

6-2016

Clinical Decision Support Malfunctions

Steven Z. Kassakian

Follow this and additional works at: <http://digitalcommons.ohsu.edu/etd>

Recommended Citation

Kassakian, Steven Z., "Clinical Decision Support Malfunctions" (2016). *Scholar Archive*. 3783.
<http://digitalcommons.ohsu.edu/etd/3783>

This Thesis is brought to you for free and open access by OHSU Digital Commons. It has been accepted for inclusion in Scholar Archive by an authorized administrator of OHSU Digital Commons. For more information, please contact champieu@ohsu.edu.

Clinical Decision Support Malfunctions

By

Steven Z. Kassakian, MD, FACP

A Thesis

Presented to Department of Medical Informatics and Clinical Epidemiology

School of Medicine

Oregon Health and Science University

In partial fulfillment of the requirements of

Master of Science

June 2016

School of Medicine
Oregon Health & Science University

Certificate of Approval

This is to certify that the Master's Thesis of

Steven Z. Kassakian

“Clinical Decision Support Malfunctions”

Has been approved

David A. Dorr

Paul N. Gorman

Thomas R. Yackel

TABLE OF CONTENTS

Title Page	i
Certificate of Approval	ii
Table of Contents	iii
Acknowledgments	iv
Abstract	v
Background and Significance	1
Materials and Methods	3
Results	8
Discussion	15
Conclusions	20
References	21
Appendices	23
Tables	26
Figures	28

Acknowledgements

This work would not have been possible without the help of many individuals. Firstly, I am indebted to my thesis advisor Dr. David Dorr for his guidance and support. Additionally, Drs. Gorman and Yackel have both provided valuable insight and mentorship.

Specifically, I would like to extend my gratitude to Diana Pennington, Denise Sandell and Neil Edillo who provided countless hours of effort investigating various clinical decision support rules.

Dr. Adam Wright's previous work in the field and current efforts on the topic were critical in laying the groundwork for the performance of this research. Furthermore, his collaboration and guidance on this effort is very much appreciated.

I would also like to extend my appreciation to Mrs. Jennifer Anderson and Dr. James Morrison for their critical reading and comments on this thesis itself.

To all of the faculty, staff, and students at Department of Medical Informatics and Clinical Epidemiology at OHSU I am indebted for all the support and assistance they have provided. In particular I would like to thank Virginia Lankes for her steadfast support, keen guidance, and most importantly her friendship. Also, I need to express my deep appreciation to Dr. Alfredo Almerares who has at times been my harshest critic but always a steadfast friend and supporter.

Finally, I wish to express my sincere gratitude to Dr. William Hersh for providing me the opportunity to study clinical informatics at OHSU.

Research reported in this publication was supported by the National Library of Medicine of the National Institutes of Health under Award Number T15LM007088. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Abstract

Objective: To determine if clinical decision support (CDS) malfunctions occur in a commercial electronic health record (EHR) system, characterize their pathways and explore methods of detection.

Materials and Methods: We retrospectively examined the firing rate for 226 alert type CDS rules for detection of anomalies using both expert visualization and statistical process control (SPC) methods over a five year period. Candidate anomalies were investigated and validated. We defined an anomaly as alert firing rates which deviated from historical or expected trends. A CDS malfunction was defined as a CDS alert which was not functioning as designed or intended. A false positive CDS anomaly was an alert that altered firing from an intentional change.

Results: In 8,300 alert-months twenty-one candidate CDS anomalies were identified. Of these candidate anomalies, four were confirmed as CDS malfunctions, eight as false-positives, and nine could not be classified. The four CDS malfunctions were a result of errors in knowledge management: (1) inadvertent addition and removal of a medication code to the electronic formulary list; (2) a seasonal alert which was not activated; (3) a change in the base data structures; (4) direct editing of an alert related to its medications. 154 CDS rules were amenable to SPC methods and the test characteristics were calculated as a sensitivity of 95%, positive predictive value of 29% and F-measure 0.44.

Discussion: CDS malfunctions were found to occur in our EHR. All of the pathways for these malfunctions can be described as knowledge management errors. Expert visualization is a robust method of detection, but is resource intensive. SPC-based methods, when applicable, perform reasonably well retrospectively.

Conclusion: CDS anomalies were found to occur in a commercial EHR and visual detection along with SPC analysis represents promising methods of malfunction detection.

Background and Significance

Clinical Decision Support

Clinical decision support (CDS) tools are focused on improving medical decision making. A large body of evidence supports the effectiveness of these tools to improve process outcomes and reduce errors.¹⁻⁵

Overall, CDS is considered an essential part of realizing the potential benefits of health information technology.

CDS tools can take many forms including alerts, reminders, drug-dose calculations, information retrieval tools, knowledge management resources, or any targeted information intended to facilitate decision making. Many of these tools are found within the context of the electronic health record (EHR).

Taxonomy to describe CDS tools was recently developed and wide variation in use of types of CDS tools in commercial EHR systems was found.^{6,7} While there is ample evidence that individual CDS tools can improve outcomes, the best methods for organizations to implement and curate their suite of CDS tools is a less well defined and more complicated endeavor. Knowledge management has been identified as one of the key features of a successful CDS program.⁸ Building on this work, a qualitative study of leading CDS-utilizing organizations stressed the importance of developing tools to provide ongoing monitoring of CDS.⁹

Real-world use and effects of CDS tools present a far more complicated picture.¹⁰ While CDS related to medication prescribing has been shown to reduce errors, there is substantial evidence that the level of overriding of CDS is high, from 40-96%.¹¹⁻¹⁴ Overriding is generally defined as when the provider intentionally chooses not to follow the CDS guidance. This high level of overriding appears to be tightly coupled with the phenomenon known as alert fatigue, whereby, via increasing exposure to CDS alerts, provider responsiveness to them rapidly declines.¹⁵⁻¹⁷ The problem of alert fatigue is believed to arise

both from the sheer volume of alerts presented to the user as well as the high rate of false positive alerts presented.¹⁸ Alert fatigue is unlikely to decrease as more and more care processes, quality initiatives, and compliance-related issues are being “hard-wired” into the EHR via CDS. Given these concerns it is imperative that significant effort is made to optimize and curate the CDS tools so as not to contribute further to the phenomena of alert fatigue with poorly functioning CDS.

In addition to alert fatigue, several other safety concerns surrounding CDS have been uncovered and remain unresolved.^{15,19} recently, an emerging concern regarding the malfunctioning of CDS systems has been described. A malfunctioning CDS system is best described as when a CDS system “...does not function as it was designed or expected to”.²⁰ Work by Wright et al. described a small case series of four CDS malfunctions in a home-grown EHR system in which CDS malfunctions occurred secondary to a change in a laboratory test code, a drug dictionary change, inadvertent alteration of the underlying alert logic, and a software coding error in the underlying system.²⁰ Additionally, these authors carried out a survey asking Chief Medical Information Officers (CMIOs) whether similar types of malfunctions of CDS malfunctions had occurred in their systems and some 27 out of 29 CMIOs responded affirmatively. To date there has been no comprehensive analysis of CDS malfunctions within any other EHR installations and none in a commercial system, which are the predominant type in use in the U.S.

When a CDS tool malfunctions there is rarely a mechanism in place to detect the malfunction. Rather, ad-hoc user reports might uncover an issue or an administrator might retrospectively review and manually analyze the firing and response rates.²¹ Given that many organizations have hundreds or thousands of different pieces of CDS this review is all but impossible to comprehensively conduct manually. As well, the current tools in most commercial EHRs provide limited functionality to even conduct a manual review. Fortunately, the activity of CDS alerts is a heavily audited process within most EHRs and therefore the occurrence of each alert is available for review. The most defining measure of

CDS tool activity is the firing rate, i.e. how often was the rule triggered, and thus the data related to alerts is a time-series pattern. All else being equal one would expect a fairly constant firing rate or some predictable variation owing to the timing of workflow related processes. For instance, CDS alerts in the ambulatory or surgical settings might fire less on the weekends, while a sub-specialty clinic might only operate on a given weekday.

Objectives

Given the ongoing concerns about effective use of CDS and the emerging concerns regarding CDS malfunctions we set out to evaluate and characterize CDS malfunctions in a commercial EHR.^{8,20,22}

Firstly, we wanted to determine whether CDS malfunctions are occurring in our instance of a commonly utilized commercial EHR. Secondly, if CDS malfunctions are found to occur we sought to characterize the pathways through which these malfunctions happen. Thirdly, we want to describe methods for both detection and prevention of CDS malfunctions.

Materials and Methods

Overview

A retrospective observational cohort study was conducted on selected CDS alerts from 1 January 2010 through 30 July 2015. We analyzed the firing history with both a visualization method and a statistically based method. Based on this analysis alerts were characterized as either normal or CDS anomalies. All CDS anomalies were manually validated to determine whether they were true CDS malfunctions or false positives. True CDS malfunctions were further investigated to elucidate the pathway through which the error occurred. Finally, we compared the performance of the anomaly detection methods and describe methods for both CDS malfunction prevention and detection.

Study Setting and Electronic Health Record System

This study utilized data from Oregon Health and Science University (OHSU), a 576 bed tertiary care facility in Portland, OR. The EHR in use at OHSU is EPIC (EPIC Systems, Verona, WI).

Clinical Decision Support Tools

While a multitude of different types of CDS tools have been described and are available to varying degrees in different EHR vendor systems, this study focuses on the use of what have been referred to in the literature as point of care alerts/reminders.⁷ These CDS tools generally fall within the category of rules wherein a pre-specified logical criteria is created and expects a specific action or set of actions to be fulfilled.²³ Within the EHR in use at OHSU, this alert type is known as Best Practice Advisories. However, it is important to note that given the multitude of ways in which an organization can customize these tools they are not always strictly alerts/reminders as described in the literature but can function without even being shown to the user to facilitate other processes.

We have chosen to focus on these alerts since the development and knowledge management of these alerts is done locally, whereas many of the other CDS tools are either inherent to the system (i.e. ordering duplication checking,) or obtained from third-party vendors (Rx-Rx interaction checking). Therefore the knowledge elucidated from this study will likewise have greater external validity. This set of CDS tools was further restricted to alerts that were still active in our system and were actively shown to the user.

Data Abstraction

The EHR utilizes a structured query language (SQL) relational database model which functions as a data warehouse for use in analytics. Utilizing Oracle SQL Developer (Oracle Corporation, Redwood Shores, CA) a SQL script was created to retrieve the alert activity history from the relational database (Appendix

A). Once the initial data abstraction from the database was complete multiple steps for data refinement were undertaken as follows using STATA v 13.1 (STATA CORP, College Station, TX) (Appendix B). Firstly, the date of alert activity variable was transformed from a string to a long datatype following the specific date convention of STATA. Additionally, the dataset was restricted to CDS tools which were still active in the system at the time of extraction and were shown to the user. Additionally, one additional variable was created to signify the type of day, i.e. weekday and weekend.

Dataset Validation

To validate the dataset created above the following steps were performed. Using a reporting tool within the EHR itself a random sample of 5 alerts were chosen and reports of their alerting activity were created for the time period of the study. Monthly plots of the alerting activity were then created and comparison made to the visualizations created below.

Visual Anomaly Detection

Visual anomaly detection was performed using Tableau 9.2 (Tableau Software, Seattle, WA). Using the alert activation date as the column and the sum of the number of records as the row, the number of alerts per unit of time was created. A filter was added on the unique alert identifier as well as the variable which described the day type, i.e. weekday or weekend.

An expert (SZK), who serves as the co-chair of the clinical decision support committee at OHSU, visually inspected the alert firing history for each unique alert. When the alert firing appeared to deviate from historical patterns or exhibited behavior that appeared inconsistent with knowledge of the targeted activity the alert firing event was deemed to be a candidate visual anomaly. For example, alerts related to influenza were expected to exhibit a seasonality to their activity and thus if this result was

encountered via the visual inspection it would not have been considered an anomaly. As necessary, the visualizations were viewed on different time scales to examine their activity.

A second reviewer was utilized to validate the visualization method. The second reviewer, Dr. David Dorr, is board certified clinical informatics and formerly co-chaired the CDS committee at OHSU. Ten representative visualizations of unique alert activity, five considered anomalies and five considered normal by the initial reviewer, were shared and classified by DD. Inter-rater reliability was calculated.

Statistical Process Control Anomaly Detection

Given that our dataset comprised count data and the underlying denominator, or area of opportunity, likely varied insignificantly, statistical process control (SPC) c- charts were created.^{24,25} To create c-charts in Tableau a parameter, standard deviation, was created which consisted of an integer with range 1-3 in 0.5 step increments. Following this process, a measure was created which defined an 'average line', which was the numerical average of the observations for the respective time period. Additionally, an upper confidence limit and lower confidence limit line were created to encompass the average line + 3*standard deviation and average line – 3* standard deviation (Appendix C). The following tests were performed to detect the presence of special cause.²⁴ Test #1 the presence of a single point outside the control limits using 3*standard deviation. Test #2 two of three consecutive points are more than 2 standard deviations from the average line and both on the same side of the average line. Test #3 eight or more consecutive points on the same side of the average line. Test #4 consisted of 6 or more values steadily increasing or decreasing. SPC anomaly detection was attempted on time points for both a weekly and monthly scale.

To determine the characteristics and performance of SPC detection methods sensitivity, specific, precision and the F measure were determined. For the purposes of test performance characteristics, since there is no established gold standard and resources precluded validation of the entire underlying

CDS cohort, we first reduced the dataset to those CDS rules where SPC was able to be utilized, i.e. there was some prior history of control for a sufficient period of time. Then with this reduced dataset we treated all detected visual anomalies, as determined by reviewer S.Z.K, as the true positive and the non-candidate CDS rules as true negatives.

Candidate Anomaly Validation

A true positive anomaly (aka CDS malfunction) occurs when the CDS rule "...does not function as it was designed or expected to".²⁰ For example, should an anomaly detection method find that a specific alert's firing rate decreased and it is then determined that this occurred because the target population is seen less frequently in the respective setting this alert would be considered a false positive. In contrast, if a candidate anomaly is identified because the firing rate decreased significantly and this was found to be secondary to a change in a laboratory test code which is part of the CDS tool logic this would be considered a true positive (aka CDS malfunction). In essence, if the CDS should have kept firing because the same situation was occurring and the same alert should still fire in that situation, then it was a true positive or malfunction.

Candidate anomaly validations were conducted in the following manner. Firstly, the alert build records were searched to determine the original date of creation and whether the records had any history of editing, as demonstrated in the time stamp data. Linked records were examined to ensure they remained released in the system. For each candidate anomaly an informal discussion with the local CDS analyst regarding the findings took place. Following this discussion, further discussion with other EHR analysts responsible for various parts of the EHR build occurred. Additionally, the CDS analyst work logs, when available, were searched to determine if notes regarding the build and subsequent alterations to the tool were available. Our institution-wide EHR change notice system was searched for related entries coinciding with changes in alert activity. As needed, we discussed alert activity with relevant clinical

users and departments to examine for possible competing changes which would have affected alert firing rates. For any alert involving medication records, extensive discussion with pharmacy informatics colleagues occurred.

Institutional Review Board

Institutional review board (IRB) approval for this study with a waiver of informed consent was obtained.

Results

We had a total of 8,300 alert months comprised of 226 alert type CDS rules which were shown to the user and still active in the system. These CDS rules formed the cohort used in this analysis. Of these 226 rules, 21 were considered visual anomalies by the first CDS reviewer. Of the 21 visual anomalies 4 were considered CDS malfunctions (aka true positives), 8 were false positives (i.e. expected changes in alerting) and 9 were unable to be classified (Table 1). Of the 226 alert type CDS rules, 154 were amenable to the SPC detection method. The remaining rules were not amenable as they did not meet the assumption of control required for SPC detection. All four CDS malfunctions pathways were considered to be the result of knowledge management processes (Table 2).

Anomaly Visualization Reviewer Agreement

A random sample of five candidate visual anomalies visualizations and five non-candidate visualizations, as classified by reviewer SZK, were shared with a second reviewer, DD, who was blinded to the first reviewer determination. The second reviewer independently assessed the alert firing visualizations to detect the presence of visual anomalies. There was 100% agreement in terms of classifying the visualizations as either anomalies or non-anomalies.

CDS Malfunction #1: Use of enoxaparin order set

Based on both visualization and SPC c-chart methods an anomaly was detected (Figure 1). The visualization method generated a likely anomaly given that the monthly alerting rate consistently ranged in the 60-100s for several years and then subsequently decreased to less than 10 firings per month starting in May of 2015. With regard to SPC detection, tests #2 and #3 were violated.

Ensuring the proper use of anti-coagulation is a major patient safety concern. As part of an institutional quality improvement process all orders for enoxaparin needed to utilize an order set to ensure compliance with a regulatory requirement denoting which provider was managing the anti-coagulation. To ensure users utilized the order set an interruptive alert was created which was triggered when an order for enoxaparin was entered and the patient did not have an accompanying order for anti-coagulation management.

Following detection of a candidate anomaly and discussing it with the pharmacy informatics they immediately described the result as "...actually an error on our part." (N. Edillo, Feb. 2016). In initially developing the alert a dummy medication record for enoxaparin was created and placed in the EHR formulary list available to users. This dummy record would redirect users to the order set and was the one specified in the alert criteria which would cause the alert to fire. However, the actual medication record still needed to exist to ensure proper functioning of a multitude of pharmacy processes. In July of 2015 during routine pharmacy Rx list maintenance a new formulary medication list was created which included the actual enoxaparin medication record instead of the dummy order. Thus, when providers now searched for enoxaparin they found the actual order and not the dummy. As the actual Rx order was not included in the alert logic, the alert was never triggered.

CDS Malfunction #2: Administration of Flu vaccine at discharge

Based on visualization a candidate anomaly was detected (Figure 2). SPC methods could not be applied since the underlying characteristics of the alert firing do not allow its use. The visualization method generated an anomaly given that a clear seasonal pattern was observed in influenza seasons 2010-2011, 2012-2013, and 2013-2014 but absent in both 2011-2012 and 2014-2015.

This alert is triggered when a patient has an active order for influenza vaccination, which has not yet been administered, and then receives an order for discharge. Thus, the alert is trying to prevent failure to administer the vaccination before discharge. The alert logic contains a specific medication record group that specifies which influenza vaccinations are appropriate for the current season.

Following the discovery of this candidate anomaly further activity of the alert was obtained as this dataset stopped at 30 July 2015. An internal EHR report was created using native EHR tools to examine the more recent activity of the alert. It was found that the alert started firing again on 15 September 2015 after its last period of activity on 29 April 2014. Following discussion with the CDS analyst it was determined that following influenza season the rule is de-activated, i.e. removed from production, along with all the other influenza CDS rules. While this seems counterintuitive, had the rule not been de-released and allowed to remain as-is in the production system it would have never fired again since the medication record criteria on which it fires requires update each flu season. The seasonal updating is required as the influenza vaccinations are adjusted to the most prevalent predicted sero-types each with new medication identifiers.

CDS Malfunction #3: No documented height in oncology patients

Based on visualization and SPC c-chart a candidate anomaly was detected (Figure 3). For the visualization method the alert activity was noted to go to zero after March 2014. With regards to the SPC c-chart detection method the alert activity violated test #1, with values more than three standard deviations away from the control line.

This alert was created to ensure that patients in the hematologic malignancy clinic had recently documented height measurements in the EHR. This is important since many chemotherapeutic medications are dosed based on body surface area which requires a height to calculate and therefore appropriately dose.

Following the discovery of this candidate anomaly the alert criteria was critically examined and it was noted that the department in which this alert was targeted was no longer active. Within the EHR, departments function not so much as virtual representations of physical space, but more as scheduling and billing entities and are in relatively frequent flux. Given this provenance they can be created or deactivated for administrative purposes while no perceptible changes occur with the clinical work or processes. On review it was noted that the deactivation of this department name occurred on 7 March 2014 and a notice was placed in the local EPIC change notice system. However, the necessity to recognize whether a change in this attribute will affect another given area in the EHR is generally left up to the analysts and therefore can be missed.

CDS Malfunction #4: Coronary artery disease and use of anti-platelet medications

Based on visualization of the alert firing history and use of SPC chart detection (Tests 1, 2 and 3) an anomaly was discovered. In visual terms the alert firing was historically occurring approximately 200

times per month. Around October of 2010 an increase of at least two-fold was seen in the alert firing rate (Figure 4).

This alert was created to ensure that patients with a diagnosis of coronary artery disease (CAD) were also prescribed an anti-platelet medication, as supported by strong evidence.²⁶ When this alert was originally created the classes of medication which satisfied or suppressed the alert were those in the type of anti-platelet Rxs as consistent with the guidelines.

During the candidate anomaly validation work it was determined that a change in rule was made on 10 October 2010, which is the exact day an increase in the alert firing rate was detected. This change included the expanding of the target medications which would satisfy, i.e. suppress, this alert from activating. Previously, this alert was suppressed when a patient with a diagnosis of CAD and an Rx for an antiplatelet medication. However, for unclear reasons the alerts criteria was expanded to include medications in the class of anticoagulants in addition to antiplatelet medications. Now with these additional medications in place a larger group of medication would suppress the alert. However, what was immediately seen was that the rate of alert firing increased. In discussion with the pharmacy informatics colleagues involved in these changes there was no clear understanding why this increase occurred. Additionally, the alert firing rate subsequently decreased to slightly below its historical level, again with no known explanation.

False Positive CDS anomalies

There were a total of eight false positive CDS anomalies identified (Table 1). To further illustrate their character several are described.

An alert to inform users regarding a shortage of the medication propofol was implemented. Via visual detection and SPC methods (Tests 1, 2 and 3) an anomaly was detected as the alert stopped firing after 29 January 2014 (Figure 5). Once the propofol shortage ceased, rather than remove the alert itself, the triggering orders were revised to stop the firing of this alert since this approach results in more efficient maintenance for the pharmacy informatics group. The alert therefore would no longer fire but was available.

An alert specific to the Emergency Department (ED) users was implemented to inform the user that an EKG that had been performed which had no corresponding order, preventing reimbursement for the procedure. Based on visualization and Test #3 of SPC methods a candidate anomaly was identified when the rate of firing substantially declined in October of 2014 (Figure 6). After discussion with a clinical manager it was determined that the alerting activity declined secondary to the implementation of a nursing-initiated order protocol which now allowed RNs to place orders for the EKG upon patient assessment in triage.

An alert was implemented specific to ED RN users to ensure that documentation related to the end times for IV medications were entered to ensure regulatory compliance related to billing. Based on both visualization and SPC methods (Test #3) a candidate anomaly was detected (Figure 7). After discussion with the ED RN manager it was determined that several build changes in the EHR were implemented which corresponded in time to the changes in alerting activity seen.

Performance of SPC c-chart anomaly detection methods

SPC c-chart detection methods were applied to 19 out of 21 anomalies detected by visualization and were able to detect 18 of 19 anomalies (Table 1). Two candidate anomalies were not amenable to SPC

methods. As previously described, to calculate the test characteristics of SPC c-charts we first reduced the dataset to only those CDS rules in which the firing rates were amenable to SPC given that they had a sufficient number of historical control points. There were 154 CDS rules which we were able to analyze by SPC control. Following this screening process we then treated all anomalies detected by visual analysis as the set of true positives and the remaining set of CDS rules not considered anomalies as true negatives, i.e. we treated the detection of visual anomalies as the gold standard test. Additionally, we made this analysis using an aggregate measure for SPC c-chart detection with a positive test being one in which either Test #1, Test #2 or Test #3 was violated. Using these assumptions the sensitivity of SPC c-chart detection methods in our study was 0.95 (18/19). The precision or positive predictive value was 0.29 (18 / (18+44)). The F measure is 0.44.

Additional Findings from this review

Two candidate CDS anomalies which remained unclassified present interesting cases. While their classification remains unknown, their lack of functionality has resulted in duplicative efforts. In both cases after finding that these rules were functioning anomalously it was also subsequently determined that the targeted users groups had efforts underway to create new de-novo rules completely unaware of the already existing CDS rules.

In the first case, an existing alert recommends to providers that for patients on second generation anti-psychotics, they should check their creatinine, electrolytes, lipids and liver function tests with some regularity (i.e. every 3, 6 or 12 months). This alert was identified with visualization methods as a candidate visual anomaly as it fired at a rate of 60-100 times per month for about 2 years then decreased to only 10-20 times per month onward. When attempting to validate this candidate anomaly a discussion occurred with the clinical content and alert targeted user group (psychiatry). Insufficient

information regarding the cause of the change in alert firing activity was obtained and thus the candidate anomaly was considered unclassified. However, two important findings were uncovered in the validation attempt. Firstly, the departmental informatics contact person had no knowledge of this alert. Secondly, they reported that they were in the process of trying to implement a CDS tool which would alert users that patients who were taking second generation anti-psychotics should have their lipids checked regularly - something already covered in the older alert.

The second alert maintenance example is related to an alert telling users that the patient they are admitting is a member of a large HMO and requires admission to a specific hospitalist service. An alert for this situation was implemented in 2010. This alert was identified as a candidate anomaly since it fired about 60-80x/month and then ceased firing in September of 2012. The validation process was unable to uncover any specific reason that this alert ceased functioning and therefore it was categorized as unclassified. However, it was determined that an extremely similar alert was implemented in late 2015 to accomplish the same task (i.e. alert providers admitting patients that a particular patient was a member of this particular HMO and required admission to a specific service). The methods used to identify the patients in each instance were different, as the former alert uses the insurer associated with the patient and latter alert uses the patient's PCP and subsequently whether that particular PCP is employed by the HMO. In discussing the new alert with the CDS analyst she relayed that when building the newer alert there was no knowledge of the existence of the old alert.

Discussion

In 8,300 alert months we were able to find and validate four individual CDS malfunctions from a subset of 226 CDS rules in our EHR. Additionally, we identified nine CDS rules as anomalies that remain

unclassified. It is entirely possible, if not likely, that several of these rules represent additional CDS malfunctions. Prior to this study none of the identified CDS malfunctions found had been previously identified by either our analysts or users. To our knowledge this work represents the first examination of CDS malfunctions in a commercial EHR. Recent work by Wright et al. found four individual instances of CDS malfunctions in a commercial EHR. Recent work by Wright et al. found four individual instances of CDS malfunctions in a home-grown EHR system out of some 201 examined (A. Wright, May 2016).²⁰ Furthermore, via a survey of CMIOs, they found that these types of errors were possibly much more widespread.

We collectively referred to the patterns of malfunctions found in our CDS library as knowledge management errors. All of these errors involved active alternation to some aspect of the EHR system except in one case where a lack of action occurred (i.e. Influenza rule not activated, resulting in a malfunction). The first malfunction uncovered was related to use of the medication enoxaparin and occurred when routine maintenance on the electronic medication formulary list was being performed and a medication code was inadvertently switched. The second malfunction occurred when the CDS analyst inadvertently did not activate a CDS rule to coincide with the beginning of influenza season. This failure to activate occurred on two separate occasions, resulting in the non-functioning of this rule over two separate influenza seasons. The third CDS malfunction occurred when the targeted department of the rule was discontinued. Finally, the fourth occurred when a rule was directly edited to add an additional class of medications which would cause it to not fire.

Following these findings the two main follow-up questions focus on detection and prevention. With regards to prevention, it would be overly simplistic to suggest that the CDS rules need to be tested any time a change in the system is made. While this is clearly prudent when an analyst make a direct edit to a CDS rule, as was the case prior to the fourth CDS malfunction, in cases where changes to other parts of the system are made this is likely not feasible given the frequency of alterations. Particularly germane

to this pathway is the CDS malfunction which occurred as the result of the discontinuation of a department. We believe this type of malfunction could have occurred through a change in any number of attributes which are used to target the CDS to specific provider or patient populations. In the EHR, departments are virtual entities that are created mostly to enable the billing and scheduling process. Departments are created and discontinued with some frequency and these can occur with virtually no perceptible changes in the physical world. Our institution has a notification system which is utilized to alert analysts when changes are made in our EPIC build which may affect other areas. However, this system requires manual review by the analyst to see if any changes will affect their domain. We have over 200 notifications per month in our system. This problem essentially arises because many of these local attributes are essentially hard coded in multiple areas of the EHR build.

The malfunction related to inadvertently not activating the CDS rule related to influenza is clearly a knowledge management issue. As it currently stands in our institution, analysts are essentially left up to the task of remembering which rules requires manual activation and deactivation on a seasonal basis. In this particular case the analyst normally responsible for the deactivation and reactivation of these alerts was out on leave in both instances when it was not activated. As ironic as it might be, an improved knowledge management system which can track and remind the analysts regarding these types of required changes would likely prevent this type of CDS malfunction.

While we identified four CDS malfunctions, the fact that an additional nine anomalies remain unclassified is concerning and demonstrates the significant resources required to validate and test CDS rules. In regards to the nine unclassified anomalies, after following all the methods of validation, it remains uncertain whether their detected change in alerting activity is expected or the result of a malfunction. To fully validate these remaining unclassified anomalies would require complete mapping and analysis of every attribute which defines the rule - something which is prohibitively resource

intensive. One of the most complicating parts of the validation is its retrospective nature. In many cases we are trying to look for a needle that was dropped into the haystack 3, 4 or even 5 years ago. We surmise that had these anomalies been uncovered in near real-time they would be much easier to validate. Adding to this difficulty is the fact that in our EHR system the records related to the CDS rules are directly edited, overwriting prior entries. There is a method available in this vendor system to create new records upon editing, which would preserve prior configurations. Implementation of this method would likely improve the ability to trace the root cause of malfunctions.

The CDS malfunction that occurred following routine maintenance on the EHR formulary list highlights the necessity for a prospective method of detection as a particular prevention method for this remains unclear. Additionally, in further support of a prospective method of detection there are undoubtedly other pathways of CDS malfunctions which we have yet to be elucidated. For detection methods, we chose to utilize both visualization and SPC c-charts as these methods have support for their use from the literature in other similar domains, as well as ease of application helping to promote generalizability. Visualization has been shown to be a very strong detection method in multiple domains.²⁷ While SPC c-chart detection could only be applied to 152 of the 226 CDS rules we believe this number could easily be improved if there was a prospective system in place. One of the limitations for using SPC c-charts retrospectively is that we did not have the knowledge of when external changes occurred which would cause an expected reset of the control process. Therefore, many CDS rules which were retrospectively not amenable to SPC based methods would be so if CDS reviewers marked times of expected changes.

Strength and Limitations

We believe one of the greatest strengths of this work is that it occurred in a vendor system, which is by far the predominant type of EHR currently in use in the U.S. Additionally, while this analysis reviews a

single vendor system, there is no particular reason to believe that these types of malfunctions would be limited to our instance of this system or even to this particular vendor. As well, all of the methods used and software tools which we utilized to carry out this work are readily available to any healthcare system, increasing the external validity of our work.

In order to utilize SPC c-chart methods we had to assume that the denominator of interest, i.e. the potential population of patients on whom the CDS was targeted, did not vary significantly.²⁴ We believe this assumption was both reasonable and necessary as calculating denominators for the target populations would have been extremely resource intensive as each alert has a unique set of provider, patient and departmental-level characteristics that define when the alert would be shown. Any attempt at analysis on this level would require a massive undertaking and would potentially introduce significant error with little clear benefit. It has been established the use of c-chart whereby only the event itself is measured is valid so long as the denominator or area of opportunity doesn't vary by more than 20%.²⁴

For the SPC c-charts analysis method we utilized one aggregate average and therefore created two aggregate control limits, an upper and a lower. Given that we had no preconceived idea when the change in process would occur, if at all, we were unable to create a control line prior to the change. One of the limitations of our study is that we were unable to validate the entire cohort, meaning there are potentially a number of false negatives, or CDS malfunctions that went undetected by our methods. SPC c-chart demonstrated reasonable performance with respect to retrospective detection of CDS anomalies. While we did not have a gold standard, we do think that the visualization method likely represents a strong method of detection and therefore serves as a reasonable proxy. Additionally, given that upon first constructing detection methods for these types of errors high sensitivity is desirable, and owing to the fact that the alerting behavior of false positive CDS anomalies is fairly identical in many

case to true CDS malfunctions, we feel that it was prudent to utilize the CDS anomalies identified by visualization as “true positives” for the purposes of SPC c-chart characteristics.

Conclusion

From a systematic examination of 8,300 CDS alert months we uncovered 21 CDS anomalies. Following validation four were determined to be CDS malfunctions, eight false positive and nine remain unclassified. All of the validated CDS anomalies appear to follow the pathway of what could be termed knowledge management errors. We did not find any errors which results from any intrinsic issues with the EHR system, issues with external system integration or third-party content. This likely represents important work as these anomalies are likely occurring in other installations of this vendor system and in other vendor systems as well. Furthermore, use of SPC c-chart analysis represents a promising method for prospective monitoring of CDS alert rules, augmented by manual review for those rules that were are not amenable to SPC.

References

1. Bright TJ, Wong A, Dhurjati R, et al. Effect of clinical decision-support systems: a systematic review. *Annals of internal medicine*. 2012;157(1):29-43.
2. Lobach D. *Enabling Health Care Decisionmaking Through Clinical Decision Support and Knowledge Management*. 2012.
3. McDonald CJ. Protocol-based computer reminders, the quality of care and the non-perfectability of man. *The New England journal of medicine*. 1976;295(24):1351-1355.
4. Moja L, Kwag KH, Lytras T, et al. Effectiveness of computerized decision support systems linked to electronic health records: a systematic review and meta-analysis. *Am J Public Health*. 2014;104(12):e12-22.
5. Roshanov PS, Fernandes N, Wilczynski JM, et al. Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials. *BMJ*. 2013;346:f657.
6. Wright A, Goldberg H, Hongsermeier T, Middleton B. A description and functional taxonomy of rule-based decision support content at a large integrated delivery network. *Journal of the American Medical Informatics Association : JAMIA*. 2007;14(4):489-496.
7. Wright A, Sittig DF, Ash JS, et al. Development and evaluation of a comprehensive clinical decision support taxonomy: comparison of front-end tools in commercial and internally developed electronic health record systems. *Journal of the American Medical Informatics Association : JAMIA*. 2011;18(3):232-242.
8. Ash JS, Sittig DF, Dykstra R, et al. Identifying best practices for clinical decision support and knowledge management in the field. *Studies in Health Technology & Informatics*. 2010;160(Pt 2):806-810.
9. Wright A, Sittig DF, Ash JS, et al. Governance for clinical decision support: case studies and recommended practices from leading institutions. *Journal of the American Medical Informatics Association*. 2011;18(2):187-194.
10. Carspecken CW, Sharek PJ, Longhurst C, Pageler NM. A clinical case of electronic health record drug alert fatigue: consequences for patient outcome. *Pediatrics*. 2013;131(6):e1970-1973.
11. Lin CP, Payne TH, Nichol WP, Hoey PJ, Anderson CL, Gennari JH. Evaluating clinical decision support systems: monitoring CPOE order check override rates in the Department of Veterans Affairs' Computerized Patient Record System. *Journal of the American Medical Informatics Association : JAMIA*. 2008;15(5):620-626.
12. Hsieh TC, Kuperman GJ, Jaggi T, et al. Characteristics and consequences of drug allergy alert overrides in a computerized physician order entry system. *Journal of the American Medical Informatics Association : JAMIA*. 2004;11(6):482-491.
13. van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. *Journal of the American Medical Informatics Association : JAMIA*. 2006;13(2):138-147.
14. Weingart SN, Toth M, Sands DZ, Aronson MD, Davis RB, Phillips RS. Physicians' decisions to override computerized drug alerts in primary care. *Archives of internal medicine*. 2003;163(21):2625-2631.
15. Ash JS, Sittig DF, Campbell EM, Guappone KP, Dykstra RH. Some unintended consequences of clinical decision support systems. *AMIA Annu Symp Proc*. 2007:26-30.
16. Shah NR, Seger AC, Seger DL, et al. Improving acceptance of computerized prescribing alerts in ambulatory care. *Journal of the American Medical Informatics Association : JAMIA*. 2006;13(1):5-11.

17. Embi PJ, Leonard AC. Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study. *Journal of the American Medical Informatics Association : JAMIA*. 2012;19(e1):e145-148.
18. Cash JJ. Alert fatigue. *Am J Health Syst Pharm*. 2009;66(23):2098-2101.
19. Ash JS, Sittig DF, Dykstra R, Campbell E, Guappone K. The unintended consequences of computerized provider order entry: findings from a mixed methods exploration. *International journal of medical informatics*. 2009;78 Suppl 1:S69-76.
20. Wright A, Hickman TT, McEvoy D, et al. Analysis of clinical decision support system malfunctions: a case series and survey. *Journal of the American Medical Informatics Association : JAMIA*. 2016.
21. Ash JS, Sittig DF, Campbell EM, et al. Some unintended consequences of clinical decision support systems. *AMIA Annu Symp Proc*. 2007;Annual Symposium Proceedings/AMIA Symposium.:26-30.
22. Osheroff JA, Teich JM, Levick D, et al. *Improving Outcomes with Clinical Decision Support An Implementers Guide*. Second ed: Healthcare Information Management Society; 2012.
23. Wright A, Sittig DF, Ash JS, Sharma S, Pang JE, Middleton B. Clinical decision support capabilities of commercially-available clinical information systems. *Journal of the American Medical Informatics Association : JAMIA*. 2009;16(5):637-644.
24. Carey RG. *Improving healthcare with control charts : basic and advanced SPC methods and case studies*. Milwaukee, WI: ASQ Quality Press; 2003.
25. Ben-Gal I. *Outlier Detection*. Kluwer Academic Publishers; 2005.
26. Smith SC, Jr., Benjamin EJ, Bonow RO, et al. AHA/ACCF Secondary Prevention and Risk Reduction Therapy for Patients with Coronary and other Atherosclerotic Vascular Disease: 2011 update: a guideline from the American Heart Association and American College of Cardiology Foundation. *Circulation*. 2011;124(22):2458-2473.
27. Keim DA. Information Visualization and Visual Data Mining. *IEEE Trans Vis Comput Graph*. 2002;8(1):1-8.

Appendix A

Structured Query Language script for data abstraction from EHR database.

```
☐ Select
  a.alt_id,
  b.contact_date,
  a.alert_desc,
  a.bpa_locator_id

  From alert a inner join alt_history b on a.alt_id = b.alt_id

  WHERE b.contact_date BETWEEN ('1-JAN-10') and ('30-July-15')
  AND general_alt_type_c =1
  AND b.was_shown_c=0
```

©2016 EPIC Systems Corporation. Used with permission.

Appendix B

STATA Do-file script with dataset refinement commands

```
2
3 clear
4 //Imports Full BPA Query
5 import delimited //Insert file name here
6
7 //creating STATA Dates
8 gen alert_date = date(contact_date, "DM20Y")
9 format alert_date %td
10 sort alert_date alt_id
11
12 //Tagging the first occurrence of alt_id with the variable first_alert==1
13 //Then dropping all other occurrences of the same alt_id
14 by alt_id, sort: gen first_alert=_n==1
15 drop if first_alert!=1
16 drop first_alert
17
18 //Dropping BPAs not shown and no longer released
19 drop if bpa_locator_id==13 drop if bpa_locator_id==19 drop if bpa_locator_id==24
20 drop if bpa_locator_id==925 drop if bpa_locator_id==981 drop if bpa_locator_id==981
21 drop if bpa_locator_id==986 drop if bpa_locator_id==1018 drop if bpa_locator_id==1030
22 drop if bpa_locator_id==1047 drop if bpa_locator_id==1067 drop if bpa_locator_id==1101
23 drop if bpa_locator_id==1240 drop if bpa_locator_id==1285 drop if bpa_locator_id==1295
24 drop if bpa_locator_id==1299 drop if bpa_locator_id==1332 drop if bpa_locator_id==1355
25 drop if bpa_locator_id==1698 drop if bpa_locator_id==1751 drop if bpa_locator_id==1800
26 drop if bpa_locator_id==1810 drop if bpa_locator_id==1820 drop if bpa_locator_id==1825
27 drop if bpa_locator_id==2103 drop if bpa_locator_id==2120 drop if bpa_locator_id==2157
28 drop if bpa_locator_id==2127 drop if bpa_locator_id==2185 drop if bpa_locator_id==2197
29 drop if bpa_locator_id==2199 drop if bpa_locator_id==2201 drop if bpa_locator_id==2206
30 drop if bpa_locator_id==2247 drop if bpa_locator_id==2249 drop if bpa_locator_id==2253
31 drop if bpa_locator_id==2257 drop if bpa_locator_id==2259 drop if bpa_locator_id==2359
32 drop if bpa_locator_id==201103032 drop if bpa_locator_id==201103090 drop if bpa_locator_id==201104010
33 drop if bpa_locator_id==201106010 drop if bpa_locator_id==201106040
34 drop if bpa_locator_id==201108010 drop if bpa_locator_id==201108012
35 drop if bpa_locator_id==201108040 drop if bpa_locator_id==201109012
36 drop if bpa_locator_id==201110050 drop if bpa_locator_id==304100001
37 drop if bpa_locator_id==304300550 drop if bpa_locator_id==304500051
38 drop if bpa_locator_id==408010007 drop if bpa_locator_id==408010016
39 drop if bpa_locator_id==408010019 drop if bpa_locator_id==1600000020
40 drop if bpa_locator_id==2101060001 drop if bpa_locator_id==2101080000
41 drop if bpa_locator_id==2101100001 drop if bpa_locator_id==30400000028
42 //END Dropping BPAs
43
44 //Seperating Weekdays from weekends
45 gen day=dow(alert_date)
46 gen wkend =1 if day ==0 | day ==6
47 replace wkend=0 if wkend!=1
48
49 export delimited using //insert filename here
50
```

Appendix C

Statistical Process Control c-chart creation

Average Line TrimmedBPA2000-2015_first_alerts_wkend-wkdays

Results are computed along Table (Across).
`WINDOW_AVG(SUM([Number of Records]))`

The calculation is valid. Sheets Affected Apply OK

Default Table Calculation

LCL TrimmedBPA2000-2015_first_alerts_wkend-wkdays

Results are computed along Table (Across).
`[Average Line]-[Standard Deviation]*WINDOW_STDEV(SUM([Number of Records]))`

The calculation is valid. Sheets Affected Apply OK

Default Table Calculation

UCL TrimmedBPA2000-2015_first_alerts_wkend-wkdays

Results are computed along Table (Across).
`[Average Line]+[Standard Deviation]*WINDOW_STDEV(SUM([Number of Records]))`

The calculation is valid. Sheets Affected Apply OK

Default Table Calculation

©2016 Tableau Software Inc. Used with permission.

Table 1

<u>CDS Rule</u>	<u>SPC Detection</u>	<u>CDS Malfunction</u>
Propofol Shortage	1,2,3	No
ER EKG Ordering	3	No
Chemotherapy Ordering	2	No
IV Rx Stop Time	3	No
Osteoporosis Screening	3	No
Foley Catheter	0	No
Osteoporosis Screening #2	3	No
Treatment Protocol	3	No
MRI and Observation status	3	Unknown
Sigmoidoscopy	1	Unknown
2nd Generation anti-psychotics	3	Unknown
Pnemococcal Vaccination	3	Unknown
Post Stroke anti-platelet Rx	1	Unknown
Daypatient status	3	Unknown
Observation Status	3	Unknown
HMO Insurance	NA	Unknown
Colorectal Cancer Screening	3	Unknown
Coronary artery disease Rx	1,2,3	Yes
Enoxaparin orderset use	2,3	Yes
Influenza at Discharge	NA	Yes
Patient Height documentation	1	Yes

Table 1. CDS rules identified as candidate visual anomalies. SPC Detection column lists the test which was violated, see methods section for more details. NA = Not applicable

Table 2

<u>CDS Rule</u>	<u>Malfunction Pathway</u>
Height documentation	Target clinical department was discontinued
Enoxaparin order set use	Medication record accidentally added to preference list
Influenza vaccination	Rule not activated for influenza season
Coronary artery disease management	Direct editing of rule logic

Table 2. CDS malfunctions and corresponding pathways.

Figure 1

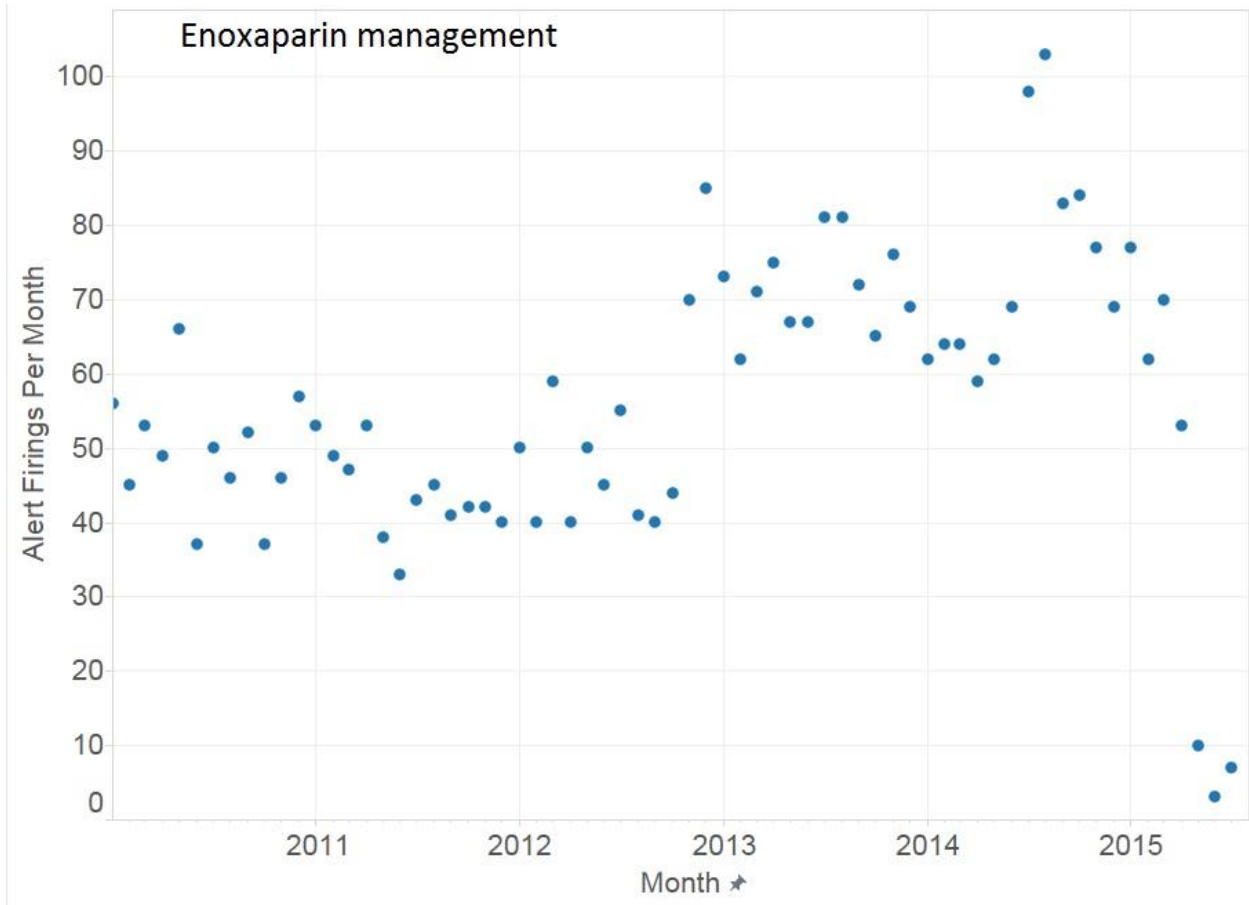


Figure 1. Alert activity from CDS rule related use of enoxaparin order set, determined to be a malfunction

Figure 2

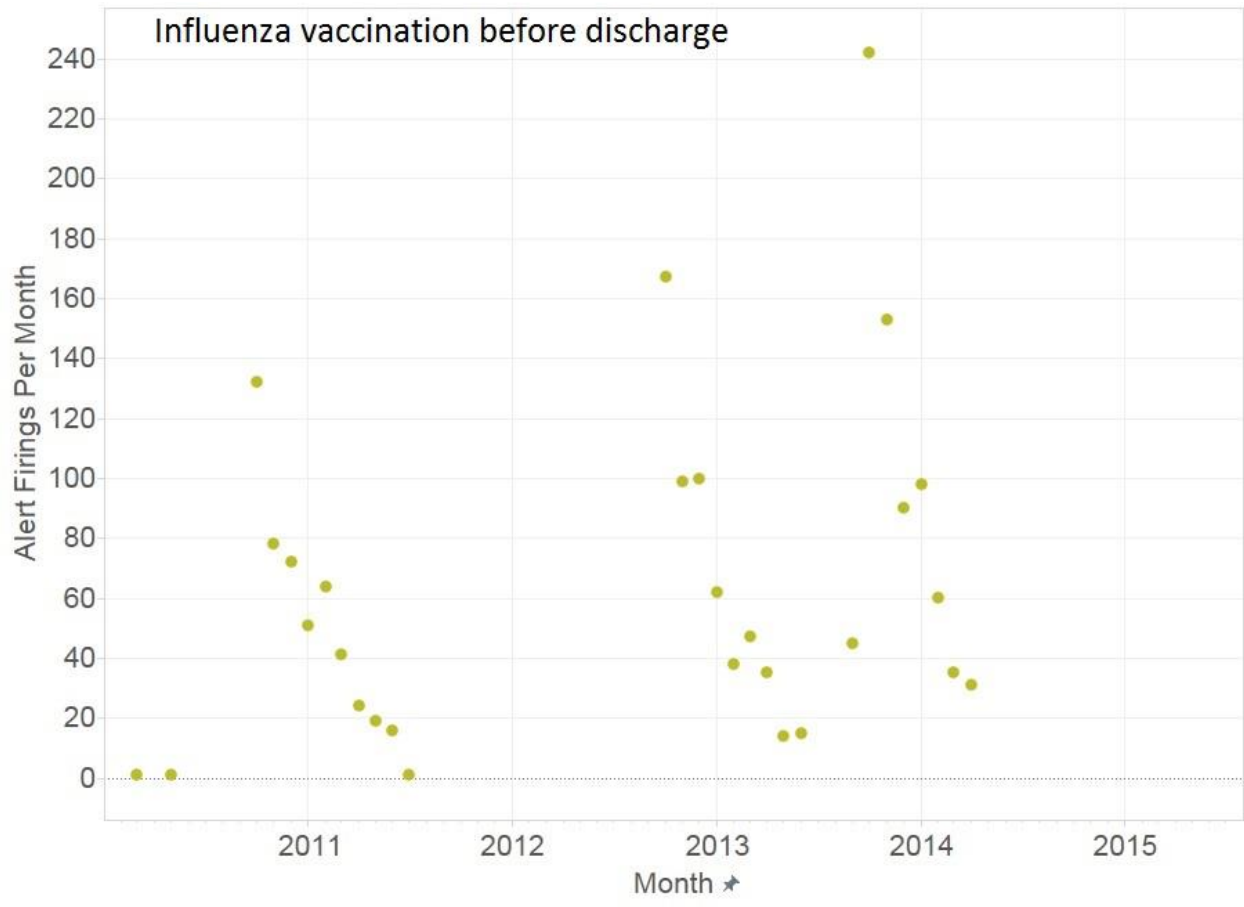


Figure 2. Alert activity from CDS rule related to influenza administration prior to discharge, determined to be a malfunction

Figure 3

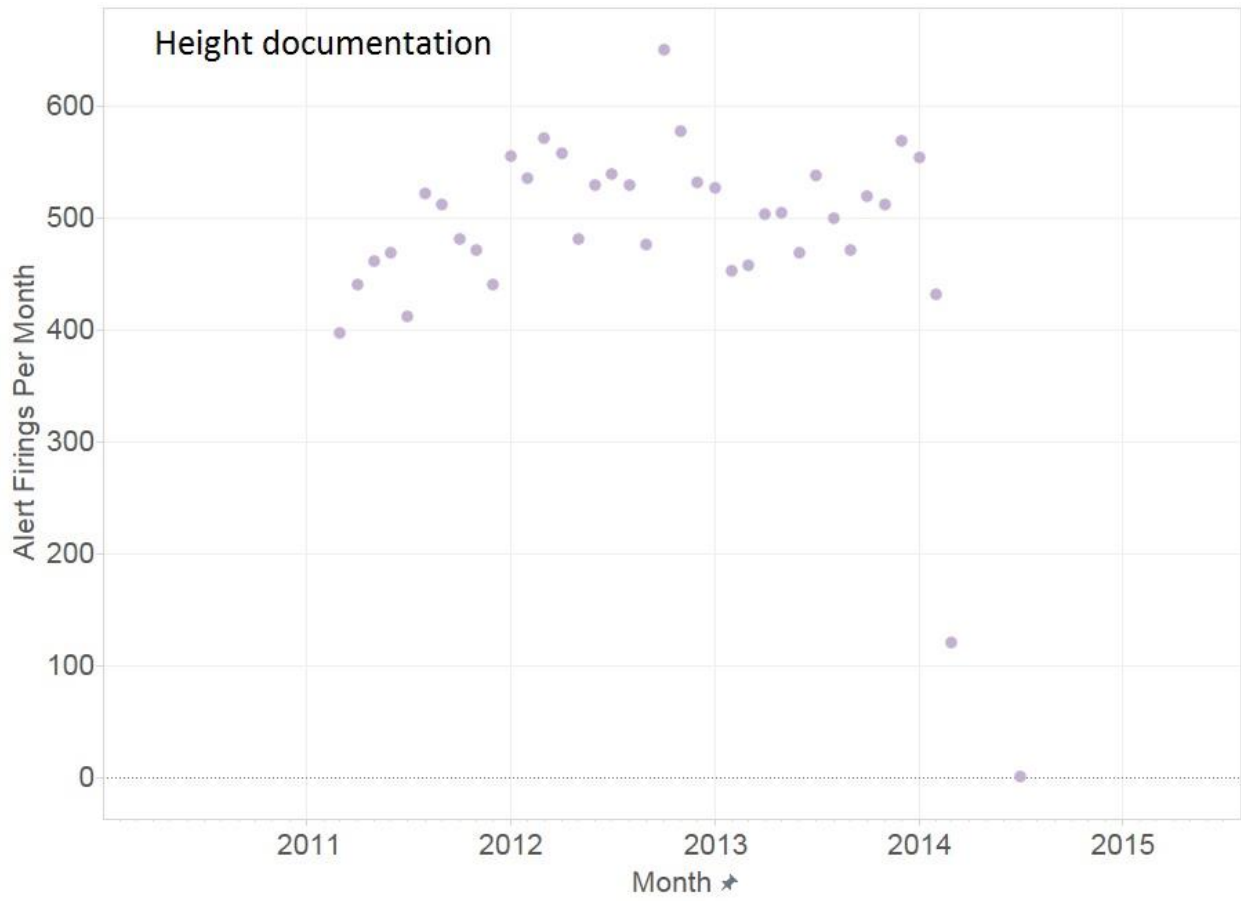


Figure 3. Alert activity from CDS rule related to height documentation in oncology clinic, determined to be a malfunction

Figure 5

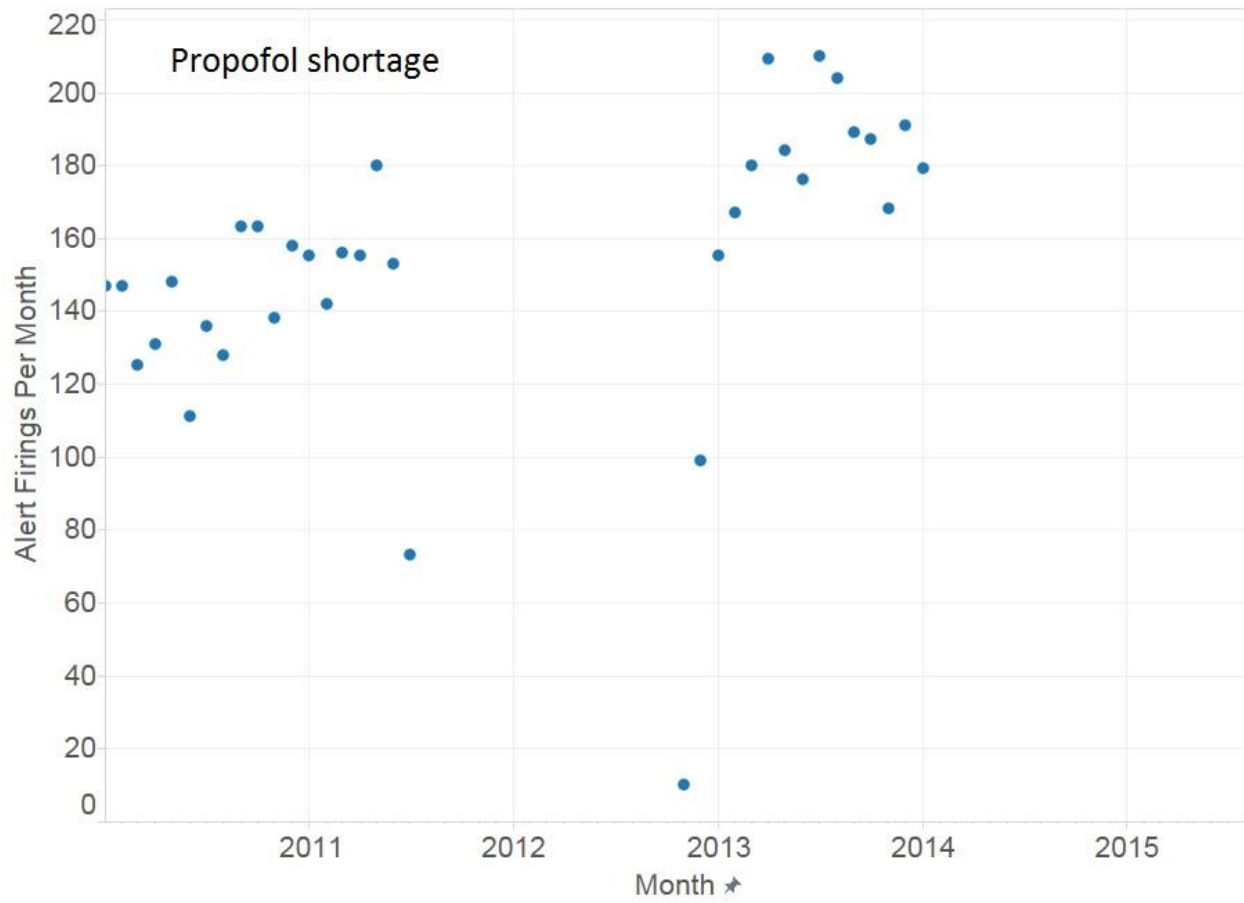


Figure 5. Propofol alert activity shown as example of visual anomaly categorized as a false positive following manual validation.

Figure 6

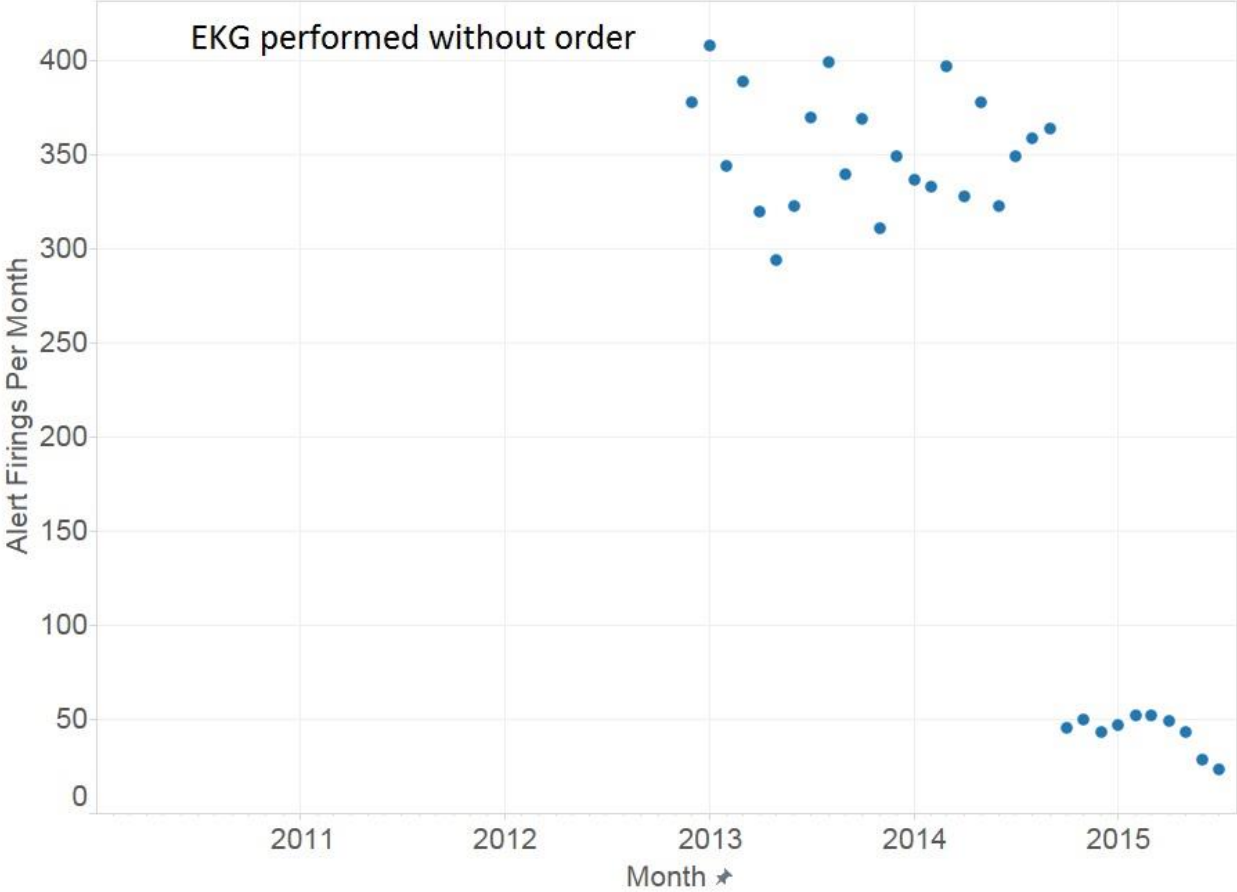


Figure 6. CDS rule alerting RN that patient has no corresponding order for EKG. This CDS rule represents an example of a visual anomaly which was found to be a false positive following manual validation.

