

# The accuracy of an electronic Pulmonary Embolism Severity Index auto-populated from the electronic health record

## Setting the stage for computerized clinical decision support

D.R. Vinson<sup>1,2,3</sup>; J.E. Morley<sup>4</sup>; J. Huang<sup>3</sup>; V. Liu<sup>1,3,5</sup>; M.L. Anderson<sup>1,2</sup>; C. E. Drenten<sup>6</sup>; R.P. Radecki<sup>7</sup>; D.K. Nishijima<sup>4</sup>; M.E. Reed<sup>3</sup>; on behalf of the Kaisers Permanente CREST Network

<sup>1</sup>The Permanente Medical Group, Oakland, California; <sup>2</sup>Department of Emergency Medicine, Kaiser Permanente Roseville Medical Center, Roseville, California; <sup>3</sup>Kaiser Permanente Division of Research, Oakland, California; <sup>4</sup>Department of Emergency Medicine, University of California Davis School of Medicine, Sacramento, California; <sup>5</sup>Department of Pulmonary and Critical Care Medicine, Kaiser Permanente Santa Clara Medical Center, Santa Clara, California; <sup>6</sup>Department of Emergency Medicine, Sutter General Medical Center, Sacramento, California; <sup>7</sup>Department of Emergency Medicine, The University of Texas Medical School, Houston, Texas

### Keywords

Clinical decision support systems; electronic health record; risk assessment; pulmonary embolism; emergency medicine; data completeness

### Summary

**Background:** The Pulmonary Embolism (PE) Severity Index identifies emergency department (ED) patients with acute PE that can be safely managed without hospitalization. However, the Index comprises 11 weighted variables, complexity that can impede its integration into contextual workflow.

**Objective:** We designed a computerized version of the PE Severity Index (e-Index) to automatically extract the required variables from discrete fields in the electronic health record (EHR). We tested the e-Index on the study population to determine its accuracy compared with a gold standard generated by physician abstraction of the EHR on manual chart review.

**Methods:** This retrospective cohort study included adults with objectively-confirmed acute PE in four community EDs from 2010–2012. Outcomes included performance characteristics of the e-Index for individual values, the number of cases requiring physician editing, and the accuracy of the e-Index risk category (low vs. higher).

**Results:** For the 593 eligible patients, there were 6,523 values automatically extracted. Fifty one of these needed physician editing, yielding an accuracy at the value-level of 99.2% (95% confidence interval [CI], 99.0%-99.4%). Sensitivity was 96.9% (95% CI, 96.0%-97.9%) and specificity was 99.8% (95% CI, 99.7%-99.9%). The 51 corrected values were distributed among 47 cases: 43 cases required the correction of one variable and four cases required the correction of two. At the risk-category level, the e-Index had an accuracy of 96.8% (95% CI, 95.0%-98.0%), under-classifying 16 higher-risk cases (2.7%) and over-classifying 3 low-risk cases (0.5%).

**Conclusion:** Our automated extraction of variables from the EHR for the e-Index demonstrates substantial accuracy, requiring a minimum of physician editing. This should increase user acceptability and implementation success of a computerized clinical decision support system built around the e-Index, and may serve as a model to automate other complex risk stratification instruments.

**Correspondence to:**

David R Vinson, MD  
Department of Emergency Medicine  
Kaiser Permanente Roseville Medical Center  
1600 Eureka Road  
Roseville, CA 95662, USA  
Email: drvinson@ucdavis.edu

**Appl Clin Inform** 2015 ;6: 318–333<http://dx.doi.org/10.4338/ACI-2014-12-RA-0116>

received: December 16, 2014

accepted: March 27, 2015

published: May 13, 2015

**Citation:** Vinson DR, Morley JE, Huang J, Liu V, Anderson ML, Drenten CE, Radecki RP, Nishijima DK, Reed ME. The accuracy of an electronic pulmonary embolism severity index auto-populated from the electronic health record. *Appl Clin Inf* 2015; 6: 318–333  
<http://dx.doi.org/10.4338/ACI-2014-12-RA-0116>

## 1. Background

Select patients in the emergency department (ED) with acute pulmonary embolism (PE) designated low risk by the PE Severity Index can be safely managed without hospitalization [1, 2]. The PE Severity Index has wide generalizability because it comprises variables readily accessible at the bedside, without need of specialized laboratory or imaging results. The variables, however, 11 in total, are each weighted, and collectively sum to a score that needs to be matched to a five-tiered ranking of estimated 30-day all-cause mortality (► Table 1). This complexity can make the PE Severity Index difficult to remember and cumbersome to use.

To facilitate ease of use of the PE Severity Index by emergency and internal medicine clinicians at the point of care, we sought to design a computerized clinical decision-support (CDS) tool that could extract data already captured in discrete fields in the electronic health record (EHR) during routine care [3]. To minimize the tedious and error-prone activity of manual data entry, reduce distractions from clinical workflow and increase acceptability of the CDS, we designed it to automatically extract the Index's 11 variables directly from the patient's EHR and present to the user a pre-populated electronically-derived PE Severity Index, which we refer to hereafter as the e-Index.

Diagnostic coding in administrative databases, however, is known to be inaccurate to varying degrees, a shortcoming that could have large effects on the quality of clinical research [4–10]. Drawing from the EHR to populate the comorbidity elements of the e-Index then may predispose to errors of risk miscategorization, as others have demonstrated with different risk instruments [4]. Since an accurate list of patient diagnoses is critical for many reasons (patient care, clinical decision support, population health reporting, quality improvement, and research), some medical groups have put considerable effort into improving and maintaining an accurate, up-to-date, physician-managed Problem List in the EHR [11, 12]. But the accuracy of tapping an electronic Problem List as the source of diagnoses for a CDS tool, as we will be doing, has not been well studied.

In addition to coding errors and omissions in the Problem List, the e-Index has other inherent limitations: it is unable to identify all relevant vital signs, including pre-arrival measurements from emergency medical services as well as unstructured measurements during the ED stay, and it cannot easily identify the presence of altered mental status, since documentation of this examination finding is neither uniform nor standardized. The e-Index, then, would need to be editable by the clinician user in order to correct these shortcomings. The amount of editing required for the e-Index is unknown, but the degree of inaccuracy could well effect its usability.

## 2. Objectives

Prior to incorporating the e-Index in a CDS system used in a live environment, characterization of its accuracy is necessary. We undertook this study of adult ED patients with objectively-confirmed PE to simulate use of the e-Index and test its performance characteristics by measuring the degree of editing of each variable undertaken by physician abstractors viewing the documentation in the EHR from the patient's index ED encounter. The physician-edited version of the e-Index (what we call the p-Index) is the gold standard against which the e-Index was compared and the basis for our assessment of e-Index accuracy.

## 3. Methods

### 3.1. Study Design and Setting

This retrospective cohort study was undertaken in four community EDs between 2010 and 2012 within Kaiser Permanente Northern California, a large integrated health care delivery system. The health system is supported by an Epic-based (Verona, WI) EHR implemented in 2005 [13].

### 3.2. Study Population

This current project represents one arm of a larger study that had two pre-defined aims: to characterize the performance of the e-Index, presented here, and to improve the risk stratification capabilities of the PE Severity Index for a low-risk population, published elsewhere [14]. The second arm of the study shaped the eligibility criteria. The process of cohort assembly is detailed in ► Figure 1. The three risk stratification instruments we used to electronically identify patients to be examined for eligibility were the PE Severity Index [15], the simplified PE Severity Index [16], and the PE Triage Score [17]. Our final cohort consisted of adult ( $\geq 18$  years) ED patients with acute objectively-confirmed PE, based on the official board-certified radiologist or nuclear medicine interpretation of a positive computed tomography pulmonary angiogram or high-probability ventilation perfusion scan, respectively.

### 3.3. The electronic-Index (e-Index)

The patient comorbidities of the e-Index were drawn from the active Problem List in the EHR using codes from *International Classification of Diseases*, ninth edition (ICD-9) (► Table 2). The accuracy of the Problem List in Kaiser Permanente Northern California has been the result of an ongoing system-wide effort since its inception. Primary care providers and specialists have continued to actively maintain the Problem List so it can serve as a reliable snapshot of patient health and a useful tool for patient care.

Vital signs were drawn from structured nursing documentation, and limited to those recorded following the patient's arrival in the ED. Identification of altered mental status was inferred from abnormal assessments in four fields of the nursing flowsheet documentation: (a) level of consciousness (obtunded, or stuporous, or comatose); (b) orientation (confused); (c) the mentation element of the Schmid fall-risk assessment (periodic confusion, confused at all times, or comatose); (d) altered mental status (yes). A singly documented instance of (a) or two or more instances of (b), (c), or (d), in one or a combination of these three fields, counted as positive on the altered mentation variable of the e-Index.

### 3.4. The physician-edited Index (p-Index)

To simulate the implementation of the e-Index in clinical practice, the physician abstractors were presented with the results of the 11 automatically extracted variables of the e-Index. Individual abstractors edited those values as needed during manual chart review to create the p-Index using only data judged available to the treating clinician at the time of the patient encounter (► Figure 2).

The five physician abstractors were practitioners in the Kaiser Permanente Northern California integrated healthcare system and were well-versed in the system's EHR. Four of them are emergency physicians and one is a pulmonology and critical care specialist, all of whom have extensive experience in the diagnosis and management of patients with acute PE. They were each trained in the definitions of the variables of the PE Severity Index and the process of data collection with a standardized electronic data collection tool that was pre-populated with the results of the e-Index. Ambiguities in interpretation of the EHR were discussed with the principal investigator and resolved by consensus.

The abstractors reviewed both structured and unstructured data related to the index patient encounter. Additional data available to manual abstractors from the physicians' notes, but not extracted by the e-Index, included pre-arrival vital signs and unstructured ED vital sign measurements, unstructured assessments of mental status, and elements of medical history incompletely documented in the active Problem List in the EHR.

### 3.5. Statistical Analysis

We compared the values extracted for the e-Index against those in the edited p-Index. We report the sensitivity, specificity, positive predictive value, negative predictive value, and accuracy (with 95% confidence intervals) of the e-Index, for each variable, each category of variables, and collectively.

We calculated variable-level accuracy as (total values – [false negative values + false positive values])/total values. We also report the number of cases that required physician editing of e-Index values.

PE Severity Index scores  $\leq 85$  points (Classes I-II) were categorized as low risk and scores  $> 85$  points (Classes III-V) were categorized as higher risk [15]. We calculated risk categories for each patient using both the e-Index and p-Index. Clinical misclassifications of low-risk vs. not low-risk are reported, and patient-level accuracy was calculated as (total patients – [over-classification + under classification])/total patients.

We randomly selected 10% of cases ( $n=60$ ) for review by a second investigator for inter-rater reliability of the 11 variables of the PE Severity Index, reported as percent agreement.

## 4. Results

During the three-year study period we identified 593 individual ED patients with acute objectively-confirmed PE who met study eligibility criteria (► Figure 1). The risk spectrum of our population according to the p-Index was 73.2% low risk ( $n=434$ ) and 26.8% higher risk ( $n=159$ ) patients. Five hundred fifty one patients (92.9%) were active health plan members at the time of the index ED visit.

Of the 6,523 values extracted for the e-Index, 51 were corrected by physician abstractors. There were 11 false positive (0.2%) and 40 false negative (0.6%) values in the e-Index, resulting in a variable-level accuracy of 99.2% (95% confidence interval [CI], 99.0%-99.4%). The 51 corrected values were distributed among 47 cases: 43 cases required correction through manual chart review of one variable and four cases required correction of two variables. The majority of these edits (33/51; 65%) were attributable to an incomplete or erroneous Problem List in the EHR. Fewer cases ( $n=13$ ; 2.2% of the total cohort) had normal ED vital signs but abnormal pre-arrival or unstructured ED vital signs that crossed the Index point threshold (► Table 1) and needed correction.

The performance characteristics of the auto-populating e-Index are reported in ► Table 3. Patients who were not active health plan members at the time of the index ED visit were no more likely to have false negative comorbidity findings on their e-Index than active members: 4.7% vs 3.8% ( $P=0.67$ ).

When the patients were stratified by PE Severity Index risk category (low vs higher), the e-Index misclassified 19 cases (3.2%), resulting in a risk-category accuracy of 96.8% (95% CI, 95.0%-98.0%). Sixteen higher-risk cases (2.7%) were mistakenly classified by the e-Index as low-risk (that is, they were under-classified) and 3 low-risk cases (0.5%) were mistakenly classified by the e-Index as higher-risk (that is, they were over-classified).

The inter-rater reliability, measured as percent agreement, was 99.4% at the value level (656/660). At the variable category level, the results were as follows: demographics 100% (120/120), comorbidities 98.9% (178/180), vital signs 99.3% (298/300), and mentation 100% (60/60).

## 5. Discussion

We developed and tested an electronic, auto-populating version of the PE Severity Index, which leveraged existing structured data elements in the EHR, against a version edited by physician abstractors during manual chart review. Despite the limitations of the e-Index (e.g., an inability to discern errors in the Problem List and to access pre-arrival vital signs), we found it highly accurate at the level of individual values (99.2%). The e-Index was also highly accurate at risk assessment, properly categorizing nearly 97% of cases to the appropriate risk strata: low risk and higher risk. This level of categorization is important because a low-risk status is an integral component in determining patient eligibility for home management [1].

Though the misclassification errors of the unedited e-Index were infrequent, they could have clinical implications for site-of-care management decisions. This is why use of the e-Index as an unedited, stand-alone risk assessment tool may be inadvisable. With only a minimal amount of editing, however, the e-Index can be relied upon for clinical decision support.

The small number of edits required to improve the accuracy of the auto-populated e-Index involves significantly less data entry on the part of the clinician user (51 edits for 593 patients) than if the Index required manual data entry throughout (6,523 variables). This reduced workload may well help increase the perceived ease-of-use of the tool, which is a key determinant of usability [18]. Streamlining user effort by minimizing the number of clicks to complete the Index may also improve workflow [19, 20].

The least accurate component of the e-Index, correct in 95% of cases, was the identification of comorbidities (cancer, heart failure, and chronic lung disease) from the Problem List. False negatives, which reflect incompleteness of the Problem List at the time of the ED encounter, were twice as common as false positives. These were diagnoses that were recorded in the treating physicians' documentation, based on other sources in the EHR or the ED evaluation. Less common than diagnoses missing from the Problem List were false positives. These could be explained by miscoded diagnoses or from diagnostic codes that were imprecise, that is, sensitive for the Index's comorbidity in question but not specific for it. Nevertheless, an accuracy rate of 95% compares quite favorably with published reports of diagnostic miscoding in administrative databases [4–7, 9, 10] and may reflect the advantages of maintaining an up-to-date Problem List.

The clinical implications of a mistaken Problem List can be illustrated by the following case. A 69-year-old woman with a history of asthma and breast cancer (both documented in the Problem List) was found in the ED to have an acute PE and normal vital signs. This patient profile would generate an accurate e-Index score of 109 points, placing her in Class IV, which is associated with a higher estimated 30-day mortality (Table 1) [15]. If her malignancy, however, had been missing from the Problem List, the auto-populated e-Index would fail to include the cancer and hence miscalculate her score by 30 points. The faulty tally of 79 points would misassign her to a low-risk category (Class II). Such a risk categorization would support the appropriateness of outpatient management, absent concerning contraindications [1, 14]. But as this study demonstrates, the e-Index when used in the ED setting will require review by the treating clinician to confirm the accuracy of the patient's comorbidities and to correct them as needed (► Figure 2). When the clinician elicits the history of cancer from the EHR or the patient/family interview, the e-Index can be readily corrected with one click. The updated comorbidities would then be used to electronically calculate an accurate PE Severity Index score, class, and risk category, providing the clinician with valuable clinical decision support.

Vital signs may also require editing, though in fewer cases than comorbidities (2% vs 5%). Virtually all of these were from cases with abnormal pre-arrival vital signs that were detected in the physicians' unstructured documentation. For example, a patient was noted by the paramedic to have an oxygen saturation at home of 86%. Though the treating emergency physician documented this pertinent finding in their free-text history of the present illness, the e-Index reported a falsely negative (that is, normal) oxygen saturation because the patient, on supplemental oxygen in the ED, was now saturating at 96%. Though the clinician making the site-of-care decision for an ED patient with acute PE would take into account abnormal vital signs reported by emergency medical services or the transferring outpatient clinician, few studies of the PE Severity Index include pre-arrival vital signs in their scoring. This oversight could bias the risk assessment to underclassification, as we found with the e-Index.

Altered mental status was expectedly uncommon in this predominantly low-risk population, and so the number of cases of this parameter that needed editing were small. However, only half of the patients with altered mentation were identified by the e-Index. Unlike ED vital signs, mentation is not an examination finding that is universally documented in a discrete field in the EHR. To access the patient's mental status, the e-Index queried fields in the nursing flowsheets that are not uniformly completed and used these data in a formula that proved to be specific but insensitive. This component of the e-Index was the most unreliable of the variables of the e-Index.

The CURB-65 score, a validated risk stratification instrument for the identification of low-risk patients with pneumonia eligible for outpatient management, also includes an assessment of altered mentation, though its definition ("confusion") is more restricted than that used by the PE Severity Index (Table 1). Jones BE et al identified "confusion" in the EHR by querying structured nursing documentation for evidence of a Glasgow Coma Scale score less than 15 or an abnormal orientation in the mentation element of the Schmid fall-risk assessment [21].



The PE Severity Index is not the only complex risk stratification or prognostic tool that could be predominantly auto-populated by drawing on data that is easily accessible in discrete fields in the EHR. Examples include the Pneumonia Severity Index [22] and CURB-65 [23] for pneumonia-associated mortality, CHA<sub>2</sub>DS<sub>2</sub>-VASc [24] and ATRIA [25] for stroke risk associated with atrial fibrillation, and HAS-BLED [26] for anticoagulation bleeding risks. Designing auto-populating risk stratification tools for these and other clinical decisions may help facilitate the translation of evidence-based medicine into the routines of clinical practice [27].

Others have also assessed the accuracy of data drawn from the EHR to auto-populate risk stratification instruments used in CDS tools. For example, Jones BE et al compared electronic data with results from manual chart review on a sample of 50 patients in a study of the CURB-65 [21]. They report greater than 90% agreement between electronic and paper records. The categories of variables used in the CURB-65 clinical prediction rule, however, differ from those used in the PE Severity Index. Though both include demographics, vital signs, and some measure of mental status, the CURB-65 rule also includes laboratory values but not any comorbidities.

Successful and safe implementation of the e-Index in a computerized CDS system in the ED would require educating clinician users about the sources of the auto-populated variables and the inherent inaccuracies of the tool, particularly the occasional false positives and false negatives of the comorbidities, the inability of the e-Index to access unstructured vital sign measurements, and the insensitivity of the mental status assessment. Explaining these several shortcomings will highlight the importance of user input to confirm the auto-populated values and correct them as needed. How clinician-editors perform in real-time compared with investigator-editors during manual chart review will be an important measure.

The site-of-care aspect of patient management is critical because issues of patient safety and resource allocation hinge on such crossroads decisions of patient disposition, like outpatient vs inpatient care or a medical ward vs intensive care unit admission. Much research has been undertaken on improving the effectiveness of CDS systems in general [3, 28–32]. But most of this work has focused on compliance with recommended preventive services, the ordering of diagnostic tests and studies, and the prescribing of treatments [31]. But today, with research in computerized CDS on the rise [33], more attention is being paid to employing such a system to aid in site-of-care decisions [21, 34–36]. These efforts are likely to expand as we improve our abilities to employ health information technology in the service of cost-conscious, safety-driven evidence-based medicine [37–39].

Our results are tempered by several limitations. Because the study population was selected to target the low-risk end of the PE severity spectrum, we cannot say that our results will hold up with a larger population of higher-risk patients who have a greater prevalence of comorbidities. However, it is this lower risk population where site-of-care clinical decision support may be most useful. Also, the codes, sources, and formulae we used to identify comorbidities and mental status were derived for this study. They will require validation in other settings and populations. Our results, likewise, are specific to our Epic-based EHR and may not be generalizable to other EHR systems. Lastly, our patient population consisted predominantly of health plan members whose Problem List in the EHR was populated with their comorbidities prior to arrival in the ED. Even the subpopulation of patients who were not active members at the time of the index ED visit had a sufficiently complete Problem List to facilitate accuracy of the e-Index. One explanation for this finding is that these patients may have been prior members or, even if non-members, they may have frequented the medical centers sufficiently to have generated a sufficiently complete Problem List. Application of the e-Index to a population with a larger proportion of first-time patients would have a higher percentage of false negative comorbidities and would require greater user input.

## 6. Conclusions

We found that an auto-populating electronic PE Severity Index was sufficiently accurate to require minimal user input. This degree of accuracy increases the likelihood of its successful implementation in a real-time CDS system. Piloting the tool in clinical practice will help us determine if this hypothesis is true. If it functions well in clinical practice, this auto-populating computerized CDS

tool could provide a model for other complex risk stratification instruments that draw predominantly on data found in the EHR.

### **Clinical Relevance**

We designed an auto-populating electronic Pulmonary Embolism Severity Index that is sufficiently accurate so as to require minimal clinician editing. The e-Index should facilitate user acceptability and implementation success when placed in a computerized clinical decision support system. Accurate auto-population may expedite the knowledge translation of validated complex risk stratification instruments.

### **Conflicts of Interest**

The authors declare that they have no conflicts of interest in the research.

### **Human Subjects**

The 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 Kaiser Permanente Northern California Health Services Institutional Review Board.

### **Acknowledgements**

We are grateful to Kaiser Permanente Northern California Community Benefit Program for their generous grant and to Adina S Rauchwerger, MPH, for her invaluable administrative work. We also thank the leadership in The Permanente Medical Group for their continued support of our clinical research program. Vincent Liu was also supported by NIH K23GM11201.



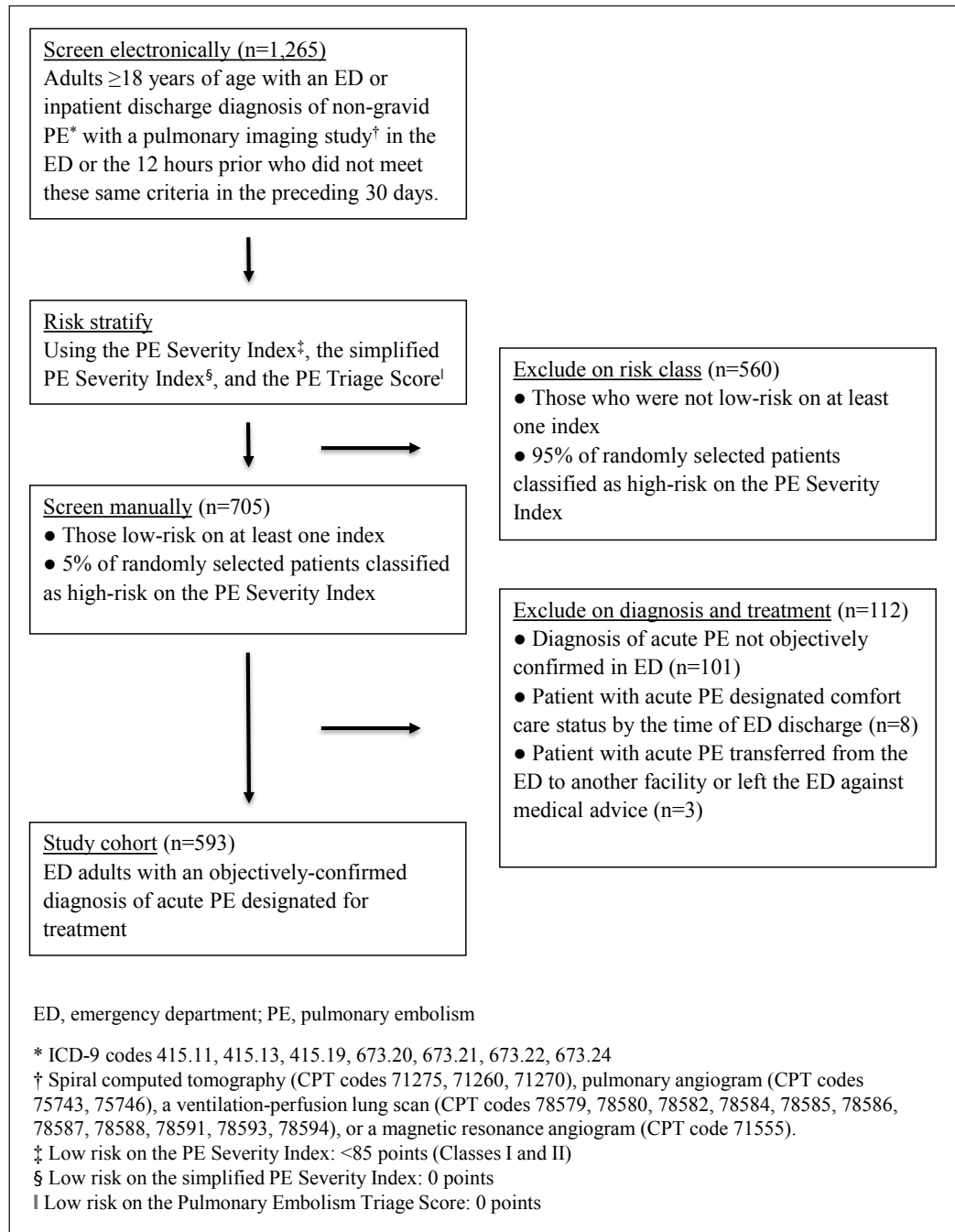
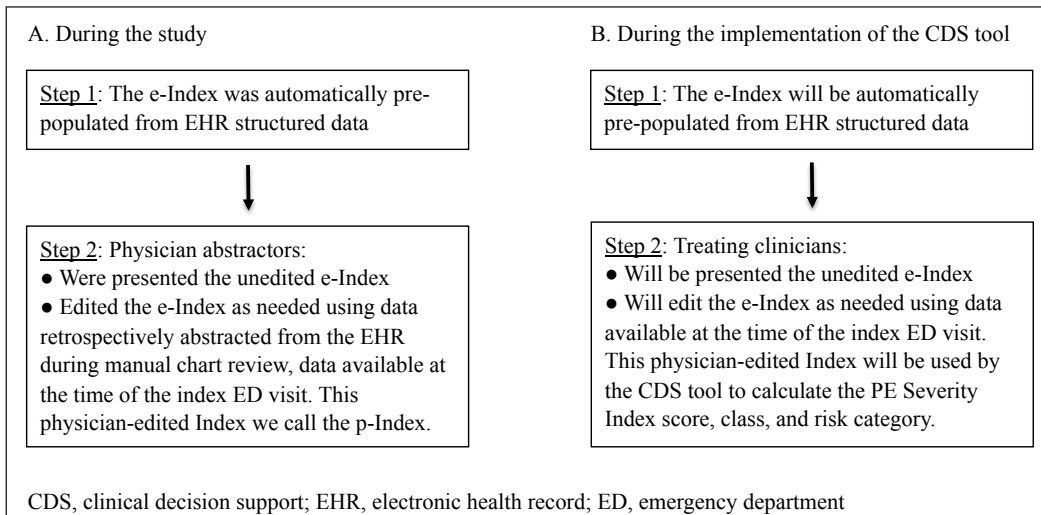


Fig. 1 Cohort assembly

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.



**Fig. 2** Editing the electronic Index (e-Index) in two different scenarios

Predictors	Points <sup>a</sup>
<b>Demographic Characteristics</b>	
Age	+ 1 per year
Male sex	+ 10
<b>Comorbid Illness</b>	
Cancer (active or history of)	+ 30
Heart Failure (systolic or diastolic)	+ 10
Chronic Lung Disease (includes asthma)	+ 10
<b>Clinical Findings<sup>b</sup></b>	
Pulse $\geq 110$ /min beats per min	+ 20
Systolic blood pressure $< 100$ mmHg	+ 30
Respiratory rate $\geq 30$ breaths per min	+ 20
Temperature $< 36^\circ$ C	+ 20
Altered mental status <sup>c</sup>	+ 60
Arterial oxygen saturation $< 90\%$ <sup>d</sup>	+ 20

**Table 1** The Pulmonary Embolism Severity Index

a. A total point score for a given patient is obtained by summing the patient's age in years and the points for each applicable characteristic. Point scores correspond with the following classes that estimate escalating risks of 30-day mortality:  $\leq 65$  class I, very low risk; 66–85 class II, low risk; 86–105 class III, intermediate risk; 106–125 class IV, high risk;  $> 125$  class V, very high risk.

b. The worst vital signs from the ED stay are used

c. Acute or pre-existing disorientation, lethargy, stupor, or coma

d. With or without supplemental oxygenation

**Table 2** Sources and definitions of the auto-populated e-Index and the physician-edited p-Index.

Variable	Auto-populated e-Index	Physician-edited p-Index
	Source	Source
Demographics (age and sex)	Electronic health record (EHR)	Emergency physician note during emergency department (ED) visit
<b>Comorbidities</b>	<b>Definition</b>	
Cancer (active or history of)	Age in years; sex: male or female	Emergency physician and hospitalist notes during ED visit
Heart Failure (systolic or diastolic)	<b>ICD-9 codes</b>	Emergency physician and hospitalist notes during ED visit
Chronic Lung Disease (includes asthma)	Problem list in the EHR at the time of the index ED visit	Emergency physician and hospitalist notes during ED visit
	Problem list in the EHR at the time of the index ED visit	Emergency physician and hospitalist notes during ED visit
	Problem list in the EHR at the time of the index ED visit	Emergency physician and hospitalist notes during ED visit
Vital Signs	EHR	Expanded to include unstructured vital signs documented in emergency physician or hospitalist notes during ED visit <sup>a</sup>
Mental status	Nursing EHR flowsheet documentation	Emergency physician, nursing, and hospitalist notes during ED visit

a. Abstractors did not reference the original pre-hospital emergency medical care documentation or the pre-ED clinic notes.

**Table 3** Performance characteristics of an auto-populating electronic Pulmonary Embolism Severity Index

Index Variable	Values n	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Accuracy % (95% CI)	Cases needing editing (N=593) (%) ; 95% CI)	Reason for prevalent inaccuracies
All 11 variables	6,523	96.9 (96.0–97.9)	99.8 (99.7–99.9)	99.1 (98.6–99.6)	99.2 (99.0–99.5)	99.2 (99.0–99.4)	47 (7.9; 6.0–10.4)	
Demographics	1,186	100.0	100.0	100.0	100.0	100.0 (99.6–100.0)	0 (0; 0–0.8)	
Age	593	100.0	--	100.0	--	100.0 (99.2–100.0)		
Sex (male)	593	100.0	100.0	100.0	100.0	100.0 (99.2–100.0)		
Comorbidities	1,779	87.1 (82.2–92.0)	99.4 (99.0–99.8)	93.9 (90.3–97.6)	98.6 (98.0–99.2)	98.1 (97.4–98.7)	33 (5.6; 4.0–7.7)	Problem List incomplete or erroneous
Cancer	593	70.0 (53.6–86.4)	99.8 (99.5–100.0)	95.5 (86.8–100.0)	98.4 (97.4–99.4)	98.3 (96.9–99.1)		
Heart failure	593	83.3 (72.1–94.6)	99.6 (99.1–100.0)	94.6 (87.3–100.0)	98.7 (97.8–99.7)	98.5 (97.1–99.3)		
Chronic lung disease	593	93.4 (88.7–98.1)	98.6 (97.5–99.6)	93.4 (88.7–98.1)	98.6 (97.5–99.6)	97.6 (96.0–98.6)		
Vital signs	2,965	95.3 (92.8–97.9)	100.0	99.6 (98.8–100.0)	99.6 (99.3–99.8)	99.6 (99.2–99.8)	13 (2.2; 1.3–3.8)	Inaccessibility of pre-arrival and unstructured values
Pulse	593	99.1 (97.2–100.0)	100.0	100.0	99.8 (99.3–99.8)	99.8 (98.9–100.0)		
Systolic blood pressure	593	97.0 (92.9–100.0)	100.0	100.0	99.6 (99.1–100.0)	99.7 (98.7–100.0)		

**Table 3** Performance characteristics of an auto-populating electronic Pulmonary Embolism Severity Index

Index Variable	Values n	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Accuracy % (95% CI)	Cases needing editing (N=593) (%; 95% CI)	Reason for prevalent inaccuracies
Respiratory rate	593	100.0	100.0	100.0	100.0	100.0 (99.2–100.0)		
Temperature	593	100.0	100.0	100.0	100.0	100.0 (99.2–100.0)		
Arterial oxygen saturation	593	81.6 (70.8–92.5)	99.8 (99.5–100.0)	97.6 (92.8–100.0)	98.4 (97.3–99.4)	98.3 (96.9–99.1)		
Altered mental status (AMS)	593	44.4 (12.0–76.9)	100.0	100.0	99.2 (98.4–99.9)	99.2 (98.0–99.7)	5 (0.8; 0.3–2.0)	Insensitivity of our AMS identification formula

Sens, sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval



## References

1. Aujesky D, Roy PM, Verschuren F, Righini M, Osterwalder J, Egloff M, Renaud B, Verhamme P, Stone RA, Legall C, Sanchez O, Pugh NA, N'Gako A, Cornuz J, Hugli O, Beer HJ, Perrier A, Fine MJ, Yealy DM. Out-patient versus inpatient treatment for patients with acute pulmonary embolism: an international, open-label, randomised, non-inferiority trial. *Lancet* 2011; 378: 41-48.
2. Vinson DR, Zehtabchi S, Yealy DM. Can selected patients with newly diagnosed pulmonary embolism be safely treated without hospitalization? A systematic review. *Ann Emerg Med* 2012; 60: 651-662.
3. Moja L, Liberati EG, Galuppo L, Gorli M, Maraldi M, Nanni O, Rigon G, Ruggieri P, Ruggiero F, Scaratti G, Vaona A, Kwag KH. Barriers and facilitators to the uptake of computerized clinical decision support systems in specialty hospitals: protocol for a qualitative cross-sectional study. *Implement Sci* 2014; 9: 105.
4. Navar-Boggan AM, Rymer JA, Piccini JP, Shatila W, Ring L, Stafford JA, Al-Khatib SM, Peterson ED. Accuracy and validation of an automated electronic algorithm to identify patients with atrial fibrillation at risk for stroke. *Am Heart J* 2015; 169: 39-44.
5. Thigpen JL, Dillon C, Forster KB, Henault L, Quinn EK, Tripodis Y, Berger PB, Hylek EM, Limdi NA. Validity of International Classification of Disease Codes to Identify Ischemic Stroke and Intracranial Hemorrhage Among Individuals With Associated Diagnosis of Atrial Fibrillation. *Circ Cardiovasc Qual Outcomes* 2015; 8: 8-14.
6. 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: e104519.
7. Nouraei SA, Hudovsky A, Frampton AE, Mufti U, White NB, Wathen CG, Sandhu GS, Darzi A. A Study of Clinical Coding Accuracy in Surgery: Implications for the Use of Administrative Big Data for Outcomes Management. *Ann Surg* 2014 Dec 2 [Epub ahead of print].
8. Weiskopf NG, Hripcsak G, Swaminathan S, Weng C. Defining and measuring completeness of electronic health records for secondary use. *J Biomed Inform* 2013; 46: 830-836.
9. 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: 388-394.
10. Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T. Identifying patients with physician-diagnosed asthma in health administrative databases. *Can Respir J* 2009; 16: 183-188.
11. Wright A, Pang J, Feblowitz JC, Maloney FL, Wilcox AR, McLoughlin KS, Ramelson H, Schneider L, Bates DW. Improving completeness of electronic problem lists through clinical decision support: a randomized, controlled trial. *J Am Med Inform Assoc* 2012; 19: 555-561.
12. Bakel LA, Wilson K, Tyler A, Tham E, Reese J, Bothner J, Kaplan DW. A quality improvement study to improve inpatient problem list use. *Hosp Pediatr* 2014; 4: 205-210.
13. Bornstein S. An integrated EHR at Northern California Kaiser Permanente: Pitfalls, challenges, and benefits experienced in transitioning. *Appl Clin Inform* 2012; 3: 318-325.
14. Vinson DR, Drenten CE, Huang J, Morley JE, Anderson ML, Reed ME, Nishijima DK, Liu V, on behalf of the Kaisers Permanente CREST Network. Impact of Relative Contraindications to Home Management in Emergency Department Patients with Low-Risk Pulmonary Embolism. *Ann Am Thorac Soc* 2015 Feb 19 [epub ahead of print].
15. Aujesky D, Obrosky DS, Stone RA, Auble TE, Perrier A, Cornuz J, Roy PM, Fine MJ. Derivation and validation of a prognostic model for pulmonary embolism. *Am J Respir Crit Care Med* 2005; 172: 1041-1046.
16. Jimenez D, Aujesky D, Moores L, Gomez V, Lobo JL, Uresandi F, Otero R, Monreal M, Muriel A, Yusen RD. Simplification of the pulmonary embolism severity index for prognostication in patients with acute symptomatic pulmonary embolism. *Arch Intern Med* 2010; 170: 1383-1389.
17. Briese B, Schreiber D, Lin B, Liu G, Fansler J, Goldhaber S, O'Neil B, Slattery D, Hiestand B, Kline J, Pollack C. Derivation of a simplified Pulmonary Embolism Triage Score (PETS) to predict the mortality in patients with confirmed pulmonary embolism from the Emergency Medicine Pulmonary Embolism in the Real World Registry (EMPEROR). *Acad Emerg Med* 2012; 19: S143-S144.
18. Lobach D, Sanders GD, Bright TJ, Wong A, Dhurjati R, Bristow E, Bastian L, Coeytaux R, Samsa G, Haselblad V, Williams JW, Wing L, Musty M, Kendrick AS. Enabling health care decisionmaking through clinical decision support and knowledge management. *Evid Rep Technol Assess (Full Rep)* 2012: 1-784.
19. Bishop RO, Patrick J, Besiso A. Efficiency achievements from a user-developed real-time modifiable clinical information system. *Ann Emerg Med* 2015; 65: 133-142.
20. Horsky J, Phansalkar S, Desai A, Bell D, Middleton B. Design of decision support interventions for medication prescribing. *Int J Med Inform* 2013; 82: 492-503.
21. Jones BE, Jones J, Bewick T, Lim WS, Aronsky D, Brown SM, Boersma WG, van der Eerden MM, Dean NC. CURB-65 pneumonia severity assessment adapted for electronic decision support. *Chest* 2011; 140: 156-163.

22. Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, Coley CM, Marrie TJ, Kapoor WN. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med* 1997; 336: 243–250.
23. Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, Lewis SA, Macfarlane JT. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 2003; 58: 377–382.
24. Olesen JB, Lip GY, Hansen ML, Hansen PR, Tolstrup JS, Lindhardsen J, Selmer C, Ahlehoff O, Olsen AM, Gislason GH, Torp-Pedersen C. Validation of risk stratification schemes for predicting stroke and thromboembolism in patients with atrial fibrillation: nationwide cohort study. *BMJ* 2011; 342: d124.
25. Singer DE, Chang Y, Borowsky LH, Fang MC, Pomernacki NK, Udaltsova N, Reynolds K, Go AS. A new risk scheme to predict ischemic stroke and other thromboembolism in atrial fibrillation: the ATRIA study stroke risk score. *J Am Heart Assoc* 2013; 2: e000250.
26. Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJ, Lip GY. A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey. *Chest* 2010; 138: 1093–1100.
27. Murphy EV. Clinical decision support: effectiveness in improving quality processes and clinical outcomes and factors that may influence success. *Yale J Biol Med* 2014; 87: 187–197.
28. Moja L, Kwag KH, Lytras T, Bertizzolo L, Brandt L, Pecoraro V, Rigon G, Vaona A, Ruggiero F, Mangia M, Iorio A, Kunnamo I, Bonovas S. Effectiveness of Computerized Decision Support Systems Linked to Electronic Health Records: A Systematic Review and Meta-Analysis. *Am J Public Health* 2014: e1-e11.
29. Khorasani R, Hentel K, Darer J, Langlotz C, Ip IK, Manaker S, Cardella J, Min R, Seltzer S. Ten commandments for effective clinical decision support for imaging: enabling evidence-based practice to improve quality and reduce waste. *AJR Am J Roentgenol* 2014; 203: 945–951.
30. Sedlmayr B, Patapovas A, Kirchner M, Sonst A, Muller F, Pfistermeister B, Plank-Kiegele B, Vogler R, Criegee-Rieck M, Prokosch HU, Dormann H, Maas R, Burkle T. Comparative evaluation of different medication safety measures for the emergency department: physicians' usage and acceptance of training, poster, checklist and computerized decision support. *BMC Med Inform Decis Mak* 2013; 13: 79.
31. Bright TJ, Wong A, Dhurjati R, Bristow E, Bastian L, Coeytaux RR, Samsa G, Hasselblad V, Williams JW, Musty MD, Wing L, Kendrick AS, Sanders GD, Lobach D. Effect of clinical decision-support systems: a systematic review. *Ann Intern Med* 2012; 157: 29–43.
32. Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, Volk L, Spurr C, Khorasani R, Tanasijevic M, Middleton B. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *J Am Med Inform Assoc* 2003; 10: 523–530.
33. Jia PL, Zhang PF, Li HD, Zhang LH, Chen Y, Zhang MM. Literature review on clinical decision support system reducing medical error. *J Evid Based Med* 2014; 7: 219–226.
34. Dean NC, Jones BE, Ferraro JP, Vines CG, Haug PJ. Performance and utilization of an emergency department electronic screening tool for pneumonia. *JAMA Intern Med* 2013; 173: 699–701.
35. Jones BE, Jones JP, Vines CG, Dean NC. Validating hospital admission criteria for decision support in pneumonia. *BMC Pulm Med* 2014; 14: 149.
36. Dean NC, Jones BE, Jones JP, Ferraro JP, Post HB, Aronsky D, Vines CG, Allen TL, Haug PJ. Impact of an Electronic Clinical Decision Support Tool for Emergency Department Patients With Pneumonia. *Ann Emerg Med* 2015 Feb 26 [Epub ahead of print].
37. Mann DM, Kannry JL, Edonyabo D, Li AC, Arciniega J, Stulman J, Romero L, Wisnivesky J, Adler R, McGinn TG. Rationale, design, and implementation protocol of an electronic health record integrated clinical prediction rule (iCPR) randomized trial in primary care. *Implement Sci* 2011; 6: 109.
38. Handel DA, Wears RL, Nathanson LA, Pines JM. Using information technology to improve the quality and safety of emergency care. *Acad Emerg Med* 2011; 18: e45–e51.
39. Mishuris RG, Linder JA, Bates DW, Bitton A. Using electronic health record clinical decision support is associated with improved quality of care. *Am J Manag Care* 2014; 20: e445–e452.