Risk factors and prediction score for chronic pancreatitis: A nationwide population-based cohort study

Yen-Chih Lin, Chew-Teng Kor, Wei-Wen Su, Yu-Chun Hsu

AIM
To explore the risk factors of developing chronic pancreatitis (CP) in patients with acute pancreatitis (AP) and develop a prediction score for CP.

METHODS
Using the National Health Insurance Research Database in Taiwan, we obtained large, population-based data of 5971 eligible patients diagnosed with AP from 2000 to 2013. After excluding patients with obstructive pancreatitis and biliary pancreatitis and those with a follow-up period of less than 1 year, we conducted a multivariate analysis using the data of 3739 patients to identify the risk factors of CP and subsequently develop a scoring system that could predict the development of CP in patients with AP. In addition, we validated the scoring system using a validation cohort.

RESULTS
Among the study subjects, 142 patients (12.98%) develop-
Acute pancreatitis (AP) and chronic pancreatitis (CP) are common diseases with a worldwide prevalence. These diseases have become an important public health issue in several countries because of the high mortality rates and a considerable burden being laid on the healthcare system. AP is an inflammatory condition of the pancreas that has been considered as a self-limiting disease, with an incidence ranging from 5 to 10 per 100 000 to 70-80 per 100 000 in western countries, which appears to have increased in recent years[1]. In contrast, CP involves a persistent destructive, inflammatory process that eventually leads to an irreversible damage to the endocrine and exocrine functions of the pancreas, and the subsequent development of diabetes mellitus and frequent hospitalizations have become one of the burdens of public health. CP has a poor prognosis, with the mortality rate being approximately two-fold higher than that in the general population. Furthermore, a worldwide epidemiological survey conducted in 1993 revealed that the standardized incidence rate of pancreatic cancer is as high as 26-fold in patients with CP, suggesting that the risk of pancreatic cancer is significantly higher in subjects with CP[2].

An emerging consensus is that AP and CP are a continuum of diseases, and the intermediate stage between them is recurrent AP (RAP)[3]. Several studies have discussed about the natural course of pancreatitis as well as the risk factors and protective factors that contribute to the transition from AP to RAP and CP, although the majority of them have been conducted in western countries[3,4]. The major risk factors for CP include smoking, in addition to alcohol consumption. Moreover, it has been reported that cigarette smoking accelerates the progression of alcoholic CP[5,6]. Furthermore, a recent study has revealed that alcohol consumption of > 13.5 g/d and cigarette use of > 5.5 cigarettes/d are associated with the development of CP[7]. Because only a small proportion of patients with AP progress to CP, and CP has been proven to be an important risk factor of pancreatic duct adenocarcinoma (PDAC), it is critical to predict the development of CP in a patient with AP. However, till date, no prediction scores for CP have been addressed in the English literature, although there were some prognosis scores for CP and AP[8,9]. Therefore, in this population-based, large-scale cohort study, we developed and validated a scoring system for predicting CP using data from the National Health Insurance Research Database (NHIRD) in Taiwan.

**MATERIALS AND METHODS**

**Data source**

We obtained data from the Taiwan NHIRD. The NHIRD is one of the most comprehensive databases in the world and includes all claims data from the National Health Insurance program, such as demographic data, number of ambulatory cases, records of clinic visits, hospital admissions, dental services, prescriptions, and disease status. The National Health Insurance program, which

---

**INTRODUCTION**

Acute pancreatitis (AP) and chronic pancreatitis (CP) are...
was initiated by the government of Taiwan in March 1995, covers > 99% of the total population or approximately 23 million people. Diagnostic codes used in the NHIRD for identifying diseases are based on the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM), which has been proven to be highly accurate and valid. This study was exempted from full review and was approved by the Institutional Review Board of the Changhua Christian Hospital (approval number: 171112). Study population
A total of 5971 patients with one or more episodes of AP (ICD-9-CM code 577.0) recorded in the inpatient claims data from 2000 to 2013 were identified from the database. A 4-year look-back period was applied from 1996 to 1999 to ensure that all cases in our cohort were newly diagnosed and to reduce false incident cases. Patients with a previous diagnosis of AP during the look-back period were excluded. Patients who had CP before the index date, those aged < 18 or > 100 years, those with a follow-up duration of < 1 year, and those with biliary pancreatitis or obstructive pancreatitis were also excluded because these patients rarely progress to CP. Accordingly, 3739 patients with nonobstructive, nonbiliary AP were identified for subsequent analysis. Next, we developed a model to predict the progress to CP in randomly selected two-thirds of this cohort (derivation cohort) and validated the model in the remaining one-third of this cohort (validation cohort).

Outcome measures and relevant variables
Outcomes and comorbidities were identified based on ICD-9-CM codes. CP was defined using ICD-9-CM codes (ICD-9-CM code 577.1).

To avoid over-estimation of CP by ICD-9-CM coding alone, we excluded all patients without abdominal computed tomography (CT) or abdominal magnetic resonance imaging (MRI) performed within 3 mo before the diagnosis of CP.

Patients were followed up from the index date (i.e., the date of first AP diagnosis) to the date when they withdrew from the insurance program or to the end of 2013. Major comorbid diseases diagnosed before the index date were defined as baseline comorbidities based on claims data. These comorbidities included obesity, hypertension, hyperlipidemia, diabetes mellitus, alcoholism [alcohol use-related codes: ICD-9-CM codes 291, 303, 305.0, 357.5, 571.0, 571.1, 571.2, and 790.3 (V11.3)], smoking habit (smoking-related codes: ICD-9-CM 305.1, V15.82, 491, 492, 493, and 496), and chronic kidney disease. If the patients with AP enrolled in our study have drinking-related coding or smoking-related coding during their follow-up period after the first AP episode, we considered them have drinking or smoking habit. To evaluate the effects of socioeconomic factors on disease development, monthly income, and place of residence of patients were recorded. To quantify baseline comorbidities, Charlson’s comorbidity index (CCI) score was used. The history of long-term use of medications that have been reported as possible risk factors for AP, including statins, angiotensin-converting enzyme inhibitors, prednisolone, hydrochlorothiazide, sex hormones, and metformin, was also evaluated.

Statistical analysis
Demographic and clinical characteristics of the study patients are summarized as proportions and mean ± standard derivation (SD) values. The chi-square test and the t test were used to compare the distributions of discrete and continuous variables, respectively. The risk of CP in patients with nonobstructive, nonbiliary AP was estimated using the Cox proportional hazards model. Variables in the Cox model included the presence of RAP or the number of episodes of AP, smoking, alcohol consumption, age, gender, all comorbidities, CCI scores, and long-term use of medications. The significant coefficients from the Cox model with backward selection procedure were used to construct an integer-based risk score for stratifying the risk of progress to CP. The referent for each variable was assigned a value of 0, and the coefficients for the other variables were calculated by dividing by the smallest coefficient in the model and then rounding to the nearest integer. Individual scores were assigned by summing the individual risk factor scores, and the cumulative incidence rate of each risk score was calculated. For easy application in clinical practice, the total risk scores were classified into low-risk category, moderate-risk category, and high-risk category based on similar magnitudes of hazard.

Within the derivation cohort, the discrimination was assessed using the time-dependent area under the receiver operating characteristic (ROC) curve. The internal validation of this risk score was conducted by 1000 bootstrap simulations. The bootstrap simulations in the derivation cohort were carried out by sampling with replacement for 1000 iterations. Each bootstrap sample was of the same size as the derivation cohort, the computed risk score, and the generated area under the ROC. Furthermore, we validated the risk score externally using the remaining one-third of the random sample. The risk score model was assessed, and the discrimination was assessed by a time-dependent ROC curve analysis. All statistical analyses were performed using SAS 9.4 software (SAS Institute Inc., Cary, NC, United States). Two-tailed P values of < 0.05 were considered to be statistically significant.

RESULTS
Characteristics of the study population
The flowchart depicted in Figure 1 describes the patient selection process, while Table 1 shows the characteristics of the study patients. A total of 3739 patients were identified as newly diagnosed nonobstructive, nonbiliary...
AP cases. Among these patients, 174 (4.65%) developed CP during the mean follow-up period of 6.13 ± 3.53 years. The mean age at the first onset of nonbiliary, nonobstructive AP was 53.04 ± 17.00 years. In addition, according to the inpatient claims data, 1094 patients had more than one episode of nonbiliary, nonobstructive AP, whereas 2645 patients had only one episode of disease attack. Regarding the behavior-related variables, 21.56% and 33.73% of patients had alcohol-use-related codes and smoking-related codes, respectively. After a random sampling, the derivation cohort consisted of 2493 patients, and the validation cohort consisted of 1246 patients, and the baseline characteristics and demographic variables were comparable between both cohorts (Table 1). The prevalence of CP was similar between the derivation and validation cohorts, at 113 (4.53%) and 61 (4.90%), respectively (Table 1).

**Risk score construction for predicting CP**
Table 2 shows the results of the multivariate Cox proportional hazard analysis. From the Cox model, four variables, including smoking habit, age < 55 years, alcohol consumption, and RAP/number of episodes of AP, were associated with risks of CP (all \( P < 0.05 \)). Due to the insignificance associated with the risk of CP comorbidities and history of medication use were excluded from the final model after performing a backward elimination procedure. The risk score based on these factors was constructed as shown in Table 3. Two scoring systems, namely risk score 1 and risk score 2, were separately developed based on the presence or absence of RAP (point 5 if RAP is present) or the number of AP episodes (point 4 for two episodes, point 5 for three episodes, and point 7 for more than three episodes), as well as alcohol-use-related codes (point 3), age < 55 years (point 2), and smoking-related codes (point 1).

**Personalized risk prediction**
The total score for each patient was calculated by summing each risk factor point. The risk scores 1 and 2 ranged between 0-11 and 0-13, respectively (Table 3). Figure 2 shows the time-dependent ROC curve assessing the discrimination for predicting CP using the risk score at different endpoints of time. The risk score 1 had excellent discrimination for predicting the 3-year, 5-year, and overall time period CP incidence, with the area under the ROC curve of 0.84, 0.84, and 0.79, respectively. The 95% confidence intervals (CIs) yielded by the 1000 bootstrap simulation validation were 0.80-0.88, 0.81-0.87, and 0.75-0.83, respectively (Figure 2A-C, solid line). The risk score 2 also had excellent discrimination for predicting the 3-year, 5-year, and
Risks stratification

As the risk score increased, the incidence rate as well as the hazard of CP increased (Supplementary Table 1). On the basis of similar magnitudes of hazard, the risk score 1 of 0-5 was classified as a low-risk category, the risk score 1 of 6-7 was classified as a moderate-risk category, and the risk score 1 of > 7 was classified as a high-risk category. Figure 3 presents the incidence rates for CP over the risk category. As demonstrated in Figure 3A, the incidence rates of CP using risk score 1 were 1.27, 7.89, and 31.37 per 1000 person-years for the low-, moderate-, and high-risk categories, respectively. The hazards of CP were 6.14 (3.05, 12.35) and 23.93 (13.4, 42.73) for the moderate- and high-risk categories, respectively (Supplementary Table 1 and Figure 3).

Similarly, the values of 0-5, 6-7, and 8-13 in risk score 2 (which includes the number of AP episodes) were classified as low-, moderate-, and high-risk categories, respectively. The incidence rates of CP were 1.30, 9.26, and 32.22 per 1000 person-years for the low-, moderate-, and high-risk categories, respectively (Figure 3B). The hazards of CP were 7.08 (3.54, 14.14) and 24.15 (13.76, 42.38) for the moderate- and high-risk categories, respectively (Figure 4).

Validation cohort

The validation cohort was used to test the risk scores. The risk scores were calculated for each patient in the validation cohort, and they were successfully classified as low-, moderate-, and high-risk categories according to the score stratifications in the derivation cohort. As the risk category increased, the incidence rate as well as the hazard of CP increased (Figures 3 and 4). The areas under the ROC curve at 3-, 5-, and overall year were 0.81, 0.83, and 0.82, respectively. The 95%CIs yielded by the 1000 bootstrap resamplings were 0.76-0.86, 0.79-0.87, and 0.78-0.87 (Figure 3D-F, solid line). In risk score 2, the areas under the ROC curve at 3-, 5-, and overall year were 0.82, 0.84, and 0.83, respectively. The 95%CIs were 0.77-0.87, 0.80-0.88, and 0.78-0.87, respectively (Figure 3D-F, dashed line). This result demonstrated that the risk score had similar performance between the derivation and validation cohorts.

Table 1  Demographic characteristics of 3739 patients (%)

| Total   | Derivation cohort | Validation cohort | P value |
|---------|-------------------|-------------------|---------|
| Sample size | 3739              | 2493              | 1246    |         |
| Age, yr | 53.04 ± 17        | 52.98 ± 17.08     | 53.18 ± 16.84 | 0.735 |
| Gender, male | 2400 (64.19)   | 1590 (63.78)      | 810 (65.01)   | 0.460 |
| Monthly income, NTD | 15423.5 ± 13018.5 | 15405.6 ± 13071.9 | 15459.1 ± 12916.2 | 0.906 |
| Geographic location |                   |                   |         |
| Northern Taiwan | 1619 (43.3)      | 1085 (43.52)      | 534 (42.86)    | 0.725 |
| Central Taiwan | 790 (20.86)       | 518 (20.76)       | 262 (21.03)     | 0.893 |
| Southern Taiwan | 1183 (31.64)     | 790 (31.69)       | 393 (31.54)     | 0.957 |
| Eastern Taiwan and Islands | 157 (4.2)      | 100 (4.01)        | 57 (4.57)       | 0.470 |
| Charlson’s comorbidity index score | 2.55 ± 2.21     | 2.56 ± 2.22       | 2.52 ± 2.19     | 0.574 |
| Comorbidities |                   |                   |         |
| Obesity | 17 (0.45)         | 11 (0.44)         | 6 (0.48)       | 0.863 |
| Hypertension | 1456 (38.94)   | 984 (39.47)       | 472 (37.88)    | 0.348 |
| Hyperlipidemia | 1048 (28.03)   | 698 (28)          | 350 (28.09)    | 0.953 |
| Diabetes mellitus | 994 (26.58)    | 658 (26.39)       | 336 (26.97)    | 0.709 |
| Chronic kidney disease | 498 (13.32)   | 332 (13.32)       | 166 (13.32)    | 0.996 |
| Alcohol use-related codes | 806 (21.56)    | 533 (21.38)       | 273 (21.91)     | 0.710 |
| Smoking related codes | 1261 (33.73)   | 849 (34.06)       | 412 (33.07)     | 0.546 |
| Long-term medication use |                   |                   |         |
| Statin | 569 (15.22)       | 385 (15.44)       | 184 (14.77)     | 0.388 |
| Angiotensin-converting enzyme inhibitor | 472 (12.62)   | 327 (13.12)       | 145 (11.64)     | 0.199 |
| Prednisolone | 74 (1.98)         | 56 (2.25)         | 18 (1.44)       | 0.097 |
| Hydrochlorothiazide | 41 (1.1)         | 29 (1.16)         | 12 (0.96)       | 0.580 |
| Sex hormone | 180 (4.81)        | 129 (5.17)        | 51 (4.09)       | 0.145 |
| Metformin | 434 (11.61)       | 290 (11.63)       | 144 (11.56)     | 0.946 |
| Number of episode RAP |                   |                   |         |
| 1 | 2645 (70.74) | 1769 (70.96) | 876 (70.3) | 0.707 |
| 2 | 599 (16.02) | 403 (16.17) | 196 (15.73) | 0.768 |
| 3 | 234 (6.26) | 153 (6.14) | 81 (6.5) | 0.718 |
| ≥ 4 | 261 (6.98) | 168 (6.74) | 93 (7.46) | 0.452 |
| Outcomes |                   |                   |         |
| Chronic pancreatitis | 174 (4.65) | 113 (4.53) | 61 (4.90) | 0.619 |
| Follow-up duration, yr | 6.13 ± 3.53 | 6.12 ± 3.52 | 6.17 ± 3.57 | 0.685 |

RAP: Recurrent acute pancreatitis; NTD: New Taiwan Dollars.
In this study, we identified RAP, alcohol consumption, age of onset of < 55 years, and smoking habit as the four major risk factors for developing CP during the mean follow-up period of 6.13 ± 3.53 years in the setting of adult patients with episodes of nonbiliary, nonobstructive AP. To our knowledge, this is the first population-based, large-scale cohort study to explore the risk factors for CP in the Chinese ethnic population\cite{14,15}. Furthermore, we developed the first prediction score model for CP, which is simple and useful in the clinical practice (Table 3).

The natural history of AP is still under debate, and the rate of progression from AP to CP varies and depends on the etiology of pancreatitis, with a mean interval of 3.5-5.5 years reported in the English literature\cite{4,16-19}. In our population-based cohort study, the rate of progression from AP to CP was 4.65% during the mean follow-up period of 6.13 ± 3.53 years (Table 1). This prevalence is similar to that reported by a prospective study conducted in Germany, with the rate of progression from AP to CP occurring in approximately 4% of all patients during a 20-year period\cite{4}. In our cohort, 12.8% (142/1094) of patients with RAP (nonobstructive, nonbiliary etiology) developed CP during the follow-up period, whereas only 1.2% (32/2645) of patients with only one episode of AP (nonobstructive, nonbiliary etiology) developed CP during the follow-up period (Figure 1). In a meta-analysis conducted in 2015 by Sankaran et al\cite{3}, it was observed 10% of patients with a first episode of AP and 36% of patients with RAP developed CP regardless of the etiology. In our study, the incidence of CP was much higher among patients who survived a second attack of

| Risk Score 1 | Model coefficient | Risk factor point |
|--------------|-------------------|------------------|
| Smoking related codes | 0.43 | 1 |
| Age of onset < 55 | 0.89 | 2 |
| Alcohol use-related codes | 1.13 | 3 |
| RAP (present or not) | 2.16 | 5 |
| Risk Score 2 | | |
| Smoking related codes | 0.39 | 1 |
| Age of onset < 55 | 0.71 | 2 |
| Alcohol use-related codes | 0.98 | 3 |
| Numbers of RAP | | |
| 2 episode | 1.62 | 4 |
| 3 episode | 2.14 | 5 |
| > 3 episode | 2.75 | 7 |
| Total | 13 |

| Risk Score 2 | Model coefficient | Risk factor point |
|--------------|-------------------|------------------|
| Smoking related codes | 0.43 | 1 |
| Age of onset < 55 | 0.89 | 2 |
| Alcohol use-related codes | 1.13 | 3 |
| RAP (present or not) | 2.16 | 5 |
| Risk Score 2 | | |
| Smoking related codes | 0.39 | 1 |
| Age of onset < 55 | 0.71 | 2 |
| Alcohol use-related codes | 0.98 | 3 |
| Numbers of RAP | | |
| 2 episode | 1.62 | 4 |
| 3 episode | 2.14 | 5 |
| > 3 episode | 2.75 | 7 |
| Total | 13 |

1Risk score was calculated by dividing each model coefficient by the smallest coefficient in the model and rounding the ratio to the integer. RAP: Recurrent acute pancreatitis.
Our multivariate analysis revealed that the risk of progression to CP was higher among patients with RAP, alcoholics, smokers, and younger patients with age of onset of < 55 years, as assessed using alcohol-use-related codes and smoking-related codes as surrogates.
for alcohol consumption and smoking habit, respectively. Recently, a majority of physicians have recognized that smoking habit is also an important risk factor for the development of CP along with RAP and alcohol...
consumption\textsuperscript{21,22}. In our population-based cohort study, the multivariate analysis reconfirmed that in addition to alcohol consumption and RAP, cigarette smoking was an independent risk factor for CP also in the Chinese ethnic population (aHR: 1.53; 95%CI: 1.04-2.25; \(P = 0.029\)).

Patients with RAP were assigned the highest score in our prediction model (Table 3). Patients with RAP in our cohort may represent a susceptible population possessing gene mutations or unfavorable alleles of the PRSS1, CFTR, SPINK1, CTRC, and CASR genes, making them more susceptible to environmental factors such as exposure to alcohol or cigarette smoking\textsuperscript{16,22-27}. Volker et al\textsuperscript{16} have reported that an interaction between the environmental and genetic factors (i.e., \(N34S + \) alcohol or PRSS1 + smoking) further increased the probability of the development of CP. In addition, Polonikov et al\textsuperscript{28} have shown that cigarette smokers with the -408CC genotype have an increased risk for AP [odds ratio (OR): 2.07], whereas nonsmoker carriers do not have the disease risk. Genetic risk factors are not rare among patients with CP, and approximately 25% of patients with CP exhibit some genetic risk factors\textsuperscript{16}. We hypothesized that a proportion of patients with RAP possess some genetic disorders, and the interaction between genetic and environmental factors, such as a smoking habit and alcohol consumption, accelerates the progression of CP. Thus, although a majority of patients with genetic risk factors display a very slow progression of the natural disease course and always present with RAP in their early life\textsuperscript{16}, environmental factors such as cigarette smoking may trigger or accelerate the development of CP in the background of genetic disorders.

We found that the age of onset of < 55 years is one of the risk factors for developing CP in this study. In a cross-sectional study, younger age (OR: 0.80; 95%CI: 0.68-0.94) was found to be independently associated with an increased risk of developing recurrent pancreatitis\textsuperscript{28}. In other words, that study demonstrated that age is a protective factor of recurrence. In a prospective study with a 30-year follow-up conducted by the Danish registries, it was observed that the risk of progression to CP decreased with increasing age in a dose-dependent manner with 2% less risk per year of age\textsuperscript{29}, suggesting that the age of onset is an important issue while evaluating the possibility of developing CP. Similar to our study cohort with an etiology of nonobstructive, nonbiliary pancreatitis, Peter layer et al\textsuperscript{29} published an article on Gastroenterology that disclosed that patients with early-onset and late-onset idiopathic CP had a median age at onset of symptoms of 19 and 56 years, respectively, whereas the majority of patients with alcoholic CP had a median age of onset of 43.9 years, suggesting that most of the patients with nonobstructive, nonbiliary CP have the disease onset in their middle age. The result of that study is consistent with our finding that the age of onset of < 55 years is one of the risk factors for developing CP.

In general, CP involves a persistent destructive, inflammatory process that eventually leads to an “irreversible” damage to the endocrine and exocrine functions of the pancreas. In contrast, early CP is a disease entity that propagated in 2009, and it has been reported that the disease course could be reversed if adequate intervention is taken\textsuperscript{30}. Using our prediction score model for CP, we were able to stratify our patients into different categories and arrange further examinations such as pancreatic functional test or endoscopic ultrasound after acute stage for the high-risk category (incidence rate of about 31 per 1000 person-years, based on our study) to detect CP\textsuperscript{31} as early as possible and determine the optimal follow-up interval for the patients with AP with a nonbiliary, nonobstructive etiology.

\textbf{Limitations of this study}

First, the definition of the disease was based on ICD-9-CM codes assigned by the NHIRD\textsuperscript{32,33}. However, the NHIRD data regarding the diagnosis of AP and other comorbidities have been used in relevant studies on AP and have been proved to be reliable\textsuperscript{34-38}. Moreover, to ensure the diagnosis of AP, we enrolled only hospitalized patients with a diagnosis of AP and excluded all the outpatients and patients with prior CP, biliary pancreatitis, and obstructive pancreatitis to avoid overestimation of the cases with AP. To ensure the objective of the diagnosis of CP, patients with CP included in our study were also required to have undergone comprehensive imaging studies such as dynamic abdominal CT or abdominal MRI within 3 mo before the diagnosis of CP. In other words, the diagnosis of CP in our study was not only dependent on ICD-9-CM coding, but it was also reconfirmed by a CT or MRI study.

Second, the personal behaviors of smoking habit and alcohol consumption were based on ICD-9-CM codes, which may have underestimated the actual prevalence of smoking and alcohol consumption. However, the prevalence of alcohol consumption and smoking habit in our study is very close to that reported by another national population-based study in Taiwan, which also used the alcohol-use-related codes and smoking-related codes as surrogates of alcohol consumption and smoking habit, respectively\textsuperscript{38}. Considering that the aim of our study was to investigate the weighting of patients’ behaviors and the number of AP episodes to the development of CP, the information extracted based on the ICD-9-CM codes was sufficient for our purpose. In addition, because we used conservative statistical methods to analyze the data, the impact of alcohol consumption and cigarette smoking could only be underestimated and not overestimated. Third, due to potential bias, the results derived from a prospective cohort study are generally of a lower statistical quality than those from perspective studies. Fourth, Because we have excluded those with biliary pancreatitis and obstructive pancreatitis at the beginning, the prediction score which we developed can not be applied in patients with biliary pancreatitis or obstructive pancreatitis. However, to assess the reliability of our results, we included the patients with biliary pancreatitis and obstructive pancreatitis in the sensitivity analysis.
The results of sensitivity analysis were consistent with those of our primary analyses, indicating the robustness of our study result (Supplementary Table 2).

In conclusion, the presence of RAP, alcohol consumption, age of onset of < 55 years, and smoking habit were found to be the important risk factors for the development of CP in our large, population-based cohort study in Taiwan. Using these parameters, we developed a simple predicting score model for CP, which could be applied easily when a clinician encounters a patient with nonbiliary, nonobstructive AP (most of them may belong to “toxic metabolic” or “idiopathic pancreatitis” according to the TiGAR-O classification). Unless more reliable biomarkers for CP are identified, we believe that this predicting score model could help clinicians in terms of decision-making, early detection of CP, and most importantly reversing the early CP and prevent the development of PDAC.

ARTICLE HIGHLIGHTS

Research background
Chronic pancreatitis (CP) involves a persistent destructive, inflammatory process that eventually leads to an irreversible damage to the endocrine and exocrine functions of the pancreas. CP has complications and poor prognosis, with the mortality rate being approximately two-fold higher than that in the general population. Incidence rate of pancreatic cancer is as high as 26-fold in patients with CP, suggesting that the risk of pancreatic cancer is significantly higher in subjects with CP. Therefore, it is critical to predict CP in patients with acute pancreatitis (AP).

Research motivation
The treatment option of CP is limited. We considered that the prediction of CP in high risk population and the early intervention in high risk category may decrease the disease burden and avoid the subsequent malignant change.

Research objectives
AP, recurrent AP (RAP), and CP are a continuum of diseases. However, only a small proportion of patients with AP progress to CP. It is critical to predict CP in patients with AP since CP is an important risk factor of pancreatic duct adenocarcinoma (PDAC).

Research methods
A total of 5971 patients with one or more episodes of AP (ICD-9-CM code 577.0) recorded in the inpatient claims data from 2000 to 2013 were identified from the database. A 4-year look-back period was applied from 1996 to 1999 to ensure that all cases in our cohort were newly diagnosed and to reduce false incident cases. Eventually, 3739 patients with nonobstructive, nonbiliary AP were included for subsequent analysis. Next, we developed a model to predict the progress to CP in randomly selected two-thirds of this cohort (derivation cohort) and validated the model in the remaining one-third of this cohort (validation cohort). Outcomes and comorbidities were identified based on ICD-9-CM codes. CP was defined using ICD-9-CM codes (ICD-9-CM code 577.1), and abdominal computed tomography (CT) or abdominal magnetic resonance imaging (MRI) performed within 3 mo before the diagnosis of CP was also considered as essential to diagnose clinical CP. The risk of CP in patients with nonobstructive, nonbiliary AP was estimated using the Cox proportional hazards model. The significant β coefficients from the Cox model with backward selection procedure were used to construct an integer-based risk score for stratifying the risk of progress to CP. The referent for each variable was assigned a value of 0, and the coefficients for the other variables were calculated by dividing by the smallest coefficient in the model and then rounding to the nearest integer. Individual scores were assigned by summing the individual risk factor scores, and the cumulative incidence rate of each score was calculated. For easy application in clinical practice, the total risk scores were classified into low-risk category, moderate-risk category, and high-risk category based on similar magnitudes of hazard.

Research results
The multivariate analysis revealed that the presence of RAP, alcoholism, smoking habit, and age of onset of < 55 years were the four important risk factors for CP. We developed a scoring system from the derivation cohort by classifying the patients into low-risk, moderate-risk, and high-risk categories based on similar magnitudes of hazard and validated the performance using another validation cohort. Using this score, we could predict the development of CP in high risk population and arrange further intervention in high risk category. However, for the lack of reliable biomarkers for predicting CP at present, we didn’t include biomarkers in our variables. In the future, if more sensitive biomarkers for CP could be identified, the biomarkers should be added into the prediction model to improve the predictive value.

Research conclusions
The presence of RAP, along with alcohol consumption, age of onset, and smoking habit have a high prediction value of CP. We developed a novel prediction score model for CP with excellent discrimination and also successfully validated this model in our study. Using this scoring system, a clinician can predict the outcome of a patient with AP and arrange further examination such as endoscopic ultrasound for the high-risk category (incidence rate of about 31 per 1000 person-years in high risk group), in order to early diagnosis of CP and the subsequent pancreatic cancer. Furthermore, healthcare providers can use this scoring system for assessing patient education in terms of alcohol and smoking abstinence, because the treatment option of CP is extremely limited.

Research perspectives
We confirmed that RAP, alcohol drinking, age of onset and smoking habit are important risk factors for CP in Chinese population. Using this score, we could predict the development of CP in high risk population and arrange further intervention in high risk category. For the lack of reliable biomarkers for predicting CP at present, we didn’t include biomarkers in our variables analysis. In the future, if more sensitive biomarkers for CP could be identified, the biomarkers should be added into the prediction model to improve the predictive value for CP in patients with AP episode.

REFERENCES
1. Banks PA. Epidemiology, natural history, and predictors of disease outcome in acute and chronic pancreatitis. Gastrointest Endosc 2002; 56: S26-S30. DOI: 10.1067/mge.2002.129022
2. Lowenfels AB, Maisonneuve P, Cavallini G, Ammann RW, Lankisch PG, Andersson Jr, Dimaggio EP, Andren-Sandberg A, Domellof L. Pancreatitis and the risk of pancreatic cancer. International Pancreatitis Study Group. N Engl J Med 1993; 328: 1433-1437. DOI: 10.1056/NEJM199305203282001
3. Sankaran SJ, Xiao AY, Wu LM, Windsor JA, Forsmark CE, Petrov MS. Frequency of progression from acute to chronic pancreatitis and risk factors: a meta-analysis. Gastroenterology. 2015; 149: 1490-1500. DOI: 10.1053/j.gastro.2015.07.066
4. Lankisch PG, Breuer N, Bruns A, Weber-Dany B, Lowenfels AB, Maisonneuve P. Natural history of acute pancreatitis: a long-term population-based study. Am J Gastroenterol 2009; 104: 2797-2805. DOI: 10.1038/ajg.2009.405
5. Maisonneuve P, Lowenfels AB, Muller-Hartmann B, Cavallini G, Lankisch PG, Andersson Jr, Dimaggio EP, Andren-Sandberg A, Domellof L, Frulloni L, Ammann RW. Cigarette smoking accelerates progression of alcoholic chronic pancreatitis. Gut 2005; 54: 510-514. DOI: 10.1136/gut.2004.039263
6. Yadav D, Hawes RH, Brand RE, Anderson MA, Money ME, Banks PA, Bishop MD, Baille J, Sherman S, DiSario J, Burton FR, Gardner TB, Amann ST, Getulio A, Lawrence C, Elinoff B, Greer JB, O’Connell M, Barnarda MM, Silvka A, Whitcomb DC; North American Pancreatic Study Group. Alcohol consumption, cigarette...
pancreatitis with/without intensive care: a nationwide population-based study in Taiwan. *Pancreas* 2011; **40**: 10-15 [PMID: 20938365 DOI: 10.1097/MPA.0b013e3181f7e750]

36 **Lai SW**, Muo CH, Liao KF, Sung FC, Chen PC. Risk of acute pancreatitis in type 2 diabetes and risk reduction on anti-diabetic drugs: a population-based cohort study in Taiwan. *Am J Gastroenterol* 2011; **106**: 1697-1704 [PMID: 21577242 DOI: 10.1038/aig.2011.155]

37 **Lin HY**, Lai JI, Lai YC, Lin PC, Chang SC, Tang GJ. Acute renal failure in severe pancreatitis: A population-based study. * Ups J Med Sci* 2011; **116**: 155-159 [PMID: 21250932 DOI: 10.3109/03009734.2010.547636]

38 **Shen HN**, Yang CC, Chang YH, Lu CL, Li CY. Risk of Diabetes Mellitus after First-Attack Acute Pancreatitis: A National Population-Based Study. *Am J Gastroenterol* 2015; **110**: 1698-1706 [PMID: 26526084 DOI: 10.1038/aig.2015.356]

39 **Lin YC et al.** Prediction score for chronic pancreatitis
