

# Automated versus Manual Data Extraction of the Padua Prediction Score for Venous Thromboembolism Risk in Hospitalized Older Adults

Juliessa M. Pavon<sup>1,2,3</sup> Richard J. Sloane<sup>1,2,3</sup> Carl F. Pieper<sup>1,2,3</sup> Cathleen S. Colón-Emeric<sup>1,2,3</sup>  
Harvey J. Cohen<sup>1,2,3</sup> David Gallagher<sup>1</sup> Miriam C. Morey<sup>1,2,3</sup> Midori McCarty<sup>1</sup> Thomas L. Ortel<sup>1</sup>  
Susan N. Hastings<sup>1,2,3,4</sup>

<sup>1</sup>Duke University, Durham, North Carolina, United States

<sup>2</sup>Geriatric Research Education Clinical Center, Durham  
Veteran Affairs Medical Center, Durham, North Carolina, United States

<sup>3</sup>Duke University Claude D. Pepper Center, Duke University, Durham,  
North Carolina, United States

<sup>4</sup>Health Services Research and Development Center of Innovation,  
Durham Veterans Affairs Health Care System, Durham,  
North Carolina, United States

**Address for correspondence** Juliessa M. Pavon, MD, Division of  
Geriatrics, Duke University Health System, Box 3003, Durham, NC  
27710, United States (e-mail: Juliessa.pavon@duke.edu).

CME/MOC-II\*

Appl Clin Inform 2018;9:743–751.

## Abstract

**Objective** Venous thromboembolism (VTE) prophylaxis is an important consideration for hospitalized older adults, and the Padua Prediction Score (PPS) is a risk prediction tool used to prioritize patient selection. We developed an automated PPS (APPS) algorithm using electronic health record (EHR) data. This study examines the accuracy of APPS and its individual components versus manual data extraction.

**Methods** This is a retrospective cohort study of hospitalized general internal medicine patients, aged 70 and over. Fourteen clinical variables were collected to determine their PPS; APPS used EHR data exports from health system databases, and a trained abstractor performed manual chart abstractions. We calculated sensitivity and specificity of the APPS, using manual PPS as the gold standard for classifying risk category (low vs. high). We also examined performance characteristics of the APPS for individual variables.

**Results** PPS was calculated by both methods on 311 individuals. The mean PPS was 3.6 (standard deviation, 1.8) for manual abstraction and 2.8 (1.4) for APPS. In detecting patients at high risk for VTE, the sensitivity and specificity of the APPS algorithm were 46 and 94%, respectively. The sensitivity for APPS was poor (range: 6–34%) for detecting acute conditions (i.e., acute myocardial infarction), moderate (range: 52–74%) for chronic conditions (i.e., heart failure), and excellent (range: 94–98%) for conditions of obesity and restricted mobility. Specificity of the automated extraction method for each PPS variable was > 87%.

**Conclusion** APPS as a stand-alone tool was suboptimal for classifying risk of VTE occurrence. The APPS accurately identified high risk patients (true positives), but lower scores were considered indeterminate.

## Keywords

- ▶ decision support algorithm
- ▶ venous thromboembolism
- ▶ geriatrics
- ▶ electronic health record and systems
- ▶ inpatient

\* To earn credit, visit AMIA for details.

received  
June 1, 2018  
accepted after revision  
August 9, 2018

© 2018 Georg Thieme Verlag KG  
Stuttgart · New York

DOI <https://doi.org/10.1055/s-0038-1670678>.  
ISSN 1869-0327.

## Introduction

Venous thromboembolisms (VTEs), which encompass deep vein thrombosis and pulmonary embolism, are serious health threats for hospitalized patients.<sup>1–6</sup> Clinical practice guidelines recommend the use of risk prediction tools, such as the Padua Prediction Score (PPS), to classify risk of VTE and select the most suitable population for anticoagulant prophylaxis.<sup>6–8</sup> Without routine use of risk prediction tools, pharmacologic VTE prophylaxis is either under- or overprescribed in medically ill hospitalized patients.<sup>9–11</sup>

The PPS is calculated based on 11 clinical criteria that are weighed and summed to a score that stratifies patients as either high or low risk for VTE occurrence.<sup>8</sup> To use PPS, a clinician can use an online calculator for computation but must input a response for each individual clinical criteria. This can be cumbersome and disruptive to the clinical workflow.<sup>12,13</sup> To facilitate the use of the PPS at the point of care, automated algorithms can extract data already captured in discrete fields in the electronic health records (EHRs) during routine care. However, automated identification of PPS variables largely depends on the International Classification of Diseases 9th/10th Revision (ICD-9/10) coding. Given that diagnostic coding may be inaccurate or omit relevant conditions,<sup>14–19</sup> the performance of automated algorithms is unknown.

## Objective

VTE anticoagulation prophylaxis may not be beneficial or cost-effective in all cases, and may be inappropriately used in some older patients. Therefore, we embarked on a larger study to determine the magnitude and scope of inappropriate use of anticoagulants for VTE prophylaxis in low risk older adults. As part of our approach, we developed an automated PPS (APPS) algorithm for risk stratification of hospitalized older adults and compared it to PPS calculated from traditional manual chart review. We examined sensitivity and specificity of the APPS for VTE risk assessment and compared it to the gold standard, manual PPS. We also examined the sensitivity and specificity of individual clinical diagnoses comprising the PPS to inform the development of other automated EHR algorithms.

## Methods

### Study Design and Setting

This retrospective cohort study used data from patients aged 70 years or older admitted to an academic medical center's general medicine service over a 1-year period (January 1, 2014–December 31, 2014). We were interested in adults over the age of 70, as they have higher frequency of comorbidities and also have higher bleeding risk related to anticoagulant therapy. Thus, weighing risks and benefits of pharmacologic VTE prophylaxis use is particularly challenging in this population. Patient exclusion included length of stay < 24 hours and receiving therapeutic doses of anticoagulants prior to admission, as these patients are not in need of VTE prophylaxis.

### Padua Prediction Score

A PPS was calculated for all patients based on clinical and demographic factors available from prior encounters and within the first 24 hours of admission to coincide with the data available to clinicians at the time they make VTE prophylaxis decisions. Fourteen clinical variables were collected for each subject to determine their PPS (→Table 1). Individual scores were assigned based on a 20-point scoring system of risk factors including: active cancer, previous VTE, myocardial infarction (MI), congestive heart or respiratory failure, trauma surgery, age ≥ 70 years, reduced mobility, acute infectious or rheumatologic disease, thrombophilic condition, obesity, and hormone use (see →Table 1). A score of < 4 signifies that a patient is considered low risk.<sup>8</sup>

### Automated Extraction of VTE Risk

Variables used to calculate the APPS were identified electronically from the EHR of the academic medical center via an administrative data portal, the Duke Enterprise for Data Unified Content Explorer (DEDUCE). The academic medical center's EHR is powered by EPIC of Verona, Wisconsin, United States. DEDUCE is a data warehouse and research tool that provides investigator access to patient-level administrative clinical information including medications and clinical encounters across the health system. The algorithms were developed by a data analyst and verified by the study team. We used the PPS validation paper as our guide for operationalizing variables except where further interpretation was required.<sup>8</sup> All variable definitions used criteria that could be applied to the structured data in DEDUCE, and relied only on the use of data elements available to providers at the time of admission, which is typically when a PPS is calculated and clinical decisions are made about VTE prophylaxis. As shown in →Table 1, algorithm definitions for previous VTE, heart failure, respiratory failure/chronic obstructive pulmonary disease (COPD), rheumatological disorder, and thrombophilic condition used ICD-9 diagnostic codes found in the Admitting Diagnoses or Problem lists. Acute conditions related to the current hospitalization, such as active cancer, MI, stroke, and infection were captured using algorithm definitions based on ICD-9 codes found solely in the Admitting Diagnoses list (→Table 1). We limited the data source to only the Admitting Diagnoses list to reduce the possibility of capturing remote histories of cancer or MIs that may have been left on a problem list, but were no longer active or related to the current hospital stay. The PPS variable “reduced mobility” was attributed to anyone with an activity order for bed rest within the first 24 hours of admission (→Table 1). The variable “obesity” was ascribed to any patient with a body mass index of ≥ 30 kg/m<sup>2</sup> within 1 year of admission or the ICD-9 code for obesity present on the Admitting Diagnoses or Problem lists. Estrogen-based hormone use was extracted from the outpatient and inpatient pharmacy orders using therapeutic class and pharmaceutical subclass names of estrogens, estrogen-progestin, and vaginal estrogens (→Table 1). Since our sample consisted only of patients aged 70 and older, everyone in the sample received 1 PPS point for the “age” variable. However, we could not

**Table 1** Sources and definitions of the automated Padua Prediction Score variables

Variable	Score <sup>d</sup>	Automated		Manual
		Source	Definition	Source
Active cancer <sup>a</sup>	3	Admitting Diagnosis list in EHR	ICD-9 149.0–172.99, 174.0–209.99	Encounter reason for admission, admitting physician history & physical (H&P) admission note at time of index hospital stay (HPI, problem list, A/P); oncology notes in the past 6 months
Previous VTE	3	Problem list and Admitting Diagnosis list in the EHR	ICD-9 415.1, 415.19, 453.8, 453.9, 671.31–671.50	Encounter reason for admission, admitting physician H&P admission note
Reduced mobility <sup>b</sup>	3	Physician orders in EHR	Bed rest (complete or strict), bed rest with bathroom privileges, bed rest with bedside commode privileges	Physician orders in EHR
Thrombophilic condition <sup>c</sup>	3	Problem list and Admitting Diagnosis list in the EHR	ICD-9 289.81	Encounter reason for admission, admitting physician H&P admission note
Recent trauma/surgery	2	NA	NA	Surgical procedure encounter, surgical notes, H&P note, or surgical history in H&P note within 30 days of admission
Age	1	Demographics in EHR	Age from date of birth	Demographics in EHR
Heart failure	1	Problem list and Admitting Diagnosis list in the EHR	ICD-9 428.x	Encounter reason for admission, Admitting physician H&P admission note
Respiratory failure	1	Problem list and Admitting Diagnosis list in the EHR	ICD-9 490–494.99, 496.x	Encounter reason for admission, admitting physician H&P admission note
Acute myocardial infarction	1	Admitting Diagnosis list in EHR; listed as primary reason	ICD-9 410.x	Encounter reason for admission, admitting physician H&P admission note
Acute stroke	1	Admitting Diagnosis list in EHR; listed as primary reason	ICD-9 430–438.99	Encounter reason for admission, admitting physician H&P admission note
Acute infection	1	Admitting Diagnosis list in EHR	ICD-9 1.x-136.99, 460–466.99, 480–488.99, 680–686.99, 995.9, 995.90–995.94,	Encounter reason for admission, admitting physician H&P admission note
Rheumatological disorder	1	Problem list and Admitting Diagnosis list in the EHR	ICD-9 274.9, 696, 696, 701.0, 715.x 725.x, 710.0–715.99, 720.x	Admitting physician H&P admission note
Obesity	1	Problem list in EHR and BMI in EHR vitals	ICD-9 278.x or BMI $\geq$ 30	Admitting physician H&P admission note and admission height and weight in EHR vitals
Hormone treatment	1	Pharmacy orders in EHR, outpatient and inpatient	Therapeutic class hormones, pharmaceutical subclass estrogens, estrogen–progestin, vaginal estrogens	Admission medication list in admitting physician H&P admission note

Abbreviations: A/P, assessment and plan; BMI, body mass index; EHR, electronic health record; HPI, history of present illness; ICD, International Classification of Disease; VTE, venous thromboembolism.

<sup>a</sup>Patients with local or distant metastases and/or in whom chemotherapy or radiotherapy has been performed in the previous 6 months.

<sup>b</sup>Bed rest with bathroom privileges (either due to patient's limitations or on physician's order) for at least 3 days.

<sup>c</sup>Carriage of defects of antithrombin, protein C or S, factor V Leiden, G20210A prothrombin mutation, and antiphospholipid syndrome.<sup>8</sup>

<sup>d</sup>Low risk of VTE: score < 4.

operationalize the PPS variable “recent ( $\leq 1$  month) trauma and/or surgery,” because trauma was not easily defined using ICD-9 codes, and surgical procedures could not be efficiently identified across all Current Procedural Terminology codes. Therefore, this variable was not included in the APPS or manual PPS scoring.

### Manual Data Extraction of VTE Risk

A trained abstractor performed manual chart abstractions, considered the gold standard for calculating the PPS as this is how it was abstracted in the validation paper.<sup>8</sup> The abstractor was instructed to review the EHR for all PPS comorbidity variables using structured encounter data (e.g., Reason for Admission and Problem Lists) and unstructured data (e.g., documentation of present illness, assessments, and plans found in the physician admission note). Problem lists are updated and maintained at the system level and autopopulated into the admission notes. To test the reliability of the manual abstraction protocol, a second coder independently verified a sample of 20 randomly selected patient charts. All findings were recorded using the Research Electronic Data Capture platform (REDCap Consortium, Vanderbilt University, Nashville, Tennessee, United States). This study was approved by the Institutional Review Board.

### Analysis

To describe the accuracy of the APPS, we examined the sensitivity and specificity of the automated algorithm and compared with the manual review. Performances of the two methods were examined for individual variables and the overall risk classification based on their total scores. We calculated (1) the proportion of high risk patients according to the manual review who were also classified as high risk by the APPS (sensitivity), and (2) the proportion of low risk patients according to the manual extraction method who were also classified as low risk patients by the APPS (specificity). Because the “recent trauma or surgery” variable was not operationalized in the automated algorithm, the overall accuracy for the total score of the PPS was based on a maximum possible score of 19 rather than 20. All analyses were performed using SAS Version 9.4 (SAS Institute, Cary, North Carolina, United States).

### Results

Of the total eligible cohort ( $N = 1,399$ ), 400 patients were randomly selected for manual chart review; 89 of these patients were not eligible because they were on anticoagulants upon admission, leaving  $n = 311$  patients in the analytic sample. Mean age for the sample was 80.6 years (standard deviation [SD], 7.3); 42% were male; and 34% were African American. Median length of stay was 4.0 days (→Table 2). According to the manual PPS, 41% ( $n = 129$ ) of patients had a total score of 4 or more and so were classified as high risk for VTE.

The mean manually abstracted PPS and APPS were 3.6 (SD, 1.8) and 2.8 (1.4), respectively. Manual PPS identified a

**Table 2** Characteristics of the study patients

	Overall sample, $n = 311$ Mean (SD) or $n$ (%)
Demographics	
Age	80.6 (7.3)
Gender (Male)	131 (42)
Race (African American)	105 (34)
Ethnicity (Hispanic)	2 (0.6)
Hospital characteristics	
Length of stay, (days), median	4
Admission diagnosis	
1. Acute infection	37 (12)
2. Altered mental status	28 (9)
3. Cardiovascular disease	9 (3)
4. COPD	9 (3)
5. Kidney failure	3 (1)
6. Dysphagia	50 (16)
7. Cancer	3 (1)
8. Other diagnoses <sup>a</sup>	171 (55)

Abbreviations: COPD, chronic obstructive pulmonary disease; SD, standard deviation.

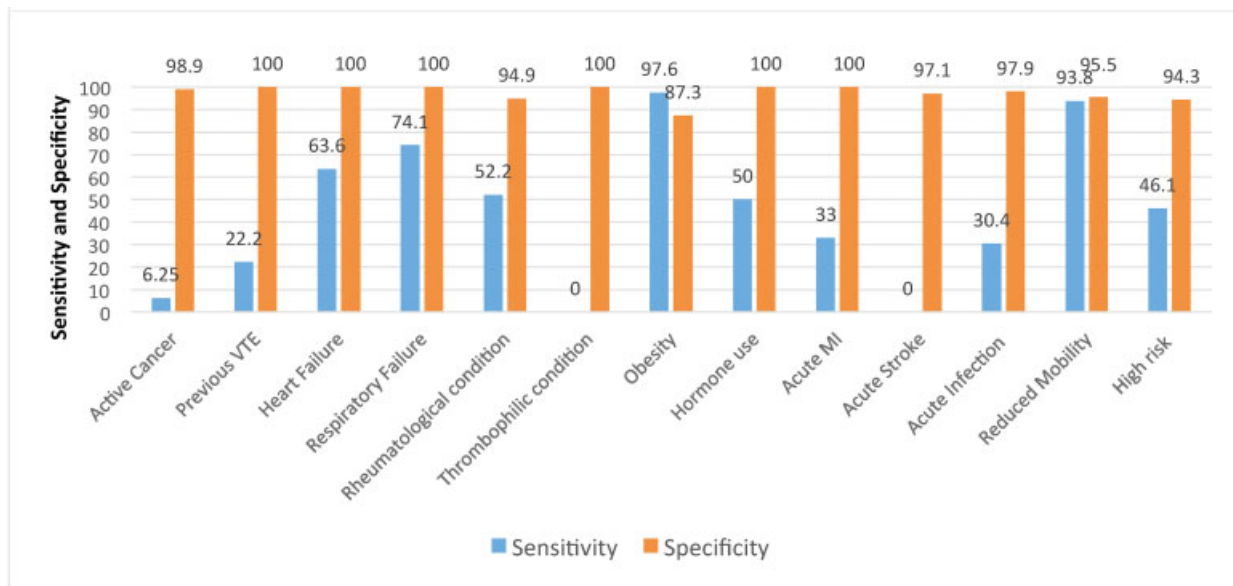
<sup>a</sup>Examples of other diagnoses include: gastrointestinal conditions and symptoms, cardiac/respiratory symptoms (i.e., dyspnea, chest pain, cough), neurological symptoms (i.e., difficulty walking, gait instability, weakness, falls), and anemia, missing  $n = 1$ .

higher percentage of patients at high risk for VTE compared with APPS (41% vs. 24%, respectively). Sensitivity of the APPS algorithm for detecting patients at high risk for VTE was 46% (→Fig. 1); 17% of high risk patients were misclassified as low risk by the automated algorithm. However, the algorithm was highly specific (94%) and rarely misclassified a patient at low risk for VTE as being high risk (false positives).

The performance of individual APPS variables fluctuated. The automated extraction had poor sensitivity (range: 6–34%) when capturing acute conditions (MI, infection, and active cancer; →Fig. 1, →Table 3). The automated extraction had moderate sensitivity (range: 52–74%) when identifying chronic conditions (COPD, heart failure, and rheumatological diseases; →Fig. 1, →Table 3). For conditions that relied on structured data fields (obesity and restricted mobility), the APPS algorithm had excellent sensitivity (range: 94–98%; →Fig. 1, →Table 3). Specificity of the APPS algorithm was > 87% for each PPS variable (→Fig. 1).

### Discussion

In this study of 311 hospitalized older adults, an APPS had low sensitivity but high specificity in classifying patients as high risk for VTE. Whereas a score of 4 or more on the APPS indicating a “high risk” patient was reliable, there were considerable false negatives in the “low risk” group. A previous study comparing APPS to manual calculation found observed mean scores to be slightly higher in the manual group, although the difference was not statistically



**Fig. 1** Sensitivity and specificity of automated Padua Prediction Score individual variables and risk categorization. High risk, Padua Prediction Score  $\geq 4$ ; MI, myocardial infarction; PPS, Padua Prediction Score; VTE, venous thromboembolism.

**Table 3** Comparison between the automated and manual Padua Prediction Score

	Automated PPS extraction <i>N</i> = 311, <i>n</i> (%)	Manual PPS extraction <i>N</i> = 311, <i>n</i> (%)
Active cancer	7 (2)	33 (11)
Recent trauma		32 (10)
Previous VTE	6 (2)	28 (9)
Heart failure	51 (16)	84 (27)
Respiratory failure	63 (20)	94 (30)
Rheumatological illness	88 (28)	130 (42)
Thrombophilic history	0 (0)	1 (0.3)
Obesity	143 (46)	101 (33)
Ongoing hormone use	8 (3)	5 (2)
Acute MI	2 (0.6)	4 (1)
Acute stroke	10 (3)	3 (1)
Acute infection	71 (23)	139 (45)
Restricted mobility	40 (13)	33 (11)
Padua Score, mean SD <sup>a</sup>	2.8 (1.4)	3.6 (1.8)
High risk, score $\geq 4$	75 (24)	129 (41)

Abbreviations: MI, myocardial infarction; PPS, Padua Prediction Score; SD, standard deviation; VTE, venous thromboembolism.

<sup>a</sup>Automated PPS out of possible 19 points, manual PPS out of possible 20 points.

significant (mean score, 5.1 [SD, 2.6] vs. 5.5 [SD, 2.1], respectively,  $p = 0.073$ ).<sup>20</sup> That study did not report proportion of each group that would have been identified as high or low risk for VTE occurrence, so we cannot compare directly to our findings. Our primary objective was to examine how accurately the automated algorithm predicted high- versus low risk stratification of patients, as this is how the PPS is used clinically.

Successful translation of PPS clinical criteria for use in a subsequent EHR system requires carefully constructed variable definitions, which in turn affects the sensitivity and specificity of individual PPS items. The poor sensitivity of our algorithm is partially explained by the reliance on diagnostic codes and structured data from the EHR. Similar to results by Elias et al, we found that the differences in the performance between manual and automated data extraction methods for the PPS were largest among the individual variables for acute conditions, such as active cancer, acute infection, as well as for previous VTE.<sup>20</sup> The automated definitions for these variables rely heavily on diagnostic coding in the EHR. Risk for undercoding diminishes the likelihood that automated algorithms will capture certain PPS conditions. Previous studies show that diagnostic coding in the EHR and administrative databases is often inaccurate, with poor sensitivity being a particular issue.<sup>14,21–23</sup> For example, the sensitivity of VTE diagnosis using ICD-9/10 coding is 75%.<sup>15</sup> With our algorithm, 7% of patients identified with no history of VTE by the automated algorithm did have prior VTE. Constructing a variable definition for acute infection solely using ICD-9/10 codes can also lead to miscategorization. Upon admission, many infections are coded using symptom categories (e.g., fevers, shortness of breath, cough, etc.) rather than by actual infectious conditions (e.g., urinary tract infection, pneumonia, etc.).<sup>24</sup> Errors in risk miscategorization are common

among risk instruments, with true cases often being misclassified as “controls.” This tends to bias the risk assessment tools to underclassification.<sup>17,25</sup> Both builders and users of APPS algorithms need to be aware of individual variables that are hard to operationalize, and of poorly sensitive items. Efforts to improve sensitivity of automated extraction for PPS conditions are needed that address other common reasons why the automated algorithm did not identify the variables with poor sensitivity. This might include using more comprehensive lists and combinations of diagnosis codes, other data elements such as medications, and laboratory values, and other extraction methods such as machine learning algorithms. Our results suggest that the APPS rarely misclassifies a patient as high risk. However, others with an APPS score of < 4 should be considered as indeterminate risk and a score recalculated after clinician input of poorly sensitive items.

The data sources used to extract variables, such as Problem Lists, Orders, or Flowsheets, also influence the accuracy of APPS variables. For our automated algorithm, the search for active cancer was restricted to ICD-9 codes on the Admission Diagnoses list. This decision resulted in the automated algorithm missing almost 94% (15/16; sensitivity 6%) of patients with active cancer. Expanding the data sources to Problem lists would have increased the sensitivity at a possible cost of decreased specificity. Use of other proxy measures in the EHR, such as prescription for chemotherapy medications in the preceding 3 months, would be another approach to attempt to increase the sensitivity. In a second example, the original PPS manuscript did not define “immobile.”<sup>8</sup> Our automated algorithm assigned immobility to anyone having a bed rest order following definitions used in previous PPS studies.<sup>26</sup> This definition has resulted in 5 to 11% of patients being identified as immobile. If defined using nursing documentations, such as Flowsheets, the percent of patients classified as immobile is much higher.<sup>20</sup> The use of nursing documentation, however, requires sophisticated algorithm programming which may prevent local hospitals and health systems from implementing APPS. These examples illustrate the importance of standardizing algorithm definitions and data sources to optimize the sensitivity and specificity of results.

### Limitations

There are several limitations of this study. The first is generalizability; this is a single-center study and our results are specific to an EPIC-based EHR. Second, this study uses retrospective data; future studies will be required to test algorithms prospectively. Third, a limitation of using only diagnostic codes for variable definitions is the sensitivity and specificity of the diagnostic code itself related to undercoding (e.g., absence of a specific code does not mean absence of the condition), or overcoding (e.g., clinicians selecting a condition by best match to help communicate care decisions though it may not entirely reflect the presenting condition). Fourth, ICD-10 has now replaced ICD-9, and it is unclear whether this will affect the results. Although the validity of ICD-9 and ICD-10 administrative data in coding clinical

conditions was generally similar when tested in 2008, the validity of ICD-10 codes was expected to improve as coders gained experience with them.<sup>19</sup> As a result, automated algorithms like ours would be expected to perform better using ICD-10 coding. Fifth, the variable recent trauma and/or surgery was not included in our analysis. Future studies should examine incorporating this variable into automated algorithms by using proxy measures, such as admission to a surgical or trauma service in the preceding 30 days, or scheduled operating room times. Lastly, patients with acute stroke and acute MI are underrepresented in our sample because at our medical center they are admitted to subspecialty services rather than to general medicine.

### Implications

Our results do not support use of an APPS as a stand-alone tool to predict VTE risk; however, the specific test characteristics revealed (low sensitivity/high specificity) suggest the utility of a hybrid approach combining automated and manual data inputs. For example, PPS values that can be extracted automatically with high reproducibility could be autopopulated, leaving a more concise set of data values for clinicians to extract. As an illustration of this, during the admission order entry for a patient, a screen could be presented to the clinician asking them to fill in PPS data elements that the algorithm has not been able to determine, and then proceed to calculate the PPS. Instead of having to answer all 11 PPS variable fields, this hybrid method would require fewer manual inputs. Alternatively, a “high risk” designation based on the APPS could be considered actionable, and any scores below this threshold considered indeterminate and in need of further clinician review. Both of these examples could save time while improving recognition of high risk individuals, because in current practice, all VTE risk assessments and decision-making about prophylaxis are based on clinician review. Our results suggest that the following considerations should be made when developing and testing PPS algorithms: (1) determine inclusion or exclusion of diagnostic codes and their data sources (i.e., Problem lists, Admission Diagnosis), which can affect false negative and false positive rates, and consider adding other EHR data elements (i.e., medications, laboratories, notes), so that more data can be available to an algorithm to achieve better results; (2) create structured data fields for essential elements of the PPS to facilitate automated extraction; and (3) explore other methods of extraction, such as the use of natural language processing algorithms to extract data from texts in physician’s notes, as they contain a large part of clinically relevant information.<sup>27</sup> In particular, we found that some algorithm elements, such as active cancer, previous VTE, and acute infection, are more difficult than others to capture.

### Conclusion

The accuracy of the APPS was suboptimal as a stand-alone tool for classifying risk of VTE occurrence. The APPS accurately identified patients as high risk for VTE occurrence

(true positives), but lower scores were indeterminate. Future work should examine how other areas of the EHR and other data extraction methods would improve the sensitivity of the APPS. A hybrid method combining both automated and manual data extraction methods merits further study as a method for improving recognition of VTE risk and decision-making around VTE prophylaxis.

### Clinical Relevance Statement

We designed an automated Padua Prediction Score (PPS) that had low sensitivity but high specificity in classifying patients as high risk for venous thromboembolism (VTE). A “high risk” designation based on the automated PPS could be considered actionable, and any scores below this threshold considered indeterminate. PPS values that can be extracted automatically with high reproducibility could be autopopulated, leaving a more concise set of data values for clinicians to review, thereby saving time while improving recognition of high risk individuals.

### Multiple Choice Questions

1. You are building an automated Padua Prediction Score algorithm, and the variable “active cancer” needs to be operationalized. What is the best strategy to improve the sensitivity of the algorithm for capturing “active cancer”?
  - a. Manually extract information from the provider notes.
  - b. Use a limited and narrow list of ICD-9/10 cancer diagnostic codes.
  - c. Use a proxy measure, such as orders for chemotherapy in the prior 3 months.
  - d. Use dates of encounter instead of ICD-9/10 diagnostic codes.

**Correct Answer:** The correct answer is option c. Previous studies show that diagnostic coding in the EHR and administrative databases is often inaccurate, with poor sensitivity a particular issue.<sup>23</sup> Use of other proxy measures in the EHR, such as prescription for chemotherapy medications in the preceding 3 months, would be another approach to attempt to increase the sensitivity of the variable “active cancer” (c). Manual extraction of information can be cumbersome (a). Efforts to improve sensitivity of automated extraction include using a more comprehensive list of ICD-9/10 diagnostic codes, and other elements of the EHR (b). Dates of encounter could be used in addition to ICD-9/10 diagnostics rather than in place of diagnostic codes (d).

2. Relying only on diagnostic codes may not result in a useful automated algorithm that provides real-time clinical decision support. Including other types of EHR data may improve the diagnostic accuracy. Of the following, which is most likely to require advanced data extraction techniques from which to infer diagnoses?
  - a. Laboratory data.
  - b. Medications.

- c. Provider notes.
- d. Vital signs.

**Correct Answer:** The correct answer is option c. Provider notes would require advanced machine learning data extraction techniques such as natural language processing. Laboratory data, medications, and vital signs can be extracted with simpler rule-based programming algorithms (a, b, d).<sup>28</sup>

3. Based on the results of this study, you are asked to build a hybrid PPS algorithm combining automated and manual data inputs. The strategy that will best improve adoption of the clinical decision support (CDS) tool would be:
  - a. Autopopulate all variables that use diagnoses codes, and use manual entry for variables that use orders or medication data.
  - b. Make the algorithm available to clinicians by selecting it from a menu of available tools.
  - c. Require clinician manual input of variables that can be accurately obtained only from them, and automate input of highly specific variables.
  - d. Require clinicians to review the accuracy of highly specific variables identified by the automated algorithm.

**Correct Answer:** The correct answer is option c. According to Bates et al, clinical decision support tools should anticipate needs and deliver in real time, as well as fit into the user’s workflow.<sup>29</sup> Therefore, making clinicians go to a menu of available tools is a task outside of their workflow (b). Another stated commandment is that you should ask for information only when you really need it. Requiring clinical manual input only for variables they can accurately obtain is an example of this (c). However, manually entering well-performing variables (a) or requiring clinicians to also review well-performing variables by the automated algorithm (d) are examples of asking for information that may not be really needed.

4. You are implementing a CDS tool in your EHR to assess patient VTE risk. This tool has high specificity for detecting high risk patients—94% of patients assigned a score  $\geq 4$  are high risk. In contrast, the tool has lower sensitivity and assigns a score of  $< 4$  to a substantial number of high risk patients. The best way to support users is to:
  - a. Report results as “high risk” or “low risk.” When a patient is assigned a score of  $\geq 4$ , it automatically places an order for VTE prophylaxis.
  - b. Report results as “high risk” or “low risk.” When a patient is assigned a score of  $< 4$ , it automatically moves to the next screen in the EHR for the clinician to enter laboratory and diet orders.
  - c. Report results as “high risk” or “indeterminant.” When a patient is assigned a score of 3, it prompts the user to manually update specific key input variables for which the automated algorithm has poor sensitivity.
  - d. Report results as “high risk” or “indeterminant.” When a patient is assigned a score of 3, it prompts the user to close out of the tool and calculate the VTE risk score manually.

**Correct Answer:** The correct answer is option c. As per our results, this CDS accurately identified patients as high risk for VTE occurrence (true positives), but lower scores were indeterminate. Therefore, results should not be reported as “high risk” or “low risk” (a, b). A hybrid approach for a CDS refers to an approach by which the values that can be extracted automatically with high reproducibility could be autopopulated, leaving a more concise set of data values for clinicians to extract (c). This is also in line with the keys to a successful CDS tool.<sup>29</sup> However, instead of having to answer 11 variable fields, this hybrid method would require fewer manual inputs by providers, and that would be meaningful in terms of saving time (d).

#### Protection of Human and Animal Subjects

This study was performed in compliance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects and was approved by the Duke University Institutional Review Board.

#### Funding

This study received funding from the following sources: NIA GEMSTAR Award (R03AG048007) (Pavon); Duke Older Americans Independence Center (NIA P30 AG028716–01); Duke University Internal Medicine Chair’s Award; Duke University Hartford Center of Excellence; Center of Innovation for Health Services Research in Primary Care (CIN 13–410) (Hastings) at the Durham VA Health Care System; T. Franklin Williams Scholars Program (Pavon); and K24 NIA P30 AG028716–01 (Colon-Emeric). The funding sources had no role in the design and conduct of the study, analysis or interpretation of the data, preparation or final approval of the manuscript before publication, or decision to submit the manuscript for publication.

#### Conflict of Interest

None.

#### Acknowledgment

The authors would like to thank Shenglan Li from Research Triangle Institute for her assistance in the data programming and database creation.

#### References

- Cohen AT, Agnelli G, Anderson FA, et al; VTE Impact Assessment Group in Europe (VITAE). Venous thromboembolism (VTE) in Europe. The number of VTE events and associated morbidity and mortality. *Thromb Haemost* 2007;98(04):756–764
- Kahn SR, Hirsch A, Shrier I. Effect of postthrombotic syndrome on health-related quality of life after deep venous thrombosis. *Arch Intern Med* 2002;162(10):1144–1148
- Pengo V, Lensing AW, Prins MH, et al; Thromboembolic Pulmonary Hypertension Study Group. Incidence of chronic thromboembolic pulmonary hypertension after pulmonary embolism. *N Engl J Med* 2004;350(22):2257–2264
- Anderson FA Jr, Zayaruzny M, Heit JA, Fidan D, Cohen AT. Estimated annual numbers of US acute-care hospital patients at risk for venous thromboembolism. *Am J Hematol* 2007;82(09):777–782
- Cohen AT, Tapson VF, Bergmann JF, et al; ENDORSE Investigators. Venous thromboembolism risk and prophylaxis in the acute hospital care setting (ENDORSE study): a multinational cross-sectional study. *Lancet* 2008;371(9610):387–394
- Gould MK, Garcia DA, Wren SM, et al. Prevention of VTE in nonorthopedic surgical patients: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* 2012;141(2, Suppl):e227S–e277S
- Qaseem A, Chou R, Humphrey LL, Starkey M, Shekelle P; Clinical Guidelines Committee of the American College of Physicians. Venous thromboembolism prophylaxis in hospitalized patients: a clinical practice guideline from the American College of Physicians. *Ann Intern Med* 2011;155(09):625–632
- Barbar S, Noventa F, Rossetto V, et al. A risk assessment model for the identification of hospitalized medical patients at risk for venous thromboembolism: the Padua Prediction Score. *J Thromb Haemost* 2010;8(11):2450–2457
- Marcucci M, Iorio A, Nobili A, et al; REPOSI (REgistro POLiterapie Società Italiana di Medicina Interna) Investigators. Prophylaxis of venous thromboembolism in elderly patients with multimorbidity. *Intern Emerg Med* 2013;8(06):509–520
- Depietri L, Marietta M, Scarlini S, et al. Clinical impact of application of risk assessment models (Padua Prediction Score and Improve Bleeding Score) on venous thromboembolism, major hemorrhage and health expenditure associated with pharmacologic VTE prophylaxis: a “real life” prospective and retrospective observational study on patients hospitalized in a Single Internal Medicine Unit (the STIME study). *Intern Emerg Med* 2018;13(04):527–534
- Pavon JM, Sloane RJ, Pieper CF, et al. Poor adherence to risk stratification guidelines results in overuse of venous thromboembolism prophylaxis in hospitalized older adults. *J Hosp Med* 2018;13(06):403–404
- Mamykina L, Vawdrey DK, Stetson PD, Zheng K, Hripcsak G. Clinical documentation: composition or synthesis? *J Am Med Inform Assoc* 2012;19(06):1025–1031
- Hripcsak G, Vawdrey DK, Fred MR, Bostwick SB. Use of electronic clinical documentation: time spent and team interactions. *J Am Med Inform Assoc* 2011;18(02):112–117
- Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T. Identifying individuals with physician diagnosed COPD in health administrative databases. *COPD* 2009;6(05):388–394
- Alotaibi GS, Wu C, Senthilselvan A, McMurtry MS. The validity of ICD codes coupled with imaging procedure codes for identifying acute venous thromboembolism using administrative data. *Vasc Med* 2015;20(04):364–368
- Jones BE, Jones J, Bewick T, et al. CURB-65 pneumonia severity assessment adapted for electronic decision support. *Chest* 2011;140(01):156–163
- Vinson DR, Morley JE, Huang J, et al; Kaisers Permanente CREST Network. The accuracy of an electronic pulmonary embolism severity index auto-populated from the electronic health record: setting the stage for computerized clinical decision support. *Appl Clin Inform* 2015;6(02):318–333
- Benesch C, Witter DM Jr, Wilder AL, Duncan PW, Samsa GP, Matchar DB. Inaccuracy of the International Classification of Diseases (ICD-9-CM) in identifying the diagnosis of ischemic cerebrovascular disease. *Neurology* 1997;49(03):660–664
- Quan H, Li B, Saunders LD, et al; IMECCHI Investigators. Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database. *Health Serv Res* 2008;43(04):1424–1441
- Elias P, Khanna R, Dudley A, et al. Automating venous thromboembolism risk calculation using electronic health record data upon hospital admission: the automated Padua Prediction Score. *J Hosp Med* 2017;12(04):231–237



- 21 Navar-Boggan AM, Rymer JA, Piccini JP, et al. Accuracy and validation of an automated electronic algorithm to identify patients with atrial fibrillation at risk for stroke. *Am Heart J* 2015;169(01):39–44
- 22 McCormick N, Lacaille D, Bhole V, Avina-Zubieta JA. Validity of heart failure diagnoses in administrative databases: a systematic review and meta-analysis. *PLoS One* 2014;9(08):e104519
- 23 Weiskopf NG, Hripcsak G, Swaminathan S, Weng C. Defining and measuring completeness of electronic health records for secondary use. *J Biomed Inform* 2013;46(05):830–836
- 24 Barber C, Lacaille D, Fortin PR. Systematic review of validation studies of the use of administrative data to identify serious infections. *Arthritis Care Res (Hoboken)* 2013;65(08):1343–1357
- 25 Wang LE, Shaw PA, Mathelier HM, Kimmel SE, French B. Evaluating risk-prediction models using data from electronic health records. *Ann Appl Stat* 2016;10(01):286–304
- 26 Greene MT, Spyropoulos AC, Chopra V, et al. Validation of risk assessment models of venous thromboembolism in hospitalized medical patients. *Am J Med* 2016;129(09):1001.e9–1001.e18
- 27 Deleger L, Brodzinski H, Zhai H, et al. Developing and evaluating an automated appendicitis risk stratification algorithm for pediatric patients in the emergency department. *J Am Med Inform Assoc* 2013;20(e2):e212–e220
- 28 Ford E, Carroll JA, Smith HE, Scott D, Cassell JA. Extracting information from the text of electronic medical records to improve case detection: a systematic review. *J Am Med Inform Assoc* 2016;23(05):1007–1015
- 29 Bates DW, Kuperman GJ, Wang S, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *J Am Med Inform Assoc* 2003;10(06):523–530