HHS Public Access

Author manuscript

Thromb Res. Author manuscript; available in PMC 2016 April 01.

Published in final edited form as:

Thromb Res. 2015 April ; 135(4): 616–620. doi:10.1016/j.thromres.2015.01.012.

Improving Accuracy of *International Classification of Diseases* Codes for Venous Thromboembolism in Administrative Data

Kristen M. Sanfilippo, MD, MPHS[1,2], **Tzu-Fei Wang, MD**[3], **Brian F. Gage, MD, MSc**[2], **Weijian Liu, MPH**[1,4], and **Kenneth R. Carson, MD, PhD**[1,2]

[1]Saint Louis Veterans Health Administration Medical Center Research Service

[2]Washington University in Saint Louis

[3]The Ohio State University Comprehensive Cancer Center

[4]Department of Biostatistics, College for Public Health and Social Justice, Saint Louis University

Abstract

BACKGROUND—Increasingly, clinicians and researchers are using administrative data for clinical and outcomes research. However, they continue to question the accuracy of using *International Classification of Diseases 9th Revision* (ICD-9) codes alone to capture diagnoses, especially venous thromboembolism (VTE), in administrative data.

OBJECTIVES—We tested the hypothesis that incorporation of treatment data and/or common procedural terminology (CPT) codes could improve accuracy of administrative data in detecting VTE.

RESEARCH DESIGN—Using the Veterans Affairs Central Cancer Registry, we compared three competing algorithms by performing three cross-sectional studies. Algorithm 1 identified patients by ICD-9 codes alone. Algorithm 2 required VTE treatment in addition to ICD-9 codes. Algorithm 3 required a VTE diagnostic CPT code in addition to treatment and ICD-9 criteria.

RESULTS—The accuracy of ICD-9 codes alone for detection of VTE was marginal, with a PPV of 72%. The PPV was improved to 91% after addition of treatment data (algorithm 2). As compared to algorithm 2, addition of CPT codes (algorithm 3) did not significantly increase the accuracy of detecting VTE (PPV 92%), but decreased sensitivity from 72% to 67%.

Corresponding Author: Kristen M. Sanfilippo, MD, MPHS, Division of Hematology, Washington University School of Medicine, 660 S. Euclid Ave., Campus Box 8125, St. Louis, MO 63110, Phone: 314-362-0233, Fax: 314-362-8813, ksanfili@dom.wustl.edu.
Kristen M. Sanfilippo, MD, MPHS, Division of Hematology, Washington University School of Medicine, 660 S. Euclid Ave., Campus Box 8125, St. Louis, MO 63110, Phone: 314-362-0233, Fax: 314-362-8813, ksanfili@dom.wustl.edu
Tzu-Fei Wang, MD, Division of Hematology, The Ohio State University, 320 10th W. Ave, A340, Columbus OH 43210, Phone: 614-293-2887, Fax: 614-366-5970, Tzu-Fei.Wang@osumc.edu
Brian Gage, MD, MSc, Division of General Medical Sciences, Washington University School of Medicine, 660 S. Euclid Ave., Campus Box 8005, St. Louis, MO 63110, Phone: 314-454-8697, Fax: 314-454-5554, bgage@im.wustl.edu
Weijian Liu, MPH, Department of Biostatistics, Saint Louis University College for Public Health and Social Justice, Salus Center, 3545 Lafayette Ave., Cubicle 340, St. Louis, MO 63104, Phone: 314-977-8246, Fax: 314-977-8150, wliu9@slu.edu
Kenneth Carson, MD, PhD, Division of Oncology, Washington University School of Medicine, 660 S. Euclid Ave., Campus Box 8056, St. Louis, MO 63110, Phone: 314-362-5654, Fax: 314-454-7941, kcarson@dom.wustl.edu

CONCLUSIONS—Accuracy of VTE detection significantly improved with addition of treatment data to ICD-9 codes. This approach should facilitate use of administrative data to assess the incidence, epidemiology, and outcomes of VTE.

Keywords

Venous Thromboembolism; Administrative Data; Health Service Research

INTRODUCTION

Increasingly, clinicians and researchers are using administrative data for clinical and outcomes research. As compared to the assembly of prospective cohorts, research with administrative data (and electronic medical records) is cost effective and expedient[1]. Furthermore, the large volume of administrative data increases the generalizability of research findings and allows for assessment of quality of care across populations.

Incidence of venous thromboembolism (VTE) is one such quality measure identified by the Centers for Medicare & Medicaid Services (CMS). CMS identifies VTE by using *International Classification of Diseases 9th Revision* (ICD-9) codes, and uses these rates to impute hospital quality and calculate reimbursement. However, clinicians and researchers have questioned the accuracy of using ICD-9 codes alone to capture diagnoses, especially VTE[2]. A main reason for inaccuracy of ICD-9 codes is the use of an incorrect code (misdiagnosis).

The accuracy of ICD-9 codes might be improved by various means[3]. For example, one review assessed the positive predictive value (PPV) of VTE claim codes individually, and in combination[4]. The authors found that using a combination of ICD-9 codes (415, 451, 453) to identify VTE provided higher PPVs compared to using individual codes. A second study demonstrated improved accuracy by combining anticoagulant pharmacy data to VTE ICD-9 codes[5]. In that study, the PPV of a combination of ICD-9 codes (415.1 and 451–453) was 42%. After adding treatment data, the PPV increased to 65%. Thus, diagnostic algorithms might be improved by incorporating treatment data. In addition, using common procedural terminology (CPT) codes to assess for diagnostic studies used to detect VTE is another potential way to identify a VTE, and warrants investigation.

We tested the hypothesis that incorporation of treatment data with or without CPT codes could improve the accuracy of ICD-9 codes in detecting VTE in administrative data in a population of non-Hodgkin lymphoma (NHL) patients using the Veterans Health Administration (VHA) Central Cancer Registry administrative database. We linked the VHA Central Cancer Registry to the VHA EMR, allowing comparison of ICD-9 codes to the gold standard of manual chart abstraction. We focused our study on NHL patients, as patients with NHL have a 10-fold increased risk of VTE[6] and because these medical records had already been extensively reviewed as part of a prior research project by our group[7].

MATERIALS AND METHODS

Study Population

Patients diagnosed with diffuse large B-cell lymphoma between October 1, 1998 and December 31, 2008 or follicular lymphoma between October 1, 1998 and December 31, 2010 were identified in the VHA Central Cancer Registry by using ICD–O-3 codes consistent with the InterLymph classification system[8]. Patients with an ICD-9 code for atrial fibrillation (427.31) were excluded, given alternate indication for anticoagulation.

Study Design

We compared three competing algorithms for detection of VTE by performing three cross-sectional studies. Algorithm 1 identified patients by ICD-9 codes alone (Table 1). ICD-9 code for VTE was acceptable in any position from both inpatient and outpatient encounters. Algorithm 2 incorporated treatment criteria in addition to ICD-9 codes. Algorithm 3 required a VTE diagnostic CPT code in addition to treatment criteria and ICD-9 codes. ICD-9 codes used to identify VTE diagnoses and CPT codes used to identify diagnostic studies for VTE (Appendix A) were obtained from review of the ICD 9th revision 2011 and the CPT 2011 standard edition to account for codes available up to the end of study period, December 31, 2010. Treatment criteria included: prescription for outpatient anticoagulation (warfarin, enoxaparin, fondaparinux, or dalteparin), placement of an inferior vena cava (IVC) filter, or death within 30 days of VTE diagnosis. Selection of outpatient anticoagulation regimens for inclusion in the study was based on available approved anticoagulants for treatment of VTE up to 2010. Death within 30 days of VTE diagnosis was included to capture inpatients that died from their VTE before receiving an anticoagulant.

Medical Record Review

For each of the 3 algorithms, we randomly selected 150 medical records meeting criteria (algorithm positive) and 150 records that failed to meet criteria (algorithm negative) for review. Records were accessed using the Compensation and Pension Records Interchange software system. Admitting histories, radiology reports, and discharge summaries were reviewed for evidence of a VTE. Cases of VTE consisted of acute pulmonary embolism or acute deep vein thrombosis in any location (upper extremity, lower extremity, or intra-abdominally) confirmed by objective imaging: diagnostic studies to detect VTE included venous doppler ultrasound, venogram, computed tomography angiogram, pulmonary angiogram, computed tomography chest with contrast, or a ventilation/perfusion lung scan that was high probability. An elevated D-Dimer was not considered diagnostic of a VTE. A total of 900 medical records were reviewed by two hematologist-oncologists (KS and TW), who were blinded to the algorithm assignment.

Statistics

The accuracy of each algorithm was assessed following review of medical records, by calculating sensitivity, specificity, and PPV for each algorithm. To improve the precision of our PPV, we sampled an equal number of algorithm positive and algorithm negative patients (150 each) for each of the 3 algorithms. To prevent reporting inflated values of PPV from

this over-sampling, we corrected for it by calculating 95% confidence intervals of using the binomial theorem[9].

RESULTS

The study population contained 4,360 patients with a diagnosis of diffuse large B-cell lymphoma (n=3,075) or follicular lymphoma (n=1,285) between the specified dates. The average age was 65.8 years, with most Caucasian males (Table 2). The frequency of ICD-9 code distribution in the entire cohort is listed in appendix A.

Using algorithm (1), 705 VTEs and 3,655 controls were identified. Algorithm (2) identified 572 VTEs with 3,788 controls. Lastly, algorithm (3) identified 498 VTEs and 3,862 controls. Review of the 900 medical records contained no disagreements between the two reviewers and yielded the following findings. The PPV using ICD-9 codes alone (algorithm 1) was 72%. The PPV increased to 91% with addition of treatment data (algorithm 2). With the addition of CPT codes to ICD-9 and treatment data (algorithm 3), PPV was 92%, but sensitivity declined to 67% (table 3).

To better understand the reasoning for the variability in PPV among the three algorithms, we reviewed the false positives. Of the 150 patients with positive ICD-9 codes identified by algorithm 1, only 108 were true positives, yielding a PPV of 72%. The most common reasons for the 42 false positives were: random coding errors (n=13); clinical suspicion of a VTE with no positive diagnostic study (n=12); superficial thrombophlebitis (n=10); chronic deep vein thrombosis (n=5); VTE prophylaxis (n=1); and venous insufficiency (n=1). In algorithm 2 and 3, only 14 of 150 and 12 of 150 were false positive cases, respectively. In algorithm 2, the most common cause for a false positive was superficial thrombophlebitis (n=5); with remaining causes being secondary to: clinical suspicion of VTE with no diagnostic confirmation (n=3), left ventricular thrombus (n=1), and no identifiable reason (n=5). Lastly, in algorithm (3), the most common cause for a false positive was also superficial thrombophlebitis (n=4); additional reasons for false positives included clinical suspicion of VTE without diagnostic confirmation (n=3), left ventricular thrombus (n=1), and no identifiable reason (n=4).

DISCUSSION

The accuracy of ICD-9 codes alone for detection of VTE was marginal, with a PPV of 72%. The PPV was improved to 91% by the addition of treatment data (or death attributable to VTE) (algorithm 2). As compared to algorithm 2, the addition of CPT codes of VTE diagnostic studies (algorithm 3) did not significantly increase the accuracy of detecting VTE (PPV 92%), but decreased sensitivity from 72% to 67%.

Our results on the PPV of ICD-9 codes alone are comparable to prior literature. In national Medicare data, the accuracy of ICD-9 codes for VTE was 72% with a sensitivity of 61%[10]. White et al. found a 95% PPV for acute VTE when the ICD-9 code was used in the first position[3]; use of the ICD-9 code in the secondary position decreased PPV to 75%. Thus, while assessment for acute VTE using first position discharge codes has high PPV, researchers miss cases of acute VTE who had alternate discharge diagnoses listed first and

their acute VTE coded in a secondary position. Such may be the case in patients who are admitted to the hospital for surgical procedures. Henderson et al. assessed the accuracy of the Agency for Healthcare Research and Quality (AHRQ) patient safety indicator (PSI)-12 for post-operative VTE[11]. The AHRQ PSIs rely heavily on ICD-9 coding. The PPV of the PSI-12 for post-operative VTE was 54.5% with a sensitivity of 87% with the majority of false positive results superficial VTE (with improper coding). These findings were supported in a similar study by White et al. in which accuracy of AHRQ PSI-12 for post-operative VTE was found to be 79% despite a sensitivity of 95.5%[12]. Similar ICD-9 accuracies have been reported in pregnancy-associated VTE[2]. The suboptimal PPV of prior analyses and algorithm 1, argue for including treatment (as in algorithm 2 and 3) to improve accuracy and increase validity of using administrative data for assessment of patient safety and health outcomes research.

Although controversial, VTE detected by administrative data are being used to assess hospital quality[13]. The PSI-12 is used as a measure for pay-for-performance for post-operative VTE. The poor accuracy of this measure highlighted by Henderson et al. and White et al. and algorithm 1 support the need for alternates (e.g. algorithm 2 or 3) to detect VTE in administrative data and EMRs[11, 12]. The improved PPV using the approach in algorithms 2 and 3 could prevent hospitals from being financially penalized for detecting and coding trivial VTE (including superficial thrombophlebitis that are miscoded), which do not require treatment.

We found several reasons for the marginal PPV of ICD-9 codes (algorithm 1) for detecting VTE in our database. The most common false positive cases captured by ICD-9 codes alone included coding error, clinically suspected VTE with absence of confirmation on diagnostic testing, superficial thrombophlebitis, and chronic deep vein thrombosis. Inclusion of treatment data (or death attributable to VTE) into an algorithm for detection of VTE minimized false positive results and thereby improved PPV. This improvement may be because clinically suspected VTEs that are not confirmed, superficial thrombophlebitis, and chronic VTE do not receive treatment. Persistence of false positives after requirement for treatment in algorithm 2 and 3 can partially be explained by short-term treatment of superficial thrombophlebitis as well as brief therapy for clinically suspected cases until a diagnostic study could be performed to rule out thrombosis[14].

Our study has several limitations. Patients with diagnostic studies for VTE performed outside the VHA health system may have been missed in our algorithm. Additionally, the VHA EMR includes both inpatient and outpatient information. Outpatient treatment cannot be inferred in some datasets. Given that the VHA does not use ICD coding for billing purposes, there may be differences in the validity between our findings and those of Medicare based administrative data. Additionally, the cohort studied was exclusive for patients with lymphoma who have a higher prevalence of VTE than the general population. It is possible given this, that the predictive value of the algorithm would be less in populations with lower prevalence of disease. Results will need to be reassessed after introduction of the ICD-10 coding system, which would affect prospective cases of VTE after October 1, 2015[15]. Any algorithm generated for using ICD coding in administrative data will require constant modification and updating over time to account for changes and

additions to the ICD codes. Our cohort included events up until December 31, 2010. Updated versions of ICD codes become available for use October 1 of the preceding year. Thus the ICD-9 2011 version went into effect October 1, 2010. While available, the last VTE recorded in our cohort occurred September 16, 2009. Thus, any new codes present in the 2011 version, while available during the study period, did not apply as of the date of the last VTE. Review of subsequent ICD versions reveals that no changes have been made to ICD-9 codes for VTE, thus this algorithm generated is valid for claims up to October 1, 2015 in administrative data.

The VHA Central Cancer Registry database has several advantages for use in our study. It is a comprehensive database containing information on all cancer patients diagnosed and treated within any VHA medical center throughout the United States since January 1, 1995, and therefore allows for inclusion of a diverse population of patients. All data in the registry are linked to pharmacy data from both the inpatient and outpatient setting. The VHA Central Cancer Registry is linked to the VHA EMR. The linkage of our database to an EMR allowed for ascertainment of the diagnoses by review of medical records, the gold standard for diagnosis confirmation. Similarly, these reviewers captured VTE events that were not captured by the ICD-9 codes, allowing us to calculate the sensitivity of each algorithm. Third, the inclusion of IVC filter insertion as part of treatment data and the addition of CPT codes to our algorithm is a novel approach to improve the accuracy of administrative database research.

CONCLUSION

In conclusion, with increasing use of administrative databases for research and quality assessment[16], it is critical to capture VTE accurately. Here, we developed algorithms to identify VTE patients in administrative data that significantly improved PPV compared to use of ICD-9 codes alone. These algorithms will facilitate identification of VTEs using administrative databases.

Acknowledgments

FUNDING:

Funding that supported this study includes: 5K12HL087107 and U54 HL112303.

REFERENCES

1. Loke YK. Use of databases for clinical research. Archives of disease in childhood. 2014
2. White RH, Brickner LA, Scannell KA. ICD-9-CM codes poorly indentified venous thromboembolism during pregnancy. Journal of clinical epidemiology. 2004; 57:985–988. [PubMed: 15504642]
3. 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. Thrombosis research. 2010; 126:61–67. [PubMed: 20430419]
4. Tamariz L, Harkins T, Nair V. A systematic review of validated methods for identifying venous thromboembolism using administrative and claims data. Pharmacoepidemiology and drug safety. 2012; 21(Suppl 1):154–162. [PubMed: 22262602]

5. Gerstman BB, Freiman JP, Hine LK. Use of subsequent anticoagulants to increase the predictive value of Medicaid deep venous thromboembolism diagnoses. Epidemiology. 1990; 1:122–127. [PubMed: 2073498]
6. Blom JW, Doggen CJ, Osanto S, Rosendaal FR. Malignancies, prothrombotic mutations, and the risk of venous thrombosis. JAMA : the journal of the American Medical Association. 2005; 293:715–722.
7. Ganti A, Liu W, Luo S, Sanfilippo KM, Roop R, Lynch R, et al. Impact of body mass index on incidence of febrile neutropenia and treatment-related mortality in United States veterans with diffuse large B-cell lymphoma receiving rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone. Br J Haematol. 2014
8. Morton LM, Turner JJ, Cerhan JR, Linet MS, Treseler PA, Clarke CA, et al. Proposed classification of lymphoid neoplasms for epidemiologic research from the Pathology Working Group of the International Lymphoma Epidemiology Consortium (InterLymph). Blood. 2007; 110:695–708. [PubMed: 17389762]
9. Smith AK, Ayanian JZ, Covinsky KE, Landon BE, McCarthy EP, Wee CC, et al. Conducting high-value secondary dataset analysis: an introductory guide and resources. Journal of general internal medicine. 2011; 26:920–929. [PubMed: 21301985]
10. Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ, Gage BF. Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. Medical care. 2005; 43:480–485. [PubMed: 15838413]
11. Henderson KE, Recktenwald A, Reichley RM, Bailey TC, Waterman BM, Diekemper RL, et al. Clinical validation of the AHRQ postoperative venous thromboembolism patient safety indicator. Joint Commission journal on quality and patient safety / Joint Commission Resources. 2009; 35:370–376. [PubMed: 19634805]
12. White RH, Sadeghi B, Tancredi DJ, Zrelak P, Cuny J, Sama P, et al. How valid is the ICD-9-CM based AHRQ patient safety indicator for postoperative venous thromboembolism? Medical care. 2009; 47:1237–1243. [PubMed: 19786907]
13. Kinnier CV, Barnard C, Bilimoria KY. The need to revisit VTE quality measures. JAMA : the journal of the American Medical Association. 2014; 312:286–287.
14. Decousus H, Prandoni P, Mismetti P, Bauersachs RM, Boda Z, Brenner B, et al. Fondaparinux for the treatment of superficial-vein thrombosis in the legs. The New England journal of medicine. 2010; 363:1222–1232. [PubMed: 20860504]
15. Quan H, Eastwood C, Cunningham CT, Liu M, Flemons W, De Coster C, et al. Validity of AHRQ patient safety indicators derived from ICD-10 hospital discharge abstract data (chart review study). BMJ open. 2013; 3:e003716.
16. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. Annual review of public health. 2011; 32:91–108.

Appendix

ICD9 Codes Used for Algorithms as well as Frequency of Occurrence Within Data	Number of Patients	% of total
4151 Pulmonary Embolism/Infarction		
41511 Iatrogenic Pulmonary Embolism	64	0.9
41513 Saddle Pulmonary Embolism	0	0
41519 Other Pulmonary Embolism/Infarction	1373	19
451 Phlebitis/Thrombophlebitis		
4511 Phlebitis/Thrombophlebitis of deep veins of lower extremity		
45111 Phlebitis/Thrombophlebitis of the Femoral Vein	44	0.6
45119 Phlebitis/Thrombophlebitis of Other Vein	252	3.5

ICD9 Codes Used for Algorithms as well as Frequency of Occurrence Within Data	Number of Patients	% of total
4512 Phlebitis/Thrombophlebitis of Lower Extremity NOS	73	1
4518 Phlebitis/Thrombophlebitis of Other Sites		
45181 Phlebitis/Thrombophlebitis of Iliac Vein	2	0.03
45183 Phlebitis/Thrombophlebitis of Deep Veins of Upper Extremity	17	0.2
45184 Phlebitis/Thrombophlebitis of Upper Extremity NOS	19	0.3
45189 Phlebitis/Thrombophlebitis of Other Sites	8	0.1
4519 Phlebitis and Thrombophlebitis of Unspecified Sites	114	1.6
452 Portal Vein Thrombosis	14	0.2
453 Other Venous Embolism/Thrombosis		
4530 Budd Chiari Syndrome	4	0.1
4532 Embolism of Inferior Vena Cava	20	0.3
4533 Embolism and Thrombosis of Renal Vein	5	0.1
4534 Acute Venous Embolism/Thrombosis of Deep Vessels of Lower Extremity		
45340 Acute Venous Embolism/Thrombosis of Deep Vessels of Lower Extremity Unspecified	170	2.4
45341 Acute Venous Embolism/Thrombosis of Deep Vessels of Proximal Lower Extremity	86	1.2
45342 Acute Venous Embolism/Thrombosis of Deep Vessels of Distal Lower Extremity	48	0.7
4538 Acute Venous Embolism/Thrombosis of Other Specified Veins	4330	60
45382 Acute Venous Thromboembolism and Thrombosis of Deep Veins of Upper Extremity	0	0
45383 Acute Venous Thromboembolism and Thrombosis of Upper Extremity Unspecified	0	0
45384 Acute Venous Thromboembolism and Thrombosis of Axillary Vein	0	0
45385 Acute Venous Thromboembolism and Thrombosis of Subclavian Vein	0	0
45386 Acute Venous Thromboembolism and Thrombosis of Internal Jugular Vein	0	0
45387 Acute Venous Thromboembolism and Thrombosis of Other Thoracic Veins	0	0
45389 Acute Venous Thromboembolism and Thrombosis of Other Specified Veins	0	0
4539 Embolism and Thrombosis of Unspecified Site	578	8

CPT Codes Used for VTE Diagnostic Studies
Venous Doppler/Duplex Ultrasound:
93965, 93970, 93971
Venography:
75820, 75822, 75825, 75827, 75860
Nuclear Medicine Thrombosis Imaging:
78456, 78457, 78458
Ultrasound Extremity:
76880, 76536
Computed Tomography Scan:
71260, 71270, 71275
Ventilation Perfusion Lung Scan:

CPT Codes Used for VTE Diagnostic Studies
78579, 78580, 78582, 78584-78588, 78591, 78593, 78594, 78596
Computed Tomography Angiogram Chest:
71275
Pulmonary Angiogram:
75741, 75743, 75746

HIGHLIGHTS

- Use of administrative data is increasing for research and quality assessment.
- Outcomes within administrative data can be captured using ICD-9 coding.
- This algorithm captures VTE outcomes with a PPV of 91%.
- This algorithm facilitates accurate identification of VTE in administrative data.

Table 1

Algorithm's for Detection of Venous Thromboembolism

Algorithm (1)	**Algorithm (2)**	**Algorithm (3)**
• ICD-9-CM code for VTE	• ICD-9-CM code for VTE PLUS (at least one of the following criteria) • Pharmacy Script for Anticoagulation Within 30 days of ICD-9 diagnosis of VTE • Placement of an Inferior Vena Cava Filter Within 30 days of ICD-9 diagnosis of VTE • Death Within 30 days of ICD-9 diagnosis	• ICD-9-CM code for VTE PLUS (at least one of the following criteria) • Pharmacy Script for Anticoagulation Within 30 days of ICD-9 diagnosis of VTE • Placement of an Inferior Vena Cava Filter Within 30 days of ICD-9 diagnosis of VTE • Death Within 30 days of ICD-9 diagnosis PLUS • CPT code for a VTE Diagnostic Study

Table 2

Demographic and Clinical Characteristics stratified by NHL type

Characteristics	Overall (n=4360)
Mean Age (years)	65.8
Sex (Male %)	96.9
Race (%)	
White	86.8
Black	10.7
Other	1.0
Unknown	1.5
Stage (%)	
Stage I/II	39.9
Stage III/ IV	57.2
Unknown	3.0
Charlson score (mean)	2.4
BMI groups (%)	
BMI < 18.5	2.7
18.5 <= BMI < 25	31.3
25 <= BMI < 30	34.3
BMI>=30	22.3
Unknown	9.4

Table 3

Sensitivity, Specificity, and PPV for each algorithm

Algorithm	TP	FP	TN	FN	Sensitivity	Specificity	PPV (95% CI)	NPV (95% CI)
(1) ICD-9 only	108	42	148	2	0.91 (0.85, 0.96)	0.95 (0.94, 0.96)	0.72 (0.65, 0.79)	0.99 (0.98, 0.99)
(2) ICD-9 *plus* Evidence of Treatment or Death	136	14	142	8	0.72 (0.65, 0.78)	0.99 (0.98, 0.99)	0.91 (0.86, 0.95)	0.95 (0.93, 0.96)
(3) ICD-9 *plus* Evidence of Treatment or Death *plus* CPT Diagnostic Study Code	138	12	141	9	0.67 (0.60, 0.73)	0.99 (0.98, 0.99)	0.92 (0.87, 0.96)	0.94 (0.93, 0.95)