adopted because some chronic medical conditions are not common. A long period for characterization was adopted to ensure that their diagnosis was not missed. Histories of acute pancreatitis (ICD-9-CM code 557.0), chronic pancreatitis (ICD-9-CM code 557.1), alcoholic liver damage (ICD-9-CM codes 571.0, 571.1, and 571.3), nonalcoholic fatty liver diseases (ICD-9-CM code 571.8), liver cirrhosis (ICD-9-CM codes 571.2, 571.5, and 571.6), cholelithiasis (ICD-9-CM code 574), alcohol dependence syndrome (ICD-9-CM code 303), pancreatic pseudocysts (ICD-9-CM code 577.2), jaundice (ICD-9-CM code 782.4), chronic hepatitis B virus infection (ICD-9-CM codes 070.2, 070.3, and V02.61), chronic hepatitis C virus infection (ICD-9-CM codes 070.41, 070.44, 070.51, 070.54, and V02.62), cholecystitis (ICD-9-CM codes 575.0 and 575.1), cholangitis (ICD-9-CM code 576.1), gastric ulcer (ICD-9-CM code 531), and duodenal ulcer (ICD-9-CM codes 532) were identified as comorbidities before the index date.

Lifestyle behavior included smoking (yes, no), alcohol consumption (yes, no), duration of diabetes (continuous), body mass index

Table 1
HCC prediction models for general population, and patients with chronic hepatitis B, chronic hepatitis C, cirrhosis, chronic liver diseases, or type 2 diabetes.

• Author • Year (Ref) • Abbreviated name	• Population • Age range • N	Country	Risk score (range) or prediction equation	Variables in the prediction model (score)	Area under ROC or C statistics	Comments
• Hung • 2015 [12]	• HBsAg-positive and healthy individuals[a] • 20–80 • 12,377	Taiwan	Regression coefficient (0–28)	Age ≥60 (7), male (3), ALT ≥25 U/L (3), previous chronic liver disease (3), family history of HCC (3), HbsAg-positive or anti-HCV-positive (9), and cumulative smoking (1)	c-statistics' range 0.76–0.83 for 3-, 5-, and 10-year	This pooled analysis used data from 3 cohorts involving high-risk subjects characterized by male gender, HbsAg-positivity, and HCC family history. Detection bias for HCC may exist because patients with HBV and HCV had to be cared regularly and hence had a higher probability of being correctly diagnosed with HCC. The high-risk sample may be useful to explore the determinants of HCC, but the prediction model based on this type of sample may not be generalizable to the general population. Discriminant ability and calibration had been assessed.
• Michikawa • 2012 [13]	• Individuals attending health check ups[a] • 40–69 • 17,654	Japan	Algorithm proposed by Framingham heart study (−1 to 19)	Age ≥60 (3), men (2), alcohol consumption ≥450 g/week ethanol (2), BMI ≥25 kg/m^2 (1), diabetes (1), coffee consumption ≥1 cup/day (−1), hepatitis B (4), and hepatitis C (6)	c-index = 0.94 (10-year)	Among participants attending health checkups, 37% of them who provided blood sample were included in data analysis. There was a potential selection bias if differential characteristics existed between subjects who did and did not provide blood samples. It is likely to have detection bias for HCC because patients with HBV and HCV were more intensively monitored and hence had a higher chance of being correctly diagnosed with HCC. It has been reported that persons who seek health checkups are different from general population in sociodemographic factors and health status, thus this prediction model may not be generalizable to the general population. Discriminant ability was assessed, but calibration was not.
• Wen • 2012 [14]	• Individuals attending a private health screening firm[a] • ≥20 • 428,584	Taiwan	Algorithm proposed by Framingham heart study (1-39)	Age ≥60 (6), male (1), smoking ≥10 pack-year (1), drinking (1), physical activity (0), diabetes (1), HBV (6), HCV (5), ALT ≥25 IU/L (2), AST ≥60 IU/L (7), AFP levels ≥10 ng/mL (9)	Area under ROC=0.93 (10-year)	There was a potential selection bias because participants were from a private health screening firm and their socioeconomic status was above the average of general population. A potential detection bias for HCC is possible because those with HBV and HCV had a higher likelihood of seeking care. The findings may not be generalizable to the general population. Discriminant ability and calibration had been assessed.
• Rau • 2016 [15]	• Type 2 DM patients[a] • 20–90 • 2,060	Taiwan	Prediction model from artificial neural network and logistic model	Age (0.015), gender (0.153), alcoholic cirrhosis (−0.804), other cirrhosis (2.401), alcoholic hepatitis (0.108), viral hepatitis (0.740), other types of chronic hepatitis (0.713), alcoholic fatty liver disease (0.101), other types of fatty liver disease (−0.495), and hyperlipidemia (−1.511)	Area under ROC = 0.873 (6-year)	Study population were persons who were newly diagnosed with type 2 diabetes with no prior cancer diagnosis identified from a random sample of the entire Taiwan population. During follow-up period, patients with cancer were defined as cases and matched control were those without cancer. Because random sampling strategy was adopted to select the random sample, case and control groups are representative of case and control populations. Thus, a selection bias was less likely. The study results were more generalizable because it is a nationwide sample. This study did not provide a risk score system and did not consider diabetes-related variables such BMI, glucose control, and antidiabetes medication. Internal validity was performed, but not external validation. Discriminant ability was evaluated, but not calibration. The discriminant ability was good.
• Si • 2016 [16] • DM-HCC	• Type 2 DM patients[a] • ≥18 • 3,544	Korea	Prediction model from Cox's model (0–33)	Age >65 (11), triglyceride level <150 mg/dL (6) and high γ-glutamyl transferase level >80 IU/L (16)	Area under ROC = 0.86 (10-year)	This study included patients with diabetes but no chronic viral hepatitis. Thus, it can rule out the confounding of chronic viral hepatitis. The study subjects were from a single center, thus selection bias is likely to occur. Because selective factors may be associated with patients seeking care in this center, the study results may not be generalizable to other clinical settings. Both clinical and routine biomarkers were considered. Internal validity was evaluated, but not external validation. Discriminant ability was assessed but not calibration.

(*continued on next page*)

Table 1 (*continued*)

• Author • Year (Ref) • Abbreviated name	• Population • Age range • N	Country	Risk score (range) or prediction equation	Variables in the prediction model (score)	Area under ROC or C statistics	Comments
• Han • 2005 [17]	• Patients with chronic hepatitis infection, liver cirrhosis, and carrier[c] • N/A • 4,339	Korea	Prediction equation estimated by logistic regression analysis	Age over 40 years (1.3145), man (0.3), heavy alcoholics (0.5840), AFP >20 ng/ml (0.8257), ALT >40 IU/L (0.2830), liver cirrhosis (1.7219), chronic HCV infection (1.2631), chronic HBV infection (0.7754), chronic hepatitis (0.7339)		The prediction equation was based on a sample of individuals at high risk, including chronic hepatitis infection, liver cirrhosis, and carriers. Thus, the results can be generalizable to persons at high risks, but not to the general population. This study provided a risk index calculated by a prediction equation derived from logistic regression analysis and it did not provide a risk score system. No internal validation assessment, including discriminant ability for risk score and calibration had been done. C-statistic was not reported.
• Lee • 2013 [18]	• Patients with HBV[a] • 30–65, • 3,342	Taiwan	Algorithm proposed by Framingham heart study (0–14)	Age (each 5 years increment) (1), male (2), ALT ≥45 IU/L (2), family history of HCC (2), HBeAg/HBV DNA/HbsAg/genotype (7)	Area under ROC = 0.84–0.87 (5-, 10-, 15-year in validation set)	The risk score was derived from a sample of REVEAL-HBV database. Detailed HBV genotype, HBsAg, and HBeAg were measured and were the most important determinants for HCC. There was a similar problem as that by Yang et al [23], the estimated effects of these determinants may not be generalizable to the general population because their effects estimated in patients with HBV infection may not be the same as those estimated in general population. These measurements are not feasible in the general population. Discriminant ability was assessed, but calibration was not.
• Lin • 2013 [19]	• Patients with HBV[a] • >40 • 1,882	Taiwan	Algorithm proposed by Framingham heart study (0–33)	Age (5-year increments) (1), male (1), ALT ≥28 IU/L (3), AAR≥1 (6), AFP ≥5 ng/mL (4), GGT ≥41 U/L (4), albumin ≤4.1 g/dL (3), Alpha-1 globulin ≤0.2 g/dL (2), HBeAg positive (5) or HBeAg negative and HBV DNA >10,000 (3)	Area under ROC = 0.91 (6-year)	Similar to the study conducted by Lee et al [18], the risk score was derived from a sample of participants with HBV infection in the REVEAL-HBV database and liver-related seromarkers were the key predictors. The study sample was a subcohort of participants with HBV infection identified from a community-based cohort. Thus, there was a potential selection bias if differential characteristics existed between those who were and were not included in the analysis. The findings were limited to patients with HBV infection. Discriminant ability was assessed, but calibration was not.
• Sinn • 2017 [20]	• Patients with HBV and normal ALT[b] • >18 • 918 for derivation; 507 for EV	Korea	Prediction equation	Log (HBV DNA IU/mL) (2.9325), log (HBV DNA IU/mL)^2 (−0.10527), age, yr (0.07013), sex (−1.27223)	Area under ROC = 0.89/0.88 (3/5-year)	All patients with HBV infection were identified from a medical center, which was very likely to result in selection bias and limit its generalizability. The findings were limited to patients with HBV infection. The model considered the nonlinear relationship of HBV DNA, ie, quadratic term of log HBV DNA. No risk score system was provided. Discriminant ability was assessed, but calibration was not.
• Wang • 2014 [21]	• Patients with HBV[a] • NA • 1,555	Hong Kong	Risk score from Cox's model (0–30)	Age >50 (10), albumin ≤35 g/L (1), HBV DNA >200,000 UL/ml (5), liver stiffness measurement >12 kPa (14)	Area under ROC = 0.83 (5-year)	All patients with HBV infection were referred to a hospital, thus selection bias was likely, limiting its generalizability. The model considered a novel predictor, liver stiffness measurement, to improve the original risk score, CU-HCC. The findings were limited to patients with HBV infection. Discriminant ability was assessed, but calibration was not.
• Wang • 2010 [22]	• Chronic hepatitis B carriers[b] • 40–70 • 1,005	Hong Kong	Risk score from Cox's model (0–44.5)	Age >50 (3), albumin ≤35 g/L (20), bilirubin >18 μmol/L (1.5), log(HBV DNA) >6 (4), and cirrhosis (15)	Area under ROC = 0.76, 0.78 (5-, 10-year)	Patients in the derivation cohort were identified by a surveillance program of a hospital, and patients in the external cohort were recruited from the same hospital during a different time period. A selection bias was likely if there was a selective factor associated with seeking care in this hospital, which may limit its generalizability. Discriminant ability was assessed, but calibration was not.

(*continued on next page*)

(body weight divided by height^2, kg/m^2), and early onset of diabetes (yes if age of onset >45 years, no if ≤45 years). Blood biochemical indices consisted of HbA1c, fasting plasma glucose (FPG), triglycerides, HDL, LDL, serum glutamic-pyruvic transaminase, and chronic kidney disease (yes if the estimated glomerular filtration rate <60 [mL/min/1.73 m^2], no if ≥60). The coefficient of variation (CV) for HbA1c or FPG measurements from outpatient visits within the first year of the index date for each patient was calculated for those who had more than 2 HbA1c or FPG measurements in the first year. The CV of HbA1c or FPG was divided by the square root of the ratio of total visits divided by total visits minus 1 to adjust for the possibility that the number of visits might affect the variation of measurements [34].

Data on the use of medications prescribed for the treatment of diseases were calculated for the 12-month period preceding the cohort entry from outpatient prescriptions. A patient was defined

Table 1 (*continued*)

• Author • Year (Ref) • Abbreviated name	• Population • Age range • N	Country	Risk score (range) or prediction equation	Variables in the prediction model (score)	Area under ROC or C statistics	Comments
• Yang • 2011 [23]	• Patients with HBV[b] • Development cohort: 30–65 • 3,548	Taiwan, Hong Kong, Korea	Algorithm proposed by Framingham heart study (0–17)	Age per 5 years (1), male (2), ALT ≥45 U/L (2), HBeAg positive (2), HBV DNA Level of 100,000–999,999 copies per mL (5)	Area under ROC = 0.769–0.902 (3-, 5-, 10-year)	This study cohort had the largest sample size for HCC prediction model in patients with HBV. Patients in the derivation cohort were identified by REVEAL-HBV database, including a sample of participants with seropositive HBsAg at entry. The validation cohort was patients from 3 independent university hospital databases (2 in Hong Kong, and 1 in Korea). The estimated effects of these determinants may not be generalized to the general population because their effects were estimated in patients with HBV infection and this may not be the same as that estimated in a general population. These sero-HBV-related measurements are not routinely collected clinically in the general population because it is not cost-effective due to low prevalence of HBV infection in the general population. Thus, it is not feasible to have these measurements in the general population. A selection bias was likely to occur because the original cohort was not a community-based population, which may limit its generalizability. Both discriminant ability and calibration were assessed.
• Yuen • 2009 [24]	• Patients with HBV[d] • 30–65 • 820	Hong Kong,	Risk score from Cox's model	Age in years (1), male (16), log (HBV DNA levels copies/mL), core promoter mutations-mutant (19) and cirrhosis (30)	Area under ROC = 0.88, 0.89 (5-, 10-year)	The risk score was based on a sample of participants with positive HBsAg at entry in a hospital with a relative smaller size. The first prediction model considered HBV DNA level, HBV genotypes, and core promoter/precore mutations. Due to selecting participants from 1 hospital, a potential selection bias is likely to occur. In addition, the generalizability of the study finding was limited to those with HBV. Discriminant ability and calibration of the risk score were not assessed.
• Chang • 2012 [25]	• Patients with HCV after sustained response to PIR combined therapy[c] • 20–83 • 871	Taiwan	Risk score from Cox's model (0–19)	Age ≥60 (5), platelets $<150\times10^9$/L (4), AFP ≥20 ng/mL (4), fibrosis F3–F4 (6)	Area under ROC = 0.85 (8-year)	All patients with HCV infection receiving combined peg-IFN/ribavirin therapy were enrolled from a medical center, which was very likely to result in selection bias. The study results can be generalized to those with HCV after sustained response to PIR combined therapy, but not all patients with HCV. The model considered the HCV RNA and genotype (1 vs. non-1), but these variables were not significant. Discriminant ability was assessed in the same sample, but calibration was not conducted.
• Chang • 2013 [26]	• Patients with HCV after interferon-based therapy[a] • N/A • $n = 1{,}879$	Taiwan	Risk score from Cox's model (0–9)	Age ≥60 (1), male (1), platelets $<150\times10^9$/L (1), AFP ≥20 ng/mL (1), fibrosis F3-F4 (2), HCV genotype-G1b (1), sustained virological respond (2)	Area under ROC = 0.79 (10-year)	All patients with HCV infection receiving IFN or peg-IFN plus ribavirin therapies were enrolled from a medical center, thus a selection bias is likely because it is a 1-hospital-based study. The study results can only be generalized to those with HCV receiving interferon-based therapy. The model considered the HCV RNA and genotype (G1b vs. non-G1b), and only HCV genotype was significant. Discriminant ability was assessed, but calibration was not conducted.
• El-Serag • 2014 [27]	• Patients with HCV and cirrhosis[a] • N/A • $n = 11{,}721$	US	Prediction model with logistic regression model	Age, AFP, ALT, platelets (not reported)		This study included patients with HCV infection and cirrhosis registered in the VA HCV Clinical Case Registry, thus a selection bias is less likely compared with those studies conducted in a single clinical setting. Some selective factors may be associated with patients seeking care in VA hospitals, thus the study results can only be generalized to those with HCV and cirrhosis in VA hospitals. Only clinical and routine biomarkers were considered, but not HCV biomarkers. This study did not provide a risk score system. Both discriminant ability and calibration were assessed. C-statistic was not reported.

(*continued on next page*)

Table 1 (*continued*)

• Author • Year (Ref) • Abbreviated name	• Population • Age range • N	Country	Risk score (range) or prediction equation	Variables in the prediction model (score)	Area under ROC or C statistics	Comments
• Flemming • 2014 [28]	• Patients with cirrhosis[b,d] • ≥18 years • 34,932	US	Prediction model from Cox's model	Age in years (0.0532), male (0.5114), diabetes (0.2135), nonwhite or Hispanic (0.2058), alcohol/metabolic (0.3509), viral (1.246), severity indicates a CTP score (0.1170)	C-indices= 0.70 (derivation set), 0.69 (validation set)	Patients with cirrhosis in the derivation group were identified from a national liver transplantation waitlist database in the US and those in the external validation group were from the Hepatitis C Antiviral Long-term Treatment against Cirrhosis study. The advantage of this study is its large size of sample. The selection bias was less likely because they were from liver transplant centers nationwide. The study results were more generalizable to a population with cirrhosis, but not a general population. No risk score system was provided. Both discriminant ability and calibration were assessed. Due to heterogeneity of patient population, the discriminant ability was fair.
• Ikeda • 2006 [29]	• Patients with HCV and cirrhosis[b] • 28–80 • 183	Japan	Prediction model from Cox's model	Age ≥55, male, AFP ≥20 ng/ml, platelet count <100,000/mm^3		This study included patients with HCV infection and cirrhosis identified in 1 hospital between 1974 and 1990 as the derivation group. Two external validation groups were used: 1 from the same hospital between 1991 and 2003; and the other from the other hospital between 1975 and 2002. A selection bias is likely to occur because the study subjects were from 2 hospitals. Some selective factors may be associated with patients seeking care in these hospitals. The study results can only be generalized to patients with HCV and cirrhosis who had similar characteristics to the study sample. Only clinical and routine biomarkers were considered, but no HCV biomarkers. No risk score system was provided. Both discriminant ability for risk score and calibration were assessed. C-statistic was not reported.
• Lee • 2014 [30]	• Patients with seropositive anti-HCV antibodies[b] • 30–65 • 1,557	Taiwan	Risk score (0–22)	Age, 5 years increment (1), ALT >45 U/L (4), AAR ≥1 (2), liver cirrhosis (10) or without liver cirrhosis and high HCV RNA level/HCV genotype 1 (6)	Area under ROC = 0.75, 0.83, 0.83 (5-, 10-, 15-year in derivation set)	The risk score was derived from a sample of the REVEAL-HCV cohort and the validation cohort included residents attending a community-based screening program. Detailed HCV genotype and HCV RNA levels were measured and their combined variable was an important determinant of HCC with magnitude of association next to liver cirrhosis. The estimated effects of these determinants may not be generalized to patients in the general population infected with HCV because their effects were estimated in patients with HCV infection and this may not be the same as that estimated in general population. Discriminant ability was assessed, but calibration was not.
• Tamaki • 2014 [31]	• Patients with HCV infection receiving IFN not achieving a sustained response[b] • N/A • 1,046	Japan	FIB-4 Index derived in the previous study for liver fibrosis	FIB-4 index=age [years] × AST [IU/L]/(platelets [10^9/L] × AST [IU/L]$^{1/2}$)		This study included patients with HCV infection receiving double liver biopsies at a hospital to test FIB-4 index's predictive ability and those receiving a single liver biopsy during the same period as validation group. The selection bias is likely to happen when some selective factors were associated with patients' seeking care at this hospital. In addition, the study results may not be generalized to those with HCV with sustained virological response. This study included clinical biomarkers such as AST and platelets, but no HCV biomarkers. Discriminant ability for risk score was assessed, but calibration was not. C-statistic was not reported.
• Zeng • 2016 [32]	• Patients with HCV having sustained virologic response[c] • N/A • 100	China	Risk score	Liver cirrhosis (1), post-SVR albumin ≤36.0 g/L (1)	Area under ROC=0.88	Patients with HCV having sustained virologic response were identified from datasets of 2 hospitals. The selection bias was likely if selective factors were associated with patients' seeking care behaviors. The study results were less generalizable. The size of sample is the smallest among those conducted in patients with HCV infection. Detailed HCV genotype and HCV RNA levels were not considered. Both discriminant ability for risk score and calibration were not assessed.

(*continued on next page*)

Table 1 (*continued*)

• Author • Year (Ref) • Abbreviated name	• Population • Age range • N	Country	Risk score (range) or prediction equation	Variables in the prediction model (score)	Area under ROC or C statistics	Comments
• Aoki • 2016 [33] • VFMAP	• Patients with CLD[c] • N/A • 1,808	Japan	Risk score derived from Cox's model with cutoff points by ROC curve (0–5)	Age ≥55 (1), male (1), VTQ >1.33 m/s (1), FPG ≥110 mg/dL (1), AFP level ≥5 ng/dl (1)	Area under ROC = 0.82 (5-year)	The risk score derivation was based on a sample of participants with chronic liver disease at entry with measurement of virtual touch quantification (VTQ) in a clinical institution. VTQ was a novel predictor for HCC risk in patients with CLD. Because participants from 1 institution were selected, thus there was a potential for selection bias. In addition, the generalizability of the study finding was limited. Discriminant ability was assessed in the same sample, not in the validation set. Calibration was not assessed.

Abbreviations: AAR = AST to ALT ratio; AFP = alpha-fetoprotein; ALT = alanine aminotransferase level; AST = aspartate aminotransferase level; BMI = body mass index; CLD = chronic liver disease; CTP score = child-turcotte-pugh score; CU-HCC = Chinese University-Hepatocellular Carcinoma; DM-HCC = diabetes mellitus-hepatocellular carcinoma; SVR = sustained virologic response; EV = external validation; FPG = fasting plasma glucose; HBV = hepatitis B virus; HBsAg = hepatitis B surface antigen (HBsAg); HCV = hepatitis C virus; IFN = interferon; NA = not available; PIR = pegylated interferon and ribavirin; REVEAL-HBV = risk evaluation of viral load elevation and associated liver disease/cancer-hepatitis B virus; ROC = receiver operating characteristics; ROC curve = receiver operating characteristics curve; VFMAP = virtual touch quantification. All prediction models were conducted by Cox's proportional hazard model except for those by Han et al [17] and Rau et al [15] by logistic regression analysis and all studies adopted cohort study design except for those by Zeng et al [32] and Rau et al [15]: a matched case-control study.

[a] Internal validation was assessed.
[b] External validity was assessed.
[c] No validation was done.
[d] Cross-validation was done.

Table 2
Parameter estimates of regression coefficient, the means or proportions, and risk scores of predictors for HCC from the final multivariate Cox's proportional hazards model.

Risk factors	$\hat{\beta}(\widehat{SE})$	Mean or proportion	*P* value	Risk scores
Sociodemographic factors				
Age	0.05 (0.005)	60.94	<0.001	−2 to 8
Gender (male)	0.41 (0.10)	0.47	<0.001	2
Smoking habit	0.46 (0.12)	0.15	<0.001	2
Diabetes related factor and biomarker				
SGPT (u/l)				
6–45	ref	ref	ref	0
<6 or >45	1.50 (0.09)	0.16	<0.001	6
Variation of HbA1c (%)				
<8.5	ref	ref	ref	0
8.5–17.5	0.03 (0.12)	0.33	0.83	0
>17.5	0.25 (0.11)	0.34	0.02	1
Comorbidity				
Liver cirrhosis	2.23 (0.18)	0.01	<0.001	9
Hepatitis B	1.10 (0.25)	0.01	<0.001	4
Hepatitis C	0.79 (0.25)	0.01	0.002	3
Medication use				
Antidiabetes medications				
No medication	ref	ref	ref	0
Oral only	0.03 (0.23)	0.86	0.88	0
Insulin	0.64 (0.29)	0.03	0.03	2
Insulin + oral agent	0.68 (0.27)	0.06	0.01	3
Antihyperlipidemia medication				
No:				
THR: male <5; female <4.5	ref	ref	ref	0
THR: male 5–9.4; female 4.5–7	−0.29 (0.12)	0.19	0.02	−1
THR: male ≥9.5; female ≥7	0.44 (0.34)	0.01	0.17	2
Yes:				
THR: male <5; female <4.5	−0.48 (0.13)	0.20	<0.001	−2
THR: male 5–9.4; female 4.5–7	−0.95 (0.17)	0.17	<0.001	−4
THR: male ≥9.5; female ≥7	−0.67 (0.58)	0.01	0.25	−3

SGPT = serum glutamic-pyruvic transaminase; ALT = alanine aminotransferase level; THR: total/high-density lipoprotein (HDL) cholesterol ratio.

as a user of specific medication when his/her number of prescription days for each specific drug was greater than 3 months. Under this definition, a patient may have more than 1 type of antidiabetes medication. The antidiabetes medications of individual patients were further classified into 4 categories: no medication, oral antidiabetes drugs, insulin monotherapy, and insulin plus oral antidiabetes drugs. Types of antihypertension medications consisted of ACE inhibitors, ARBs, β-blockers, calcium channel blockers, and diuretics. Cardiovascular medications included antiarrhythmics, anticoagulants, antiplatelet agents, digoxin, and nitrates. Antihyperlipidemia medications considered were statins and fibrates.

Statistical analysis

Proportions were presented for categorical variables, and the means and standard deviations for continuous variables. The effect sizes were calculated to assess the comparability of baseline characteristics between derivation and validation sets. Hazard ratios of

predictor variables were estimated using Cox proportional hazard models to develop a prediction model of HCC in the derivation set.

Steps in developing the predictive model are based on the Framingham Heart Study for determining an HCC risk score system [35]. HCC risk was calculated using the following equation: $1 - P_0^{\exp(\sum \beta_i \times M_i - \sum \beta_i \times \bar{M}_i)}$, where P_0 is the baseline disease-free probability, β_i the regression coefficient for X_i, and M_i is the mean level of X_i. Receiver operating characteristic (ROC) curve analysis was applied to assess the predictive accuracy, and area under the curve (AUC) was used to assess the discriminatory ability of the predictive model. Hosmer–Lemeshow x^2 goodness-of-fit test was performed by comparing the observed and predicted HCC events. For sensitivity analysis, we used a multiple imputation method to impute missing data and examine the sensitivity of our results to missing data. For internal validation, we performed bootstrap resampling 1,000 times to correct the potential for over-fitting or "optimism." The agreement between model-predicted and observed probabilities was determined to assess model calibration. The calculation of intercept by calibration-in-large method was used to assess whether or not the predictions were systematically extremely low or high. A 0 value for the intercept indicates no systematic deviation of estimation of predicted probabilities. Furthermore, the calibration slope was estimated to assess the extremeness of predicted probabilities. A slope value of 1 indicates no overfitting of the model. We completed statistical analysis using SAS version 9.4 (SAS Institute Inc., Cary, NC). Statistical significance was considered at a P value $<.05$.

Results

After 8.33 years of follow-up, 493 (2.33%) and 255 (2.41%) HCC cases were newly diagnosed in the derivation and validation sets, respectively. The baseline characteristics of the derivation and validation sets are shown in Supplementary Table 1. These 2 sets had similar baseline characteristics and risk of developing HCC with all effect sizes less than 0.1.

Supplementary Table 2 shows the cumulative incidence rates of HCC grouped by baseline predictors. Baseline characteristics with a P value $<.20$ in the univariate analyses were then entered as covariates into the multivariate Cox model. The final model included predictors with P values $<.05$. Table 2 shows the risk scores of 10 predictors for HCC incidence according to their regression coefficients in the Cox proportional hazard model. The assigned score for each predicator was fivefold of the regression coefficient for age. The calculated risk scores ranged from −6 (lowest risk) to 40 (highest risk). The 3-, 5-, and 10-year incidence rates ranged from 0.04%, 0.07%, and 0.15%, respectively, at −6 points to 100%, 100%, and 100%, respectively, at 40 points (Table 3).

The ROC curves in the derivation set were 0.81 (95% confidence interval: 0.77, 0.84) at 3 years, 0.80 (0.77, 0.83) at 5 years, and 0.77 (0.75, 0.79) at 10 years; the corresponding values in the validation set were 0.79 (0.73, 0.84) at 3 years, 0.77 (0.73, 0.82) at 5 years, and 0.76 (0.73, 0.79) at 10 years (Fig. 2). In Fig. 2, all values for ROC curves showed good discrimination (>0.7). The HCC predictive model by Hosmer–Lemeshow x^2 tests had high calibration for 3-, 5-, and 10-year risks (all P values $>.05$), indicating a good fit (Fig. 3).

Table 3
The 3-, 5-, 10-year estimated risk for HCC of each possible sum of points.

Point total	Predicted risk of HCC (%)		
	Estimate of 3-year risk	Estimate of 5-year risk	Estimate of 10-year risk
−6	0.04%	0.07%	0.16%
−5	0.05%	0.09%	0.21%
−4	0.06%	0.11%	0.27%
−3	0.08%	0.14%	0.34%
−2	0.11%	0.18%	0.45%
−1	0.14%	0.24%	0.58%
0	0.18%	0.31%	0.75%
1	0.24%	0.40%	0.97%
2	0.30%	0.52%	1.25%
3	0.39%	0.67%	1.61%
4	0.51%	0.87%	2.08%
5	0.66%	1.12%	2.69%
6	0.85%	1.45%	3.47%
7	1.10%	1.88%	4.47%
8	1.43%	2.43%	5.75%
9	1.85%	3.13%	7.39%
10	2.38%	4.03%	9.46%
11	3.08%	5.19%	12.07%
12	3.97%	6.67%	15.35%
13	5.11%	8.56%	19.41%
14	6.56%	10.94%	24.38%
15	8.41%	13.93%	30.37%
16	10.76%	17.66%	37.42%
17	13.71%	22.24%	45.50%
18	17.38%	27.80%	54.44%
19	21.90%	34.42%	63.87%
20	27.40%	42.10%	73.25%
21	33.94%	50.72%	81.87%
22	41.55%	60.01%	89.04%
23	50.11%	69.48%	94.29%
24	59.37%	78.50%	97.55%
25	68.85%	86.34%	99.18%
26	77.92%	92.41%	99.80%
27	85.86%	96.45%	99.97%
28	92.06%	98.67%	100.00%
29	96.24%	99.63%	100.00%
30	98.57%	99.93%	100.00%
31	99.59%	99.99%	100.00%
32	99.92%	100.00%	100.00%
33	99.99%	100.00%	100.00%
34	100.00%	100.00%	100.00%
35	100.00%	100.00%	100.00%
36	100.00%	100.00%	100.00%
37	100.00%	100.00%	100.00%
38	100.00%	100.00%	100.00%
39	100.00%	100.00%	100.00%
40	100.00%	100.00%	100.00%

Discussion

With 32,585 participants enrolled in clinical centers nationwide, the Taiwan Diabetes Study cohort presented a valuable opportunity to construct a prediction model for estimating the incidence of HCC. Our study has developed a highly predictive 10-item HCC risk score system for patients with type 2 diabetes aged 30 years and above with good discrimination and predictive capability for 3-, 5-, and 10-year HCC risks.

A previously conducted systematic review indicated that when compared with patients without diabetes those with diabetes have a 2-fold increased risk of HCC [36]. Given this increased propensity to develop HCC, 2 HCC risk prediction models have been reported and currently exist for patients with type 2 diabetes. The first is the Korea DM-HCC risk score that used data from 3,544 patients with diabetes but without chronic viral hepatitis or alcoholic cirrhosis to establish 10-year cumulative incidences of HCC [16]. However, because it included as factors age, plasma triglycerides, and γ-glutamyl transferase levels, its generalization is limited since the latter is not usually a part of clinical follow-up of patients with diabetes. The second is a Web-based liver cancer prediction model for Chinese patients with type 2 diabetes developed to predict liver cancer occurrence within 6 years of diagnosis by considering the age, gender, as well as chronic liver diseases such as alcoholic cirrhosis, other cirrhosis, alcoholic hepatitis, viral hepatitis, other chronic hepatitis, alcoholic fatty liver disease, and other fatty

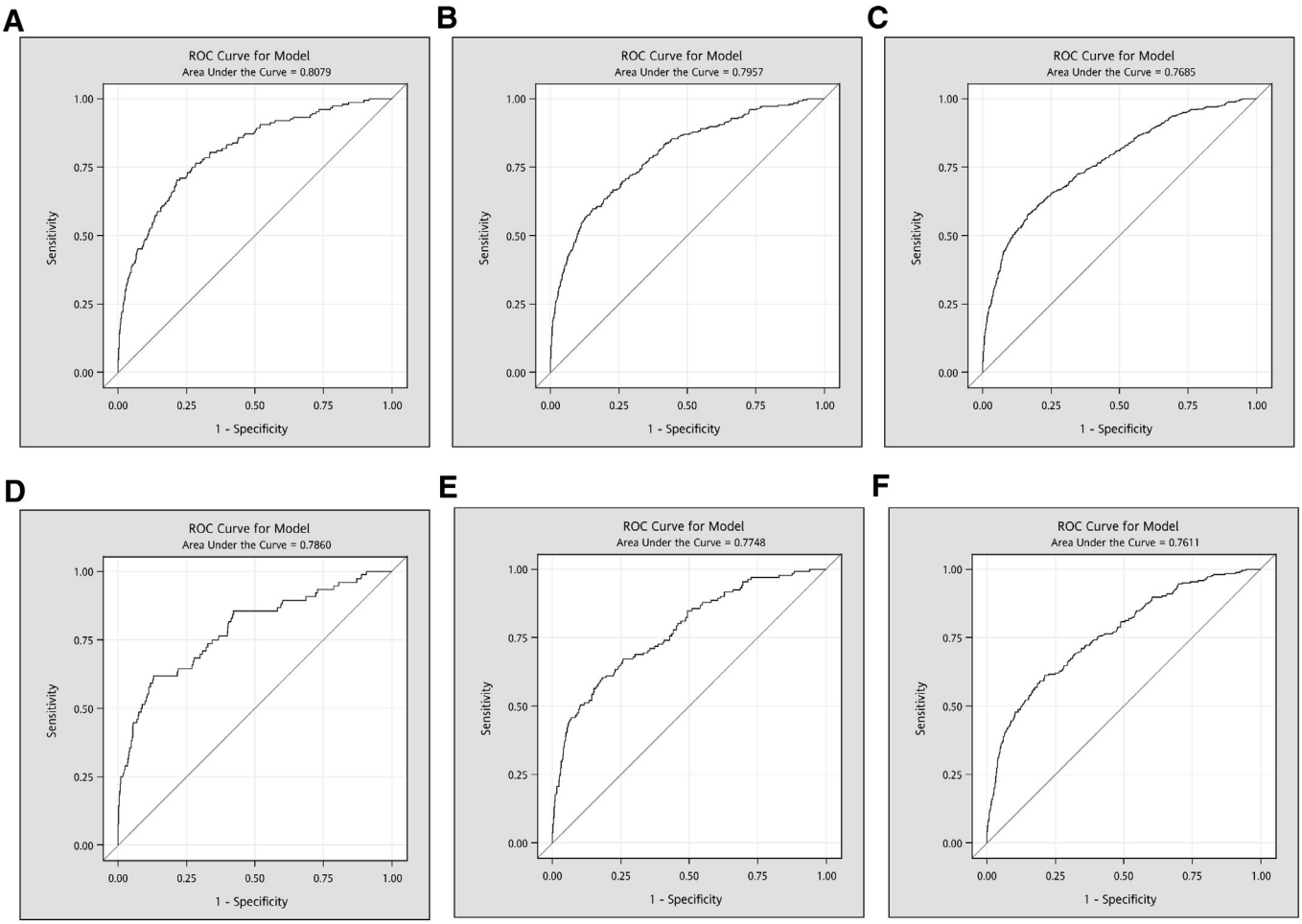

Fig. 2. Receiver operating characteristic curve (ROC) for (A) 3-year, (B) 5-year, (C) 10-year HCC risk in derivation set, and (D) 3-year, (E) 5-year, (F) 10-year HCC risk in validation set.

liver diseases [15]. While informative, this study did not consider biomarkers, such as fasting plasma glucose, HbA1c, total cholesterol, triglycerides, HDL, or LDL, that are routinely measured for diabetes care (Table 4). Compared to these 2 existing HCC prediction models in patients with type 2 diabetes, our study includes a large cohort of 32,585 participants enrolled from clinical centers in the entire country, enhancing its representativeness and generalization and allowing the estimates in our prediction model to be more precise and robust. Additionally, our model has the advantages of having adopted a higher evidence level in its design (cohort study), of including diabetes-related variables commonly acquired clinically, and of comprehensively assessing calibration and performing sensitivity analysis. Furthermore, the previously described HCC prediction models do not include as potential predictors diabetes-related factors such as HbA1c-CV, FPG-CV, and insulin or statin use [15,16], considered key variables for diabetes care. These are generally accepted risk factors that we have included in our prediction model, and that importantly can be precisely measured, ensuring our scoring system can be applied to clinical practice.

The availability of a simple and useful prediction model for HCC is important for individuals, public health, clinical practice, and medical research. One advantage of prediction models is the ability to quantify the effect of measurable and modifiable risk factors and to generate risk estimates ranging from 5 years to 10 years. These point systems enable health professionals and practitioners to use complex statistical models and simplify the estimation of risks based on complex statistical models. In addition, clinicians can be guided by these point systems in the decision-making regarding treatments and in motivating patients to modify their behaviors. Another advantage of this point system is its availability, which enables patients to easily estimate their own disease risk and monitor their risk over time.

One review indicated chronic episodes of hyperglycemia could serve as a direct or indirect mediator of the increase in tumor cell proliferation, especially in the liver, pancreas, mammary glands, and endometrium [37]. Accumulating evidence suggests hyperinsulinemia increases IGF-1 production in early stage type 2 diabetes [38]. IGF-1 stimulates hepatic cell proliferation and inhibits apoptosis in the liver [39]. In vitro studies, animal models, and epidemiologic studies have shown that insulin and IGF-1 play critical roles in liver carcinogenesis [40]. While statins used primarily as cholesterol-lowering medications for dyslipidemia may have cytostatic effects on tumor cells and improve the survival of patients with a diagnosis of HCC [41]. Given the potential importance of these diabetes-related markers on HCC risk, our study integrated these variables in developing a predictive model for HCC in patients with type 2 diabetes. Note that these variables were not considered in prior studies but that as we show, their composite scores displayed good predictive ability.

We also reported a sensitivity analysis based on multiple imputation for handling missing data. A total of 51,119 type 2 diabetic patients were included in the sensitivity analysis. The AUCs of 3-, 5-, and 10-year HCC risks were 0.81, 0.80, and 0.77, respectively, in our final models and 0.78, 0.78, and 0.76, respectively, in the sensitivity analysis. The AUCs achieved in the sensitivity analysis indicate our findings were robust. Internal val-

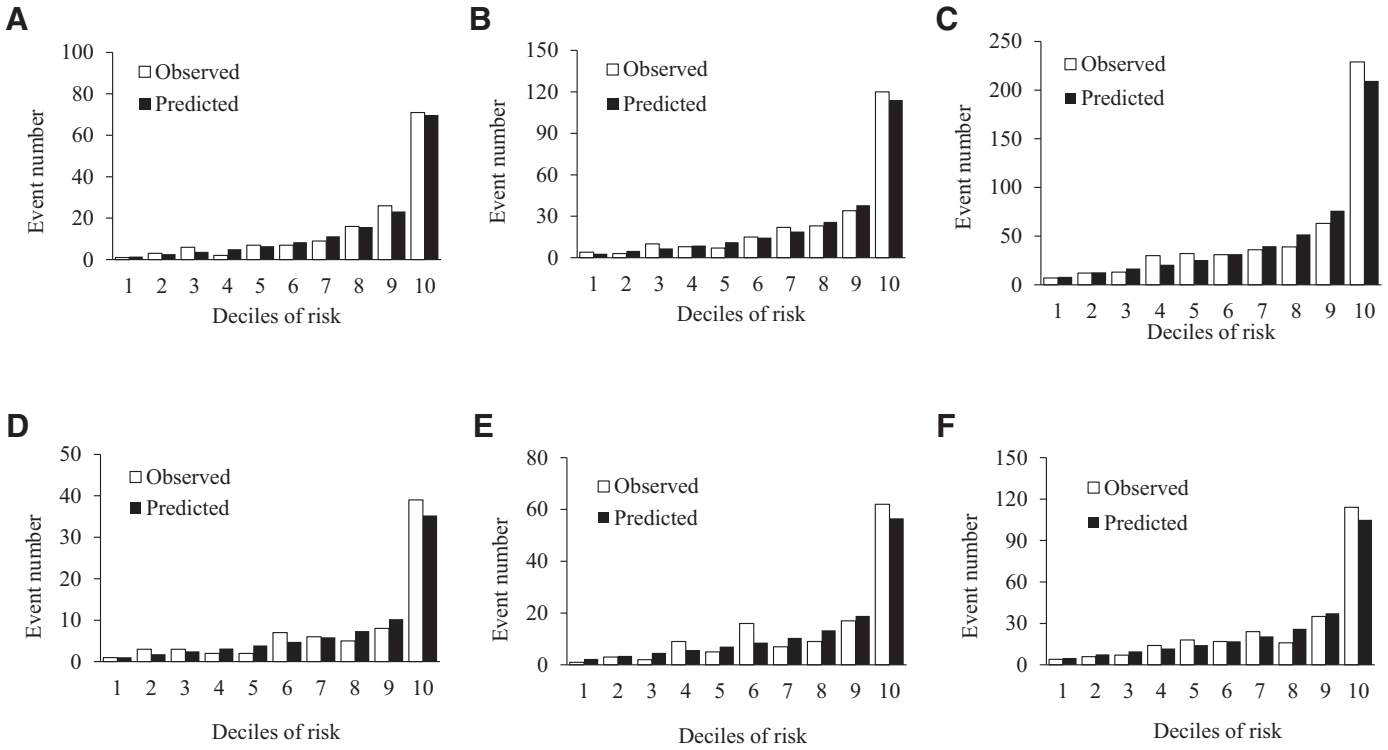

Fig. 3. Predicted versus observed HCC numbers according to deciles of (A) 3-year, (B) 5-year, and (C) 10-year risk in derivation set, and (D) 3-year, (E) 5-year, and (F) 10-year risk in validation set.

Table 4
HCC risk scores for patient with type 2 diabetes.

• Author • Year (ref no) • Abbreviated name	• Age range • Sample size	Country	Study design	Type of sample	Prediction model type	Risk score system	Unique variables in the prediction model	Area under ROC or C statistics	Calibration methods	Sensitivity analysis
• Li 2018 (present study) • TDS study	• 30–84 years • 31,723	Taiwan	Cohort study	Nationwide sample	Cox's model	Yes	Smoking, SGPT, HbA1c, insulin, insulin plus oral agent	ROC = 0.81, 0.80, 0.77 (3-, 5-, 10-year in derivation set)	The Hosmer–Lemeshow x^2 goodness-of-fit test and calibration-in-large method	Results are robust when missing data were imputed by multiple imputation approach
• Rau 2016 [15]	• 20–90 years • 2,060	Taiwan	Matched case-control study	Nationwide sample	Artificial neural network and logistic model	No	No	ROC = 0.873 (6-year)	No	No
• Si 2016 [16] • DM-HCC	• ≥18 years • 3,544	Korea	Cohort study	Single-center sample	Cox's model	Yes	γ-glutamyl transferase level	ROC = 0.86 (10-year)	No	No

ROC = receiver operating characteristics; ROC curve = receiver operating characteristics curve; SGPT = serum glutamic-pyruvic transaminase; ALT = alanine aminotransferase level.

idation of the performance of the present model was assessed based on 1,000 samples from bootstrap resampling. The optimism corrected calibration intercept was 0.003 with a mean absolute error of 0.00002, and the corresponding slope was 0.92 with a mean absolute error of 0.0009. These statistics indicated fairly good calibration for the present model, such that shrinkage of regression coefficients in the prediction model was not needed. The mean slope was slightly smaller than 1.0, indicating that the estimated risk of HCC in our prediction model was slightly overestimated.

Strengths and limitations

The strengths of our study included its conduct as a large population-based study with a long-term follow-up period, and the inclusion of novel predictors of HbA1c-CV. Antidiabetic, antihypertensive, cardiovascular, and antihyperlipidemia medications were also considered. All participants were randomly allocated into derivation and validation sets for our model's internal validations.

This study has 3 limitations. First, we only considered baseline status. Time-varying effects were not considered because of lack of data. Second, missing data may be a potential bias,

although we used a multiple imputation approach and internal validation method to assess their effects. Third, the NDCMP database lacked information on HBeAg, hepatitis B virus DNA or hepatitis C virus RNA levels, and genetic factors. Thus, we could not consider these factors. Although these biomarkers could remarkably improve the discriminatory ability of the prediction model, they were not cost-effective in patients with type 2 diabetes because of their low prevalence in this population.

Conclusion

A simple-to-use prediction tool has been developed to effectively predict 3-, 5-, and 10-year HCC risks among Taiwanese patients with type 2 diabetes. This tool may help clinicians or policy makers to plan preventive and treatment strategies for high-risk individuals to reduce the risk of developing HCC. Our study also supported previous studies regarding predictors in patients with type 2 diabetes and suggested additional important factors, including both HbA1c-CV and hyperlipidemia medications use.

Conflict of interest

The authors declare there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Acknowledgments

This study was supported primarily by the Bureau of National Health Insurance (DOH94-NH-1007), the Ministry of Science and Technology of Taiwan (NSC 101-2314-B-039 -017-MY3 & NSC 102-2314-B-039-005-MY2 & MOST 104-2314-B-039-016 & MOST 105-2314-B-039-021-MY3 & MOST 105-2314-B-039-025-MY3) and the Ministry of Health and Welfare, Taiwan (MOHW107-TDU-B-212-123004) and China Medical University Hospital (DMR-106-094).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1053/j.seminoncol.2018.07.006.

References

[1] IARC. GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012 Available from: http://globocan.iarc.fr/Default.aspx.
[2] MOHW. 2016 Statistics of Causes of Death 2016 [Available from: http://www.mohw.gov.tw/lp-3327-2.html.
[3] WHO. Global report on diabetes 2016 [07/17/207]. Available from: http://www.who.int/mediacentre/factsheets/fs312/en/.
[4] Chiang CH, Lee LT, Hung SH, et al. Opposite association between diabetes, dyslipidemia, and hepatocellular carcinoma mortality in the middle-aged and elderly. Hepatology 2014;59(6):2207–15.
[5] El-Serag HB, Tran T, Everhart JE. Diabetes increases the risk of chronic liver disease and hepatocellular carcinoma. Gastroenterology 2004;126(2):460–8.
[6] Fujino Y, Mizoue T, Tokui N, Yoshimura T. Prospective study of diabetes mellitus and liver cancer in Japan. Diabetes Metab Res Rev 2001;17(5):374–9.
[7] Printz C. Diabetes associated with increased risk of liver cancer. Cancer 2014;120(9):1288.
[8] Greten FR, Eckmann L, Greten TF, et al. IKKbeta links inflammation and tumorigenesis in a mouse model of colitis-associated cancer. Cell 2004;118(3):285–96.
[9] Hu MC, Lee DF, Xia W, et al. IkappaB kinase promotes tumorigenesis through inhibition of forkhead FOXO3a. Cell 2004;117(2):225–37.
[10] Lee DF, Kuo HP, Chen CT, et al. IKK beta suppression of TSC1 links inflammation and tumor angiogenesis via the mTOR pathway. Cell 2007;130(3):440–55.
[11] Lee DF, Kuo HP, Chen CT, et al. IKKbeta suppression of TSC1 function links the mTOR pathway with insulin resistance. Int J Mol Med 2008;22(5):633–8.
[12] Hung YC, Lin CL, Liu CJ, et al. Development of risk scoring system for stratifying population for hepatocellular carcinoma screening. Hepatology 2015;61(6):1934–44.
[13] Michikawa T, Inoue M, Sawada N, et al. Development of a prediction model for 10-year risk of hepatocellular carcinoma in middle-aged Japanese: the Japan Public Health Center-based Prospective Study Cohort II. Prev Med 2012;55(2):137–43.
[14] Wen CP, Lin J, Yang YC, et al. Hepatocellular carcinoma risk prediction model for the general population: the predictive power of transaminases. J Natl Cancer Inst 2012;104(20):1599–611.
[15] Rau HH, Hsu CY, Lin YA, et al. Development of a web-based liver cancer prediction model for type II diabetes patients by using an artificial neural network. Comput Methods Programs Biomed 2016;125:58–65.
[16] Si WK, Chung JW, Cho J, et al. Predictors of increased risk of hepatocellular carcinoma in patients with type 2 diabetes. PLoS One 2016;11(6):e0158066.
[17] Han KH, Ahn SH. How to predict HCC development in patients with chronic B viral liver disease? Intervirology 2005;48(1):23–8.
[18] Lee MH, Yang HI, Liu J, et al. Prediction models of long-term cirrhosis and hepatocellular carcinoma risk in chronic hepatitis B patients: risk scores integrating host and virus profiles. Hepatology 2013;58(2):546–54.
[19] Lin YJ, Lee MH, Yang HI, et al. Predictability of liver-related seromarkers for the risk of hepatocellular carcinoma in chronic hepatitis B patients. PLoS One 2013;8(4):e61448.
[20] Sinn DH, Lee JH, Kim K, et al. A novel model for predicting hepatocellular carcinoma development in patients with chronic hepatitis B and normal alanine aminotransferase levels. Gut Liver 2017;11(4):528–34.
[21] Wong GL, Chan HL, Wong CK, et al. Liver stiffness-based optimization of hepatocellular carcinoma risk score in patients with chronic hepatitis B. J Hepatol 2014;60(2):339–45.
[22] Wong VW, Chan SL, Mo F, et al. Clinical scoring system to predict hepatocellular carcinoma in chronic hepatitis B carriers. J Clin Oncol 2010;28(10):1660–5.
[23] Yang HI, Yuen MF, Chan HL, et al. Risk estimation for hepatocellular carcinoma in chronic hepatitis B (REACH-B): development and validation of a predictive score. Lancet Oncol 2011;12(6):568–74.
[24] Yuen MF, Tanaka Y, Fong DY, et al. Independent risk factors and predictive score for the development of hepatocellular carcinoma in chronic hepatitis B. J Hepatol 2009;50(1):80–8.
[25] Chang KC, Hung CH, Lu SN, et al. A novel predictive score for hepatocellular carcinoma development in patients with chronic hepatitis C after sustained response to pegylated interferon and ribavirin combination therapy. J Antimicrob Chemother 2012;67(11):2766–72.
[26] Chang KC, Wu YY, Hung CH, Lu SN, Lee CM, Chiu KW, et al. Clinical-guide risk prediction of hepatocellular carcinoma development in chronic hepatitis C patients after interferon-based therapy. Br J Cancer 2013;109(9):2481–8.
[27] El-Serag HB, Kanwal F, Davila JA, Kramer J, Richardson P. A new laboratory-based algorithm to predict development of hepatocellular carcinoma in patients with hepatitis C and cirrhosis. Gastroenterology 2014;146(5) 1249-55.e1.
[28] Flemming JA, Yang JD, Vittinghoff E, Kim WR, Terrault NA. Risk prediction of hepatocellular carcinoma in patients with cirrhosis: the ADRESS-HCC risk model. Cancer 2014;120(22):3485–93.
[29] Ikeda K, Arase Y, Saitoh S, Kobayashi M, Someya T, Hosaka T, et al. Prediction model of hepatocarcinogenesis for patients with hepatitis C virus-related cirrhosis. Validation with internal and external cohorts. J Hepatol 2006;44(6):1089–97.
[30] Lee MH, Lu SN, Yuan Y, Yang HI, Jen CL, You SL, et al. Development and validation of a clinical scoring system for predicting risk of HCC in asymptomatic individuals seropositive for anti-HCV antibodies. PLoS One 2014;9(5):e94760.
[31] Tamaki N, Kurosaki M, Matsuda S, Muraoka M, Yasui Y, Suzuki S, et al. Non-invasive prediction of hepatocellular carcinoma development using serum fibrosis marker in chronic hepatitis C patients. J Gastroenterol 2014;49(11):1495–503.
[32] Zeng QL, Li B, Zhang XX, Chen Y, Fu YL, Lv J, et al. Clinical model for predicting hepatocellular carcinomas in patients with post-sustained virologic responses of chronic hepatitis C: a case control study. Gut Liver 2016;10(6):955–61.
[33] Aoki T, Iijima H, Tada T, Kumada T, Nishimura T, Nakano C, et al. Prediction of development of hepatocellular carcinoma using a new scoring system involving virtual touch quantification in patients with chronic liver diseases. J Gastroenterol 2017;52(1):104–12.
[34] Kilpatrick ES, Rigby AS, Atkin SL. A1C variability and the risk of microvascular complications in type 1 diabetes: data from the Diabetes Control and Complications Trial. Diabetes Care 2008;31(11):2198–202.
[35] Sullivan LM, Massaro JM, D'Agostino RB. Presentation of multivariate data for clinical use: the Framingham Study risk score functions. Stat Med 2004;23(10):1631–60.
[36] Wang C, Wang X, Gong G, Ben Q, Qiu W, Chen Y, et al. Increased risk of hepatocellular carcinoma in patients with diabetes mellitus: a systematic review and meta-analysis of cohort studies. Int J Cancer 2012;130(7):1639–48.
[37] Ryu TY, Park J, Scherer PE. Hyperglycemia as a risk factor for cancer progression. Diabetes Metab J 2014;38(5):330–6.
[38] Gallagher EJ, LeRoith D. The proliferating role of insulin and insulin-like growth factors in cancer. Trends Endocrinol Metab 2010;21(10):610–18.
[39] Wiencke JK. Impact of race/ethnicity on molecular pathways in human cancer. Nat Rev Cancer 2004;4(1):79–84.
[40] Weng CJ, Hsieh YH, Tsai CM, Chu YH, Ueng KC, Liu YF, et al. Relationship of insulin-like growth factors system gene polymorphisms with the susceptibility and pathological development of hepatocellular carcinoma. Ann Surg Oncol 2010;17(7):1808–15.
[41] Gronich N, Rennert G. Beyond aspirin-cancer prevention with statins, metformin and bisphosphonates. Nat Rev Clin Oncol 2013;10(11):625–42.