Received: 13 January 2023 | Revised: 2 April 2023 | Accepted: 7 April 2023

https://doi.org/10.1016/j.rpth.2023.100162

ORIGINAL ARTICLE

Development of a computable phenotype using electronic health records for venous thromboembolism in medical inpatients: the Medical Inpatient Thrombosis and Hemostasis study

Ryan M. Thomas MD[1,2] | Katherine Wilkinson MS[3] | Insu Koh PhD[3] | Ang Li MD[4] | Janine S. A. Warren PhD[1] | Nicholas S. Roetker PhD[5] | Nicholas L. Smith PhD[6,7,8] | Chris E. Holmes MD, PhD[1,2] | Timothy B. Plante MD, MHS[1,2] | Allen B. Repp MD, MS[1,2] | Mary Cushman MD, MSc[1,2,3] | Neil A. Zakai MD, MSc[1,2,3]

[1]Department of Medicine, Larner College of Medicine at the University of Vermont, Burlington, Vermont, USA

[2]University of Vermont Medical Center, Burlington, Vermont, USA

[3]Department of Pathology & Laboratory Medicine, Larner College of Medicine at the University of Vermont, Burlington, Vermont, USA

[4]Department of Medicine, Baylor University Medical Center, Houston, Texas, USA

[5]Chronic Disease Research Group, Hennepin Healthcare Research Institute, Minneapolis, Minnesota, USA

[6]Department of Epidemiology, University of Washington, Seattle, Washington, USA

[7]Kaiser Permanente Washington Health Research Institute, Kaiser Permanente Washington, Seattle Washington, USA

[8]Seattle Epidemiologic Research and Information Center, Department of Veterans Affairs Office of Research and Development, Seattle, Washington, USA

Correspondence
Neil A. Zakai, Colchester Research Facility, Larner College of Medicine, University of Vermont, 360 South Park Drive, Colchester, VT 05446, USA.
Email: neil.zakai@med.uvm.edu

Handling Editor: Dr Lana Antoinette Castellucci

Abstract

Background: Accurate and efficient methods to identify venous thromboembolism (VTE) events in hospitalized people are needed to support large-scale studies. Validated computable phenotypes using a specific combination of discrete, searchable elements in electronic health records to identify VTE and distinguish between hospital-acquired (HA)–VTE and present-on-admission (POA)–VTE would greatly facilitate the study of VTE, obviating the need for chart review.

Objectives: To develop and validate computable phenotypes for POA- and HA-VTE in adults hospitalized for medical reasons.

Methods: The population included admissions to medical services from 2010 to 2019 at an academic medical center. POA-VTE was defined as VTE diagnosed within 24 hours of admission, and HA-VTE as VTE identified more than 24 hours after admission. Using discharge diagnosis codes, present-on-admission flags, imaging procedures, and medication administration records, we iteratively developed computable phenotypes for POA-VTE and HA-VTE. We assessed the performance of the phenotypes using manual chart review and survey methodology.

Results: Among 62,468 admissions, 2693 had any VTE diagnosis code. Using survey methodology, 230 records were reviewed to validate the computable phenotypes. Based on the computable phenotypes, the incidence of POA-VTE was 29.4 per 1000 admissions and that of HA-VTE was 3.6 per 1000 admissions. The POA-VTE computable phenotype had positive predictive value and sensitivity of 88.8% (95% CI, 79.8%-94.0%) and 99.1% (95% CI, 94.0%- 99.8%), respectively. Corresponding values for the HA-VTE computable phenotype were 84.2% (95% CI, 60.8%-94.8%) and 72.3% (95% CI, 40.9%-90.8%).

Conclusion: We developed computable phenotypes for HA-VTE and POA-VTE with adequate positive predictive value and sensitivity. This phenotype can be used in electronic health record data–based research.

KEYWORDS
electronic health records, inpatients, International Classification of Diseases, predictive value of tests, venous thromboembolism

Essentials

- Computable phenotypes use discrete electronic health record data to identify clinical events.
- Validated phenotypes for hospital-acquired and present-on-admission venous thromboembolism (VTE) are needed.
- We developed phenotypes for hospital-acquired and present-on-admission VTE.
- These phenotypes can be used for future studies of VTE in medical patients.

1 | INTRODUCTION

Venous thromboembolism (VTE), including deep venous thrombosis and pulmonary embolism (PE), occurs in 1 to 2 of every 1000 US adults annually [1]. Approximately 33% of VTE events are associated with hospitalization [1]. There is no required reporting of VTE events in the United States or other nations, and thus, the true incidence of VTE is unknown [2]. While providing an overview of incidence, administrative claims data lack specificity for VTE events and are unable to differentiate VTE that is hospital-acquired (HA; occurring when someone is hospitalized and not present at admission) from VTE that is present on admission (POA) [3]. HA-VTEs are usually provoked events that have different treatment approaches and recurrence risks after discontinuation of anticoagulation [4]. Further, HA-VTEs are potentially preventable, and their occurrence leads to additional treatments, increased length of stay, and higher mortality [5,6].

Large electronic health record (EHR) clinical databases have the potential to inform the practice of medicine. In the United States, trained medical coders synthesize lengthy hospital records into accessible diagnosis and procedure codes, with many diagnosis codes subject to POA reporting [7]. Diagnosis codes combined with EHR clinical data, which include flowsheets, laboratory tests, imaging studies, and medication reconciliation, can be harnessed to create specific reproducible definitions of clinical events, called "computable phenotypes." Once computable phenotypes for HA-VTE and POA-VTE are developed, the incidence of, risk factors for, and consequences of VTE can be more precisely assessed [8–14].

Here we describe the iterative development and validation of computable phenotypes for POA-VTE and HA-VTE. We hypothesized that (1) it would be possible to construct a computable phenotype for POA-VTE and HA-VTE with data available in an EHR and (2) these computable phenotypes would have high positive and negative predictive values (PPV and NPV, respectively).

2 | METHODS

2.1 | Population

All people hospitalized at the University of Vermont (UVM) Medical Center, a 540-bed tertiary acute care hospital in Burlington, Vermont, United States, between 2010 and 2019 were assessed for eligibility. All hospital admissions to the medical services (general and family medicine, cardiology, medical intensive care unit, and hematology/oncology) were included if the patient spent at least 1 midnight in the hospital (including observation admissions) and was ≥18 years old. Hospital day 1 was considered the first calendar day the person was in the hospital. The UVM Medical Center is the primary (and only) hospital for Chittenden County, Vermont (~250,000 individuals) and is a tertiary care hospital for Northwestern Vermont and Northeastern New York State (~1,000,000 individuals). The UVM Medical Center uses an Epic Systems Corporation EHR for inpatient care. Baseline characteristics of the population were assessed using EHR data known at the time of admission (Table 1). Race was assessed from that reported in the EHR, with race/ethnicity recorded in a single field.

2.2 | Computable phenotype definitions

VTE was defined as thrombus of the deep veins (either proximal or distal) of the upper or lower extremities or thrombus in the pulmonary arteries. POA-VTE was defined as a VTE identified during the first 24 hours of admission, regardless of the primary reason for hospitalization, and HA-VTE was defined as a VTE identified after the first 24 hours of an admission. EHR data used to develop the computable phenotypes included International Classification of Diseases, Ninth or Tenth Revisions, Clinical Modification (ICD-9/10-CM) discharge codes with the POA flag, Current Procedure Terminology (CPT) codes (including the date of the medical service or procedure), as well as medication administration records (eg, anticoagulant use) [10].

TABLE 1 Baseline characteristics of hospital admissions.

Admission category	Entire cohort
No. of admissions (overall)	62,468
No. of admissions by service (n, %)	
General medicine	39,021 (62.5%)
Hematology/oncology	5,053 (8.1%)
Cardiology	14,083 (22.5%)
Pulmonary/critical care	4,311 (6.9%)
Discharge disposition	
Alive	60,302 (96.5%)
Deceased	2,166 (3.5%)
Status at admission (n, %)	
Inpatient	53,565 (85.7%)
Outpatient observation	8,903 (14.3%)
Demographics	
Age (mean, SD)	65.4 (17.5)
Female (n, %)	30,687 (49.1%)
Race/ethnicity (n, %)	
White	58,465 (93.6%)
Black	920 (1.5%)
Hispanic	19 (0.03%)
Asian	774 (1.2%)
Native American	396 (0.6%)
Native Hawaiian/Other Pacific Islander	28 (0.04%)
Other/Unknown	1,865 (3.0%)

Discharge diagnosis codes for VTE were selected based on prior literature as well as manual review of the ICD-9/10-CM codebooks [3]. CPT codes for imaging studies commonly used to diagnose or incidentally diagnose VTE were also required for confirmation of VTE (Supplementary Tables 1–3). Imaging studies had to correlate to the site of the diagnosis code (ie, chest imaging was required to diagnose a PE and lower-extremity imaging was required to identify lower-extremity deep venous thrombosis).

Medication administration data were used to determine when anticoagulation treatment was initiated. Anticoagulation was defined using the following drugs: enoxaparin, dalteparin, fondaparinux, heparin, warfarin, apixaban, edoxaban, rivaroxaban, dabigatran, argatroban, bivalirudin, and lepirudin. Supplementary Table 4 defines the anticoagulation dosing used to qualify as prophylactic or full anticoagulation. The definition excluded small doses of anticoagulants such as those used for line flushes. Inferior vena cava filter placement was defined using CPT codes (Supplementary Table 3).

Initially, 2 computable phenotypes, 1 for POA-VTE and 1 for HA-VTE, were empirically defined by the authors based on clinical knowledge and experience of VTE diagnosis and management. After initial assessment of performance via manual chart review, the phenotypes were refined based on patterns of failure. For POA-VTE, 140 charts were abstracted from 5 groups (defined by various criteria such as presence of a diagnosis code, POA flag, anticoagulation, and imaging), as illustrated in Figure 1 and Table 2. For the HA-VTE computable phenotype, 130 charts were abstracted from 6 groups, as illustrated in Figure 2 and Table 3. Twenty admissions without VTE diagnosis codes were reviewed to assess for NPV.

2.3 | Validation of phenotypes

The computable phenotypes were assessed with blinded physician chart abstraction (R.T. and N.A.Z.), our gold standard. The charts were abstracted by reviewing the initial history and physical progress notes, imaging completed during the hospitalization, and the discharge summary. The diagnosis of VTE was based on the clinical determination of the treating clinicians as documented in the clinical record, unless there was an obvious typographical or syntax error. Ambiguity in the medical record was discussed and resolved with consensus. During chart review, each admission was classified into 1 of the 3 following categories: no VTE present, POA-VTE, or HA-VTE. If a POA-VTE was present, any subsequent VTE during hospitalization was not considered a HA-VTE.

2.4 | Statistical analyses

The validity of each of the 2 computable phenotypes (POA-VTE and HA-VTE) was assessed by estimating the sensitivity, specificity, PPV, and NPV, using chart abstraction as the reference standard. Since chart abstraction was performed on only a subset of the admissions for each computable phenotype, inverse probability weighting was used to estimate computable phenotype performance and certainty of the estimate [15]. Reviewed admissions were weighted based on the proportion of charts reviewed across every subcategory defined in the computable phenotype. There were a total of 5 subcategories for the POA-VTE phenotype (Table 2) and 6 subcategories for the HA-VTE phenotype (Table 3).

3 | RESULTS

3.1 | Population characteristics

Table 1 presents the characteristics of the hospitalized population. From 2010 to 2019, there were a total of 62,468 medical admissions; 39,021 (62.5%) were admitted to general medicine services, 5053 (8.1%) to hematology and/or oncology services, 14,083 (22.5%) to cardiology services, and 4311 (6.9%) to pulmonary/critical care services. Of the admissions, 8903 (14.3%) were

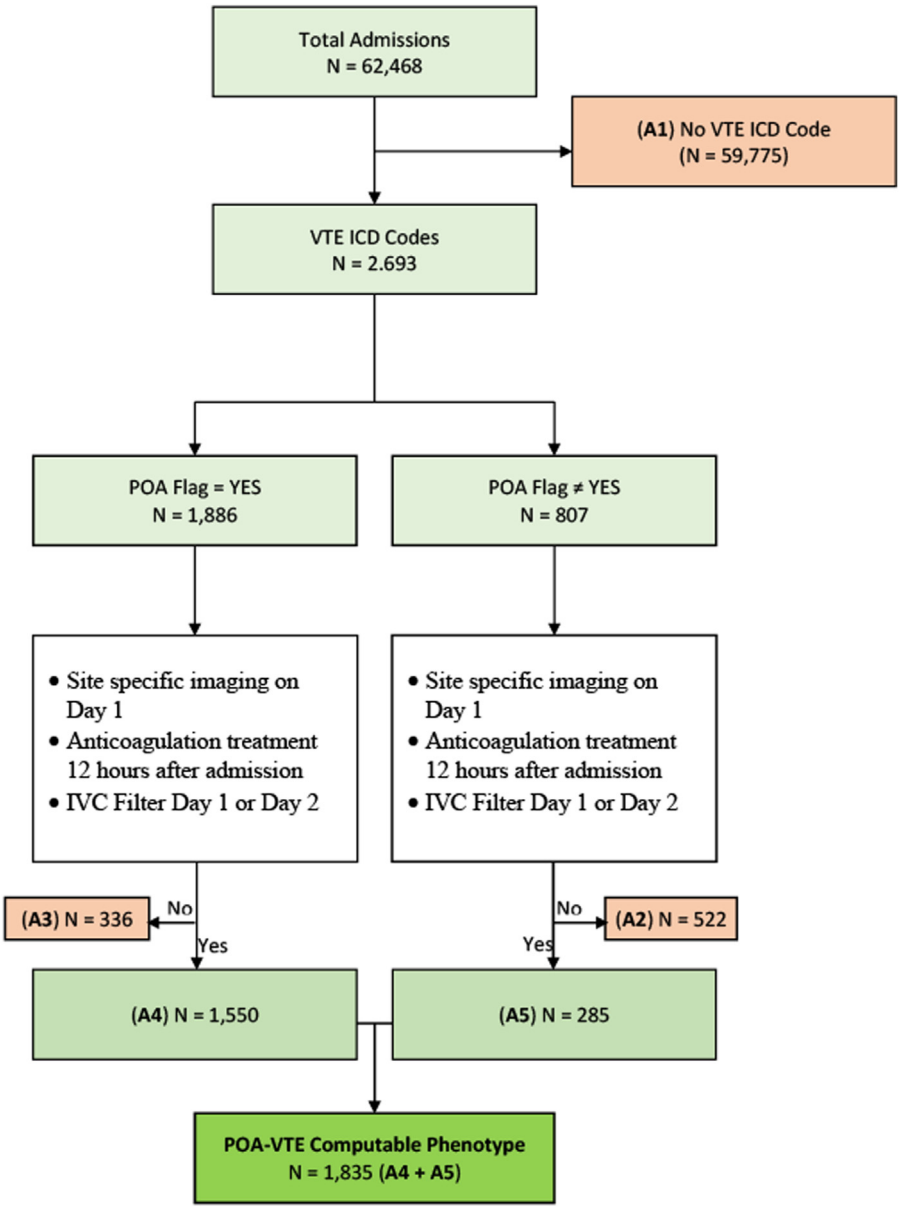

FIGURE 1 POA-VT computable phenotype flow chart. Table 2 columns correspond to numbers. ICD, International Classification of Diseases; IVC, inferior vena cava; POA, present on admission; VTE, venous thromboembolism.

observation status at admission (though could have been upgraded later). Mean (SD) age at admission was 65.4 ± 17.5 years, 41.9% of admitted patients were women, and 93.6% were identified as White persons.

3.2 | Iterative development of the VTE phenotypes

Figures 1 and 2 and Tables 2 and 3 present the results of the iterative development of the POA-VTE and the HA-VTE phenotypes.

3.2.1 | POA-VTE

For POA-VTE, the computable phenotype definition consisted of patients with an ICD-9/10-CM discharge diagnosis of VTE, the POA flag equaling "yes," and a CPT code for a corresponding site-specific imaging study on day 1 of admission. To account for the scenario where the diagnosis of VTE was established by imaging at an outside hospital prior to admission, POA-VTE also included those with a POA flag equaling "yes" and full dose anticoagulation initiated within the first 12 hours of admission or a CPT code for an inferior vena cava filter on day 1 of admission. Finally, hospital encounters with the VTE diagnosis POA flag not equaling "yes" but with a CPT code for a corresponding VTE site-specific imaging study within the first day of admission were also defined as having POA-VTE.

3.2.2 | HA-VTE

For HA-VTE, the computable phenotype definition consisted of people with a discharge diagnosis of VTE with the POA flag not equaling "yes"

TABLE 2 POA-VTE computable phenotype validation.

Population characteristic	A1 No VTE ICD code	A2 POA flag ≠ yes No imaging No treatment	A3 POA flag = yes No imaging No treatment	A4 POA flag = yes Imaging or treatment	A5 POA flag ≠ yes Imaging
				POA-VTE computable phenotype	
Total population	59,775	522	336	1550	285
No. of charts reviewed	20	9	22	58	31
Proportion of admissions reviewed	0.0003	0.0172	0.0655	0.0374	0.1088
Chart review					
No VTE	20	6	21	5	4
HA-VTE	0	3	0	1	1
POA-VTE	0	0	1	52	26

The POA-VTE computable phenotype was validated with chart review. The last 2 columns combined represent the POA-VTE computable phenotype. The total population number and total number of charts reviewed for each category are listed along with the distribution of charts in the no VTE, POA-VTE, or HA-VTE following chart review. See Figure 1 for definitions of each column (numbers A1-A5). HA, hospital acquired; ICD, International Classification of Diseases; POA, present on admission; VTE, venous thromboembolism.

and a CPT code for a site-specific imaging study completed after hospital day 1. If there was a CPT imaging code that could diagnose VTE on day 1 plus on a subsequent day, the case was excluded from the HA-VTE computable phenotype. There were only 11 patients falling into this category, and on chart review, 7 of them had a POA-VTE event, as shown in Table 3. A final rule was that if the imaging study on day 1 was a chest computed tomography (CT) scan with contrast, it would not exclude a further diagnosis of an HA-VTE if a subsequent chest CT with contrast or ventilation perfusion imaging scan was performed after day 1. CT scans were fairly common on hospital day 1 and generally provide an accurate indication if a PE is present at admission [16].

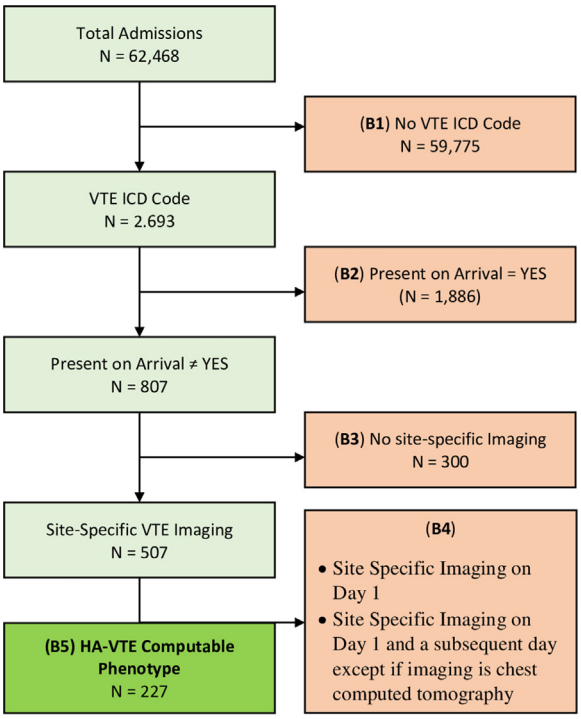

FIGURE 2 HA-VTE computable phenotype flow chart. Table 3 columns correspond to numbers. HA, hospital acquired; ICD, International Classification of Diseases; VTE, venous thromboembolism.

Tables 2 and 3 present the manual chart review validation for each computable phenotype (POA-VTE or HA-VTE). Of 230 charts, 5 charts with HA-VTE on chart review were misclassified as POA-VTE using the computable phenotype. A majority of those had multiple CPT codes representing imaging studies that could diagnose a VTE on hospital day 1; 1 VTE was incidentally discovered on a nonvascular ultrasound and, therefore, was not captured by the CPT codes in the computable phenotype. There were additional examples of miscoding; eg, an HA-VTE event having an erroneous POA flag and superficial thrombophlebitis being misclassified as a VTE event due to errors in clinical documentation and discharge coding. One case misclassified as an HA-VTE by the computable phenotype involved a high clinical suspicion followed by multiple imaging studies, but with an ultimate clinical decision that there was no HA-VTE.

3.3 | Sensitivity and specificity of the computable phenotypes

Table 4 presents the sensitivity, specificity, NPV, and PPV of the computable phenotypes. For the POA-VTE phenotype, the sensitivity was 99.1% (95% CI, 94.0%-99.9%) and the specificity was 99.7% (95% CI, 99.4%-99.8%). The NPV of the POA-VTE phenotype was 99.9% (95% CI, 99.8%-99.9%) and the PPV was 88.8% (95% CI, 79.8%-94.0%). For the HA-VTE phenotype, the sensitivity was 72.3% (95% CI, 40.9%-90.8%) and the specificity was 99.9% (95% CI, 99.8%-99.9%). The NPV of the HA-VTE phenotype was 99.8% (95% CI,

TABLE 3 HA-VTE computable phenotype validation.

Population characteristic	B1 No VTE ICD code	B2 POA flag = yes	B3 POA flag ≠ yes No imaging	B4 POA flag ≠ yes Imaging day 1 only	B4 POA flag ≠ yes Imaging day 1 + another day	B5 POA flag ≠ yes Imaging only after day 1 unless chest scan
						HA-VTE computable phenotype
Total population	59,775	1,886	300	240	40	227
No. of charts reviewed	20	40	20	20	11	19
Proportion of charts reviewed	0.0003	0.0212	0.0667	0.0833	0.275	0.0837
Chart review						
No VTE	20	8	14	6	1	2
POA-VTE	0	31	5	14	7	1
HA-VTE	0	1	1	0	3	16

The HA-VTE computable phenotype was validated with chart review. The total population number and total number of charts reviewed for each category are listed along with the distribution of charts in the no VTE, POA-VTE, or HA-VTE following chart review. See Figure 2 for the definition of each column (numbers A1-A5). HA, hospital acquired; ICD, International Classification of Diseases; POA, present on admission; VTE, venous thromboembolism.

99.6%-99.9%) and the PPV was 84.2% (95% CI, 60.8%-94.8%). The lower sensitivity was due to 1 failure out of 20 charts reviewed, reflecting an admission with an incorrect POA flag (Table 3). This failure triggered a review of 20 additional charts that revealed no further misclassifications.

3.4 | Event estimates of the computable phenotypes

Based on our computable phenotypes, we estimated that there were 1835 (29.3 per 1000 admissions) POA-VTE events and 227 (3.6 per 1000 admissions) HA-VTE events among the 62,468 admissions in this study. Using the observed sensitivity and specificity of our definitions, we estimated that the actual incidence of POA-VTE events was 1644 events (26.3 per 1000 admissions) and that for HA-VTE was 264 (4.2 per 1000 admissions) over the same timeframe.

4 | DISCUSSION

We developed and validated an EHR-based computable phenotype for HA-VTE and POA-VTE using ICD discharge codes with POA flags, imaging studies, and hospital-administered medications at an academic medical center in the United States. Unlike prior efforts [3], our methods harness the greater detail available in EHR data of hospitalized adults and allow better granularity in defining HA vs POA-VTE.

Validated computable phenotypes for VTE will allow clinicians and researchers to identify VTE events in EHR data in a standardized and reproducible manner. Extending prior research, we were able to develop separate computable phenotypes for HA-VTE and POA-VTE. Most previous algorithms using administrative data focused on primary discharge codes for inpatient events [3,17,18]. Further, in administrative data, dates of procedures (recorded as ICD procedure codes) are noted, but their exact timing and number of procedures can be problematic, limiting the ability to differentiate POA- and HA-VTE [3]. It is important to distinguish between POA- and HA-VTE because HA-VTE is by definition a provoked event and may only warrant a limited course of anticoagulation, whereas POA-VTE may be provoked or unprovoked and may warrant longer-term anticoagulation [4]. Prior efforts to define HA-VTE vs POA-VTE have relied on the position of the diagnosis codes in hospital discharges [18]. However, this approach will not distinguish the relatively common clinical event of an acute VTE co-occurring with other clinical events (ie, bronchopneumonia in someone undergoing treatment for cancer and an incidentally discovered VTE on an admission CT scan) vs occurring during hospitalization [18].

The VTE computable phenotypes rely on the fact that hospital billing in the United States employs trained chart coders who review medical records to report diagnosis codes and POA status. Unfortunately, occasional coding errors and inaccurate or ambiguous clinical documentation are major reasons why the phenotypes misclassified VTE events. From a research standpoint, as new treatment approaches evolve, imaging modalities integrated to clinical practice, and revisions to ICD coding are developed, the computable phenotypes will need to be updated [19]. These computable phenotypes may not readily translate to non-US health systems due to the phenotypes' reliance on manual coding of hospital records at discharge with POA reporting.

Our approach has strengths and weaknesses. An important strength is incorporating clinical practices and readily identifiable clinical phenotypes into our definitions. Our clinical expertise was used to develop computable phenotypes that directly reflect how VTE is diagnosed and treated in hospitalized settings. This helped improve the definition of POA- and HA-VTE and overcame a known issue with

TABLE 4 Weighted sensitivity and specificity of POA-VTE and HA-VTE computable phenotypes.

	Chart review POA-VTE (gold standard)		Chart review HA-VTE (gold standard)	
EHR computable phenotype	**Yes**	**No**	**Yes**	**No**
Yes	1629	206	191	36
No	15	60,249	73	62,168
Sensitivity (95% CI)	99.1% (94.0%-99.9%)		72.3% (40.9%-90.8%)	
Specificity (95% CI)	99.7% (99.4%-99.8%)		99.9% (99.8%-99.9%)	
Positive predictive value (95% CI)	88.8% (79.8%-94.0%)		84.2% (60.8%-94.8%)	
Negative predictive value (95% CI)	99.9% (99.8%-99.9%)		99.9% (99.6%-100.0%)	

Values in each confusion matrix were generated using inverse probability weighting on the basis of the number of charts reviewed for each computable phenotype category shown in Tables 2 and 3 and rounded to the nearest whole number. EHR, electronic health record; HA, hospital acquired; ICD, International Classification of Diseases; POA, present on admission; VTE, venous thromboembolism.

reduced specificity of secondary diagnosis codes for VTE [18]. A limitation of this study is the use of data from a single hospital and a single EHR; results could differ in other hospitals or with other EHRs. There is variability in the accuracy of discharge coding in the United States, so the computable phenotypes may not show the same performance when applied in other locations or hospitals. The information required for these phenotypes may not be available in some EHRs, and the medication administration formats are not standardized. Nevertheless, because coding forms a basis for remuneration and quality metrics, there is a strong incentive for payers and hospitals to ensure accurate coding, and documenting medication administration is a key element of patient care in hospitalization [20]. While we only reviewed 20 charts without VTE diagnosis codes, prior data from the Centers for Medicare and Medicaid Services as well as from prior work at UVM Medical Center confirm the high sensitivity of diagnosis codes for identifying HA-VTE [21,22].

To conclude, we developed and validated computable phenotypes for HA-VTE and POA-VTE using discharge diagnosis codes, CPT codes, and medication administration data found within EHR data. These phenotypes solely refer to the timing of a VTE event during a hospitalization; a POA-VTE could be a provoked event triggered by a hospitalization. These phenotypes, which require external validation at different clinical sites, may facilitate further research on VTE in hospitalized patients.

ACKNOWLEDGMENTS

We appreciate and acknowledge the support and expertise of Mr Michael Gianni, Senior Measurement Analyst, and the Data Management Office at the University of Vermont Health Network.

FUNDING

This study was funded by grant R01-HL141290 (N.A.Z.) from the National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland, USA. The sponsoring institution had no role in the analytical approach or review of data. The views presented here represent those of the authors and not necessarily of the National Institutes of Health.

ETHICS STATEMENT

The research reported conformed to the tenants of the Declaration of Helsinki and was determined to be exempt by the University of Vermont institutional review board based on exemption category 4 under the 2018 common rule [23].

AUTHOR CONTRIBUTIONS

R.M.T., K.W., I.K., N.S.R., A.B.R., and N.A.Z. contributed to the conception/design of the study with input from all the other authors. R.M.T., K.W., I.K., and N.A.Z. contributed to data acquisition. N.A.Z. secured funding. K.W., I. K., N.S.R. designed and performed the statistical analyses. R.M.T. drafted the manuscript and all authors were involved in manuscript writing and critically reviewing the manuscript and approving the final version.

RELATIONSHIP DISCLOSURE

There are no competing interests to disclose.

REFERENCES

[1] Beckman MG, Hooper WC, Critchley SE, Ortel TL. Venous thromboembolism: a public health concern. *Am J Prev Med*. 2010;38:S495–501.

[2] Virani SS, Alonso A, Aparicio HJ, Benjamin EJ, Bittencourt MS, Callaway CW, et al. Heart disease and stroke statistics-2021 update: a report from the American Heart Association. *Circulation*. 2021;143:e254–743.

[3] Sanfilippo KM, Wang TF, Gage BF, Liu W, Carson KR. Improving accuracy of International Classification of Diseases codes for venous thromboembolism in administrative data. *Thromb Res*. 2015;135:616–20.

[4] Lutsey PL, Zakai NA. Epidemiology and prevention of venous thromboembolism. *Nat Rev Cardiol*. 2023;20:248–62.

[5] Darzi AJ, Karam SG, Charide R, Etxeandia-Ikobaltzeta I, Cushman M, Gould MK, et al. Prognostic factors for VTE and bleeding in

hospitalized medical patients: a systematic review and meta-analysis. *Blood*. 2020;135:1788–810.
[6] Zhang Z, Lei J, Shao X, Dong F, Wang J, Wang D, et al. Trends in hospitalization and in-hospital mortality from VTE, 2007 to 2016, in China. *Chest*. 2019;155:342–53.
[7] Alexander S, Conner T, Slaughter T. Overview of inpatient coding. *Am J Health Syst Pharm*. 2003;60:S11–4.
[8] Richesson RL, Hammond WE, Nahm M, Wixted D, Simon GE, Robinson JG, et al. Electronic health records based phenotyping in next-generation clinical trials: a perspective from the NIH Health Care Systems Collaboratory. *J Am Med Inform Assoc*. 2013;20:e226–31.
[9] Tasker RC. Why everyone should care about "computable phenotypes. *Pediatr Crit Care Med*. 2017;18:489–90.
[10] *CPT Professional 2020*. American Medical Association; 2019.
[11] Kashyap R, Sarvottam K, Wilson GA, Jentzer JC, Seisa MO, Kashani KB. Derivation and validation of a computable phenotype for acute decompensated heart failure in hospitalized patients. *BMC Med Inform Decis Mak*. 2020;20:85.
[12] Ahmad FS, Ricket IM, Hammill BG, Eskenazi L, Robertson HR, Curtis LH, et al. Computable phenotype implementation for a national, multicenter pragmatic clinical trial: lessons learned from ADAPTABLE. *Circ Cardiovasc Qual Outcomes*. 2020;13:e006292.
[13] Richesson RL, Smerek MM, Blake Cameron C. A framework to support the sharing and reuse of computable phenotype definitions across health care delivery and clinical research applications. *EGEMS (Wash DC)*. 2016;4:1232.
[14] 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;20:364–8.
[15] Katki HA, Li Y, Edelstein DW, Castle PE. Estimating the agreement and diagnostic accuracy of two diagnostic tests when one test is conducted on only a subsample of specimens. *Stat Med*. 2012;31:436–48.
[16] Abdellatif W, Ebada MA, Alkanj S, Negida A, Murray N, Khosa F, et al. Diagnostic accuracy of dual-energy CT in detection of acute pulmonary embolism: a systematic review and meta-analysis. *Can Assoc Radiol J*. 2021;72:285–92.
[17] Fang MC, Fan D, Sung SH, Witt DM, Schmelzer JR, Steinhubl SR, et al. Validity of using inpatient and outpatient administrative codes to identify acute venous thromboembolism: the CVRN VTE study. *Med Care*. 2017;55:e137–43.
[18] White RH, Garcia M, Sadeghi B, Tancredi DJ, Zrelak P, Cuny J, et al. Evaluation of the predictive value of ICD-9-CM coded administrative data for venous thromboembolism in the United States. *Thromb Res*. 2010;126:61–7.
[19] Lyman GH, Carrier M, Ay C, Di Nisio M, Hicks LK, Khorana AA, et al. American Society of Hematology 2021 guidelines for management of venous thromboembolism: prevention and treatment in patients with cancer. *Blood Adv*. 2021;5:927–74.
[20] Burns EM, Rigby E, Mamidanna R, Bottle A, Aylin P, Ziprin P, et al. Systematic review of discharge coding accuracy. *J Public Health (Oxf)*. 2012;34:138–48.
[21] Centers for Medicare & Medicaid Services. Accuracy of coding in the hospital-acquired conditions—present on admission program final report. https://www.cms.gov/medicare/medicare-fee-for-service-payment/hospitalacqcond; 2012 [accessed December 22, 2022].
[22] Zakai NA, Callas PW, Repp AB, Cushman M. Venous thrombosis risk assessment in medical inpatients: the medical inpatients and thrombosis (MITH) study. *J Thromb Haemost*. 2013;11:634–41.
[23] 83 FR 28497 - federal policy for the protection of human subjects: six month delay of the general compliance date of revisions while allowing the use of three burden-reducing provisions during the delay period. https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/revised-common-rule-regulatory-text/index.html; 2012 [accessed December 22, 2022].

SUPPLEMENTARY MATERIAL

The online version contains supplementary material available at https://doi.org/10.1016/j.rpth.2023.100162