The 65th ASH Annual Meeting Abstracts

ORAL ABSTRACTS

332.THROMBOSIS AND ANTICOAGULATION: CLINICAL AND EPIDEMIOLOGICAL

Dynamic Modelling for Cancer-Associated Venous Thromboembolism

Simon Mantha, MDMPH[1], Rohan Singh[1], Jeffrey I. Zwicker, MD[2,1]

[1] Memorial Sloan Kettering Cancer Center, New York, NY
[2] Department of Medicine, Weill Cornell Medical College, New York, NY

Background: Cancer-associated venous thromboembolism (VTE) is an important source of morbidity and mortality. Current risk prediction models for thrombosis are handicapped by static applicability (i.e. one-time evaluation prior to the first cycle of chemotherapy). We hypothesized that event forecasting using time series modelling could account for important changes in clinical state, variations in treatment and aspects of cancer progression over time.

Methods: The cohort consisted of adult patients with a solid tumor enrolled in the MSK-IMPACT study between 2014 and 2019. VTE was defined as any symptomatic or asymptomatic pulmonary embolus or lower extremity deep vein thrombosis. We derived time series mixed-effects logistic regression models in order to predict cancer-associated VTE events using baseline patient characteristics (cancer type, time from diagnosis and age) along with time-dependent covariates including total white blood cell count, hemoglobin, platelet count, albumin, creatinine and parenteral chemotherapy type. Time-periods of one month were used and models were fitted to predict the presence or absence of a VTE episode in the following period (time = t+1) based on predictor values in the current period (time = t). The laboratory test result ascribed to a given period could be either the last value in the interval or the mean of all interval-specific values (both approaches tested). Missing laboratory results were imputed using the previously known value when available. Differences in laboratory values between periods t-1 and t were included as predictors in the initial model along with indicators of missingness (see Figure for flow of information between time-period cells). We hypothesized that absence of laboratory testing for a time-period would be a marker of stable medical state and hence associated with a lower risk of VTE. A previously defined training set comprising 80% of the cohort was used and different modelling approaches were evaluated using 5-fold cross-validation. The final selected model was evaluated in a separate dedicated validation set including the remaining 20% of the cohort.

Results: The original cohort included 35,391 patients, from which 26,075 individuals with no cancer-associated VTE episode before observation start and sufficient laboratory test information could be included in the analysis, comprising 276,579 months of observation and 1,463 VTE events. The initial logistic regression model including all available predictors was fitted on the training set, revealing associations between predictors and VTE risk largely consistent with data reported by other groups (see Table). After adjusting for multiple comparisons, significant predictors of cancer-associated VTE identified included albumin, cell count result missingness (decreased risk for missing value), treatment with a platinum analogue (increased risk), hemoglobin, white blood cell count, breast cancer (decreased risk) and high-grade glioma (increased risk). Based on model metrics and prior knowledge from published literature, the set of covariates selected for inclusion in the final model consisted of cell counts along with the corresponding missingness indicator, platelet count change over time, albumin and chemotherapy types. The c-index for the final model using 5-fold cross-validation in the training set was 0.77 (95% CI = 0.75-0.80), compared to 0.77 (95% CI = 0.74-0.79) in the validation set. The latter corresponded to a specificity of 0.83 (95% CI = 0.79-0.86) when fixing sensitivity at 50%.

Conclusions: A simple mixed-effects logistic regression time series model demonstrated promising potential to estimate the risk of cancer-associated VTE on a dynamic basis using routine laboratory test results and chemotherapy prescription data widely available in most electronic health record systems. Validation of this dynamic VTE risk model has the potential to optimize outpatient thromboprophylaxis during periods of high thrombotic risk.

Figure: Flow of Information Between Time-Period Cells
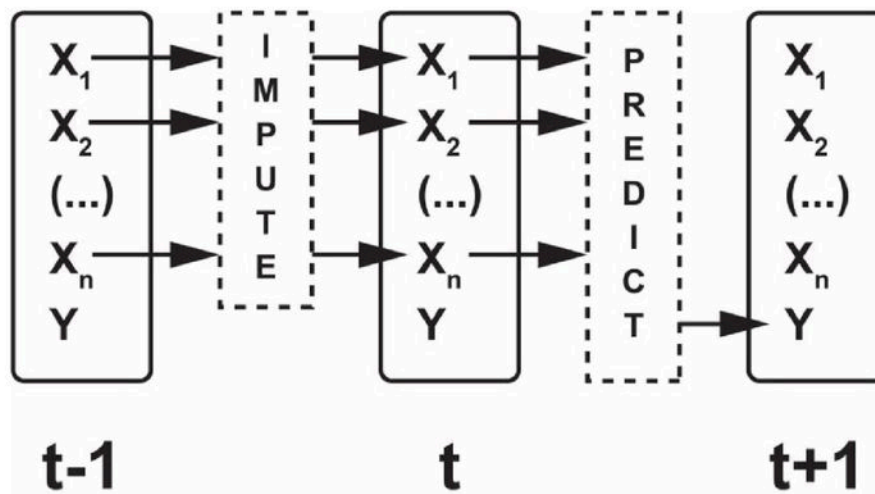
Predictors ($X_{1:n}$) and VTE outcome status (Y) are categorized into a series of successive ordered non-overlapping 30-day time-periods. Missing laboratory test results are imputed with the most recent prior values when available. VTE outcome in the following month is predicted based on predictor values in the current month.

Table: Significant Predictors of Cancer-Associated VTE

Predictor	OR	p-value	q-value
Albumin (g/dL)	0.32	<0.001	<0.001
Any missing cell count	0.38	<0.001	<0.001
Administration of a platin analogue	1.60	<0.001	<0.001
Hemoglobin (g/dL)	0.90	<0.001	<0.001
Total white blood cell count (x 1,000 cells/mcL)	1.02	<0.001	0.002
Breast cancer	0.45	<0.001	0.004
High-grade glioma	2.24	<0.001	0.02

P-values adjusted for multiple comparisons with the Benjamini-Hochberg method. Only covariates with q-value <0.10 are shown.

OR: odds ratio of VTE for characteristic present vs not in the case of binary predictors or for any increase in one unit for continuous variables.

Figure 1