Avoiding alert fatigue in pulmonary embolism decision support: a new method to examine 'trigger rates'

Anne Press,[1] Sundas Khan,[1] Lauren McCullagh,[1] Andy Schachter,[1] Salvatore Pardo,[2] Nina Kohn,[3] Thomas McGinn[1]

10.1136/ebmed-2016-110440

[1]Department of Medicine, Hofstra Northwell School of Medicine, Manhasset, New York, USA
[2]Department of Emergency Medicine, Hofstra Northwell School of Medicine, Manhasset, New York, USA
[3]Feinstein Institute for Medical Research, Manhasset, New York, USA

Correspondence to **Anne Press,** Department of Medicine, Hofstra Northwell School of Medicine, 300 Community Drive, Manhasset, NY 11030, USA; anniepress@gmail.com

Abstract

A clinical decision support system (CDSS) is integrated into the electronic health record (EHR) and allows physicians to easily use a clinical decision support (CDS) tool. However, often CDSSs are integrated into the EHR with poor adoption rates. One reason for this is secondary to 'trigger fatigue'. Therefore, we developed a new and innovative usability process named 'sensitivity and specificity trigger analysis' (SSTA) as part of our larger project around a pulmonary embolism decision support tool. SSTA will enable programmers to examine optimal trigger rates prior to the integration of a CDS tool into the EHR, by using a formal method of analysis. We performed a retrospective chart review. The outcome of interest was physician ordering of a CT angiography (CTA). Phrases that signify common symptoms associated with pulmonary embolism were assessed as possible triggers for the CDSS tool. We then analysed each trigger's ability to predict physician ordering of a CTA. We found that the most sensitive way to trigger the Pulmonary Embolism CDS tool while still maintaining a high specificity was by combining 1 or more pertinent symptoms with 1 or more elements of the Wells criteria. This study explored a unique methodology, SSTA, used to limit inaccurate triggering of a CDS tool prior to integration into the EHR. This methodology can be applied to other studies aiming to decrease triggering rates and increase adoption rates of previously validated CDSS tools.

Introduction

Clinical decision support (CDS) tools

A physician's ability to determine a patient's risk of disease can be unclear and can make clinical decisions difficult. Therefore, CDS tools can help providers in their decision-making process. Clinical prediction rules (CPR) are a type of CDS that quantify the roles of history, physical examination and laboratory results in a patient's diagnosis, prognosis and likely response to treatment.[1] Clinicians are most likely to find CPRs useful when decision-making is complex, the clinical stakes are high or there are opportunities to achieve cost savings without compromising patient care.[1]

Despite the success at validating CPR and their tremendous potential to reduce unnecessary testing and treatments, there has been limited success with implementation and dissemination to providers, so CPR remain underutilised.[2] Therefore, clinical decision support systems (CDSSs) are used to help disseminate these tools to users and bring them to the forefront of patient care.[3] A CDSS is a computerised CDS, frequently integrated into a clinical information system to aid in clinical decision-making about individual patients. With the increased use of the electronic health record (EHR), an attempt to bring CDS and CPR tools to the point of care is ever more present. Efficiency, usefulness, information content, user interface and workflow have been reported by clinicians to be the keys to effective decision support.[4]

However, a common difficulty with CDSS is alert fatigue, when users begin to ignore or over-ride the triggered tool due to a high frequency of alerts.[5] The issue arises when the alerts relate to clinically insignificant events, so many alert fatigue studies have centred on drug–drug or drug allergy, since these alerts appear frequently[6] despite a desire by users not to receive such alerts.[7] The frequent appearance of drug–drug alerts does not correspond to the <1% of serious drug–drug interactions that occur in ambulatory care, so users rate the plethora of drug–drug alerts as not useful.[8] A study by Weingart *et al*[9] found that the over-ride rates of significant drug–drug interactions to be as high as 89.4% and 92.1% for drug-allergy over-rides. Over-riding does not mean these alerts are not useful and can be turned off.

Therefore, reducing the number of unnecessary alerts for a CDSS is a crucial but difficult task. Successful workflow integration depends on careful consideration of what pieces of data can launch or 'trigger' the CDSS tool. For example, in a prior study implementing a pneumonia CPR tool into an ambulatory primary care environment, four key triggering points were identified: symptom, encounter diagnosis, orders and diagnosis/order combinations.[10] This capacity to customise triggers to reflect real-world provider habits was a driver of the high adoption of the tool.[11]

CDS and pulmonary embolism (PE)

Emergency departments (EDs) across the nation are backed up with low-risk patients with PE waiting for unnecessary CT scans. A PE patient's morbidity and mortality can be improved by timely diagnosis, triage and treatment.[12] However, since PE is a condition with major repercussions and can be difficult to diagnose, providers often overestimate patient risk and order unnecessary tests exposing patients to radiation.[13] The ready availability of CT scan, with angiography (CTA), for PE has resulted in its increased use in the ED with an associated decrease in positive studies; 80–90% of PE workups are negative.[14]

The PE Wells rule for diagnosing PE is the most validated CPR in the field of thrombosis and it considers several criteria based on history and physical examination to estimate the patient's pretest probability of PE as low, moderate or high. The rule has been extensively validated in multiple settings[15–18] and has the potential to rule out 70–80% of patients without further testing.[19 20] However, the clinical criteria and branching logic of the Wells CPR makes it challenging to apply reliably at the point of care,[21] a challenge that EHRs are well positioned to address. CDS tools for PE have been created, yet there continues to be short falls in the integration of PE CDS into the EHR. Drescher *et al*[22] (2011) found that ED use of an electronic CPR for the evaluation of suspected PE was associated with an improved yield of positive CT scans. These findings emphasise the importance of implementation of the Wells criteria in a way which will gain maximum acceptance by the treating physicians.

Therefore, our research team created an efficient and easy-to-use CDS tool for the Wells criteria, following rounds of iterative usability testing[23 24] that would integrate seamlessly into the ED environment. The first and most important question that arose was at what point in the physician's workflow the tool should be 'triggered', where the tool will pop up on the users browser and alert them to a patient's risk.[25] In order to identify the ideal triggering mechanism we developed a new method for assessing accuracy of trigger rates prior to incorporation of the tool, called the Sensitivity and Specificity Trigger Analysis (SSTA). This method entails performing a retrospective chart review to look at the possible symptoms and elements of personal history that would alert a physician, in our case to the possibility of a PE. Our goal was to find a symptom or an element of patient's history that would provide a high sensitivity of CTA ordering without triggering too often so as to cause 'trigger fatigue'. By varying the triggers in different areas of the workup, we were able to determine the triggering point that had both a high sensitivity (true positive rate) and a high specificity (true negative rates) for CTA orders.

Methods

A retrospective chart review was conducted to evaluate the utility of the selected triggers in the implementation of a PE clinical support tool in the ED. Using the Sunrise EHR system at NorthShore LIJ health system, all charts of patients who were admitted to the ED between the period of 1 January 2013 and 31 December 2013 were reviewed. Of the 73 044 charts reviewed, data on 14 553 patients were either missing or had showed invalid heart rate. This resulted in 58 491 records available for analysis.

The outcome of interest was the ordering of a CT angiogram. Trigger words or phrases in the symptom were: chest pain, cough, cough blood, haemoptysis, difficulty breathing, dyspnoea, shortness of breath (SOB) at the start of the symptom, fainted, fainting and syncope. Terms that were synonymous were combined, so that the occurrence of any of these words in the symptom was considered as a single trigger. These combination triggers were cough blood and haemoptysis; difficulty breathing, dyspnoea and SOB; and fainted, fainting and syncope. Haemoptysis and dyspnoea were also considered independently. Each element of the Wells criteria (heart rate ≥100 bpm, history of venous thromboembolism (VTE), diagnosis of cancer and immobilisation) was considered as a possible trigger point.

In order to evaluate the utility of each trigger, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated. Sensitivity is the proportion of patients with the trigger, among those patients with a CT angiogram ordered. Similarly, specificity is the proportion of patients without the trigger, among those patients who did not have a CT angiogram ordered. PPV is the proportion of patients with a CT angiogram ordered, among those patients with the trigger. NPV is the proportion of patients who did not have a CT angiogram ordered, among those patients without the trigger.

There were 58 491 charts eligible for analysis (as described above). Of these, 930 (1.6%) had a CT angiogram ordered, resulting in an incidence rate of 1.6%. The calculations of PPV and NPV given below apply only to a population with the same incidence rate. In a population with a different prevalence, PPV and NPV would have to be calculated using sensitivity, specificity, the prevalence rate of the population of interest and Bayes rule.

The receiver operating characteristic (ROC) curve was plotted for each trigger. The areas under these correlated ROC curves were then compared using the non-parametric approach of DeLong *et al*.[26] On finding a significant difference, comparisons of all possible pairs were carried out. For these comparisons a Bonferroni-like adjustment was used, such that $p<0.01$ was considered significant for all comparisons.

Results

The most common trigger word was chest pain, which occurred in 11.3% of visits. Haemoptysis was an uncommon trigger word with an occurrence rate of only 0.1%. The frequency and per cent of occurrence for individual and combined triggers and Wells criteria are summarised in table 1.

Table 1 Frequencies of individual triggers

	n (%)
Trigger word (words)	
Chest pain	6582 (11.3)
Cough	5012 (8.6)
Cough blood or haemoptysis	206 (0.4)
Haemoptysis	35 (0.1)
Presence of any of the following: difficulty breathing, dyspnoea, shortness of breath or starts with SOB	2299 (3.9)
Dyspnoea	34 (0.1)
Presence of any of the following: fainted, fainting or syncope	724 (1.2)
Wells criteria	
Heart rate ≥100 bpm	15 995 (27.3)
History of VTE or PE	889 (1.5)
Diagnosis of cancer	2281 (3.9)
Immobilised	46 (0.1)
Combination triggers	
One or more triggers present	12 549 (21.5)
One of more Wells criteria present	18 404 (31.5)
One or more triggers or one or more Wells criteria present	26 171 (44.7)
One or more triggers and one or more Wells criteria present	4782 (8.2)

PE, pulmonary embolism; SOB, shortness of breath.

The results of the study indicated that the trigger word that was most sensitive was 'chest pain' which resulted in a sensitivity of 44.8% (table 3). Similarly, chest pain had a high NPV of 99%. In contrast, the symptoms of 'difficulty breathing, dyspnoea and SOB' resulted in only 21.5% sensitivity to a physician ordering of a CTA. The highest specificity trigger was 'difficulty breathing, dyspnoea and SOB' followed by 'cough blood or haemoptysis', resulting in a sensitivity of 99% and 96%, respectively. The highest PPV was with the trigger 'difficulty breathing, dyspnoea and SOB', 8.7%.

Chest pain revealed good discriminatory characteristics with an area under the curve (AUC) of 0.67 and was higher than that for any other individual triggers ($p<0.0001$ for all comparisons). The sensitivity, specificity, NPV, PPV and AUC for each possible trigger points are listed in table 2.

Further analysis combined the different symptoms and Wells criteria to identify a more sensitive way of triggering the Wells CPR tool. The combination of triggers that displayed the highest sensitivity was 'one or more triggers or one or more Wells criteria present'. Owing to the low prevalence of CTA ordering the PPV remained low. The specificity of the trigger was 55.7%. The AUC results showed a significant difference among the four combination triggers ($p<0.0001$). The combination triggers 'one or more trigger or one or more Wells criteria' and 'one or more triggers present', showed good discrimination characteristics. The results of the combined triggers are displayed in table 3.

Discussion

Emergency room physicians have a tendency to over order CT scans to rule out PE, costing millions in unnecessary diagnostic tests and exposing the patient to unnecessary radiation. One of the ways physicians can combat this problem is by incorporating a CPR into their everyday practice. By implementing a CPR the physician can bring evidence-based medicine to the point of patient care. This is especially relevant given that other, less costly tests, without the risk of radiation exposure, are also available. The D-dimer test is an example of a test that has a high NPV and can be used in those patients who are at intermediate risk. By inserting a well-validated tool such as the Wells score criteria into the EHR in a way that is consistent with ED workflow we can help to decrease the level of CT-angiograms ordered. However, as previous studies have demonstrated, in order to do this one must place the trigger in an optimal location so as to trigger enough to alert the physicians but not overtrigger to cause alert 'fatigue'.

Table 2 Results based on individual triggers

	Sensitivity (%)	Specificity (%)	Positive predictive value (PPV), %	Negative predictive value (NPV), %	Area under the curve (AUC) (95% Wald CI)
Trigger word (words)					
Chest pain	44.8	89.3	6.3	99.0	0.671 (0.655 to 0.687)
Cough	9.7	91.4	1.8	98.4	0.506 (0.496 to 0.515)
Cough blood or haemoptysis	1.1	99.7	4.9	98.4	0.504 (0.500 to 0.507)
Presence of any of the following: difficulty breathing, dyspnoea, shortness of breath or starts with SOB	21.5	96.4	8.7	98.7	0.589 (0.576 to 0.603)
Presence of any of the following: fainted, fainting or syncope	1.6	98.8	2.1	98.4	0.502 (0.498 to 0.506)
Wells criteria					
Heart rate ≥100 bpm	21.5	72.6	1.3	98.3	0.530 (0.516 to 0.543)
History of VTE or PE	6.5	98.6	6.7	98.5	0.525 (0.517 to 0.533)
Diagnosis of cancer	8.4	96.2	3.4	98.5	0.523 (0.514 to 0.532)
Immobilised	0.2	99.9	4.3	98.4	0.501 (0.499 to 0.502)

PE, pulmonary embolism; SOB, shortness of breath.

Table 3 Results based on a combination of triggers

Combination triggers	Sensitivity (%)	Specificity (%)	Positive predictive value (PPV), %	Negative predictive value (NPV), %	Area under the curve (AUC) (95% Wald CI)
One or more triggers present	62.9	79.2	4.7	99.2	0.711 (0.695 to 0.726)
One of more Wells criteria present	31.3	68.5	1.6	98.4	0.501 (0.486 to 0.516)
One or more triggers or one or more Wells criteria present	72.6	55.7	2.6	99.2	0.641 (0.627 to 0.656)
One or more triggers and one or more Wells criteria present	21.6	92.0	4.2	98.6	0.568 (0.555 to 0.582)

Through this study, we were able to explore a new methodology to identify appropriate triggering points of CDSS tools. The new usability tool, named the SSTA allows for CDSS tools to trigger only in alignment with physician clinical reasoning.

In order to reduce CTA ordering one must capture all of the cases where a physician is likely to order a CTA. By using the SSTA, we found the most sensitive way to identify when the CPR should be triggered was by using one or more of the symptom trigger words or one or more elements of the Wells criteria. Using this trigger point we will be able to identify 72.6% of cases, which would result in a CTA ordering, while maintaining a specificity of 55.7%. Therefore, the tool will trigger when necessary. The AUC of 0.64 revealed moderate discriminatory characteristics. Similarly, the triggering combination of one or more triggers present displayed a high specificity to CTA ordering and good discriminatory characteristics. However, since the sensitivity was lower we believe the more efficient way to identify those with a CTA order is by using 'one or more of the symptom trigger words or one or more elements of the Wells criteria'.

Conclusion

This study introduces a new usability methodology, SSTA, which provides an innovative way to studying triggering points for CDS tools and can be applied to other studies integrating previously validated CDS tools into the EHR. This is especially relevant when the alerts are rare but life threatening, such as the case in PE or drug–drug interactions. The next step of this study includes a roll out of the CDS tool in the ED of a tertiary hospital to measure the rates of compliance. We hypothesise using these specific trigger points can refine the triggering rates and possibly impact the adoption rates. Analysing trigger points can help hospitals build more user-friendly CDS tools with less alert fatigue.

Acknowledgements The authors would like to thank the US Agency for Healthcare Research and Quality for providing us with a grant to conduct this research (1 R24 HS022061). The abstract from this work was presented as a poster at AcademyHealth Concordium 2015 on 21–22 September 2015 in Washington, DC.

Competing interests None declared.

Ethics approval The procedures used were reviewed and approved as being in compliance with ethical standards of the responsible institutional review committee at the authors' institution.

Provenance and peer review Not commissioned; externally peer reviewed.

References

1. **McGinn TG**, Guyatt GH, Wyer PC, *et al*. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA* 2000;**284**:79–84.
2. **Elwyn G**, Légaré F, van der Weijden T, *et al*. Arduous implementation: does the Normalisation Process Model explain why it's so difficult to embed decision support technologies for patients in routine clinical practice. *Implement Sci* 2008;**3**:57.
3. **Eslami S**, de Keizer NF, Dongelmans DA, *et al*. Effects of two different levels of computerized decision support on blood glucose regulation in critically ill patients. *Int J Med Inform* 2012;**81**:53–60.
4. **Bright TJ**, Wong A, Dhurjati R, *et al*. Effect of clinical decision-support systems: a systematic review. *Ann Intern Med* 2012;**157**:29–43.
5. **Ash JS**, Sittig DF, Campbell EM, *et al*. Some unintended consequences of clinical decision support systems. *AMIA Annu Symp Proc* 2007;**2007**:26–30.
6. **van der Sijs H**, Aarts J, van Gelder T, *et al*. Turning off frequently overridden drug alerts: limited opportunities for doing it safely. *J Am Med Inform Assoc* 2008;**15**:439–48.
7. **Feldstein A**, Simon SR, Schneider J, *et al*. How to design computerized alerts to safe prescribing practices. *Jt Comm J Qual Saf* 2004;**30**:602–13.
8. **Peng CC**, Glassman PA, Marks IR, *et al*. Retrospective drug utilization review: incidence of clinically relevant potential drug-drug interactions in a large ambulatory population. *J Manag Care Pharm* 2003;**9**:513–22.
9. **Weingart SN**, Toth M, Sands DZ, *et al*. Physicians' decisions to override computerized drug alerts in primary care. *Arch Intern Med* 2003;**163**:2625–31.
10. **Mann DM**, Kannry JL, Edonyabo D, *et al*. 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.
11. **McGinn TG**, McCullagh L, Kannry J, *et al*. Efficacy of an evidence-based clinical decision support in primary care practices: a randomized clinical trial. *JAMA Intern Med* 2013;**173**:1584–91.
12. **Aronsky D**, Jones I, Raines B, *et al*. An integrated computerized triage system in the emergency department. *AMIA Annu Symp Proc* 2008;**2008**:16–20.
13. **Moser KM**. Diagnosing pulmonary embolism. *BMJ* 1994;**309**:1525–6.
14. **Weir ID**, Drescher F, Cousin D, *et al*. Trends in use and yield of chest computed tomography with angiography for diagnosis of pulmonary embolism in a Connecticut hospital emergency department. *Conn Med* 2010;**74**:5–9.
15. **Wells PS**, Owen C, Doucette S, *et al*. Does this patient have deep vein thrombosis? *JAMA* 2006;**295**:199–207.
16. **Wells PS**, Hirsh J, Anderson DR, *et al*. Accuracy of clinical assessment of deep-vein thrombosis. *Lancet* 1995;**345**:1326–30.
17. **Kearon C**, Ginsberg JS, Douketis J, *et al*. Management of suspected deep venous thrombosis in outpatients by using clinical assessment and D-dimer testing. *Ann Intern Med* 2001;**135**:108–11.
18. **Wells PS**, Anderson DR, Bormanis J, *et al*. Value of assessment of pretest probability of deep-vein thrombosis in clinical management. *Lancet* 1997;**350**:1795–8.
19. **Kyrle PA**, Eichinger S. Deep vein thrombosis. *Lancet* 2005;**365**:1163–74.
20. **Douketis JD**. Use of a clinical prediction score in patients with suspected deep venous thrombosis: two steps forward, one step back? *Ann Intern Med* 2005;**143**:140–2.
21. **Bates SM**, Jaeschke R, Stevens SM, *et al*. Diagnosis of DVT: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* 2012;**141**(Suppl 2):e351S–418S.
22. **Drescher FS**, Chandrika S, Weir ID, *et al*. Effectiveness and acceptability of a computerized decision support system using modified Wells criteria for evaluation of suspected pulmonary embolism. *Ann Emerg Med* 2011;**57**:613–21.

23. **Press A**, McCullagh L, Khan S, *et al*. Usability testing of a complex clinical decision support tool in the emergency department: lessons learned. *JMIR Hum Factors* 2015;**2**:e14.
24. **Khan S**, McCullagh L, Press A, *et al*. Formative assessment and design of a complex clinical decision support tool for pulmonary embolism. *Evid Based Med* 2016;**21**:7–13.
25. **McGinn T**. Putting meaning into meaningful use: a roadmap to successful integration of evidence at the point of care. *JMIR Med Inform* 2016;**4**:e16.
26. **DeLong ER**, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;**44**:837–45.