Rare Adverse Event Monitoring of Medical Devices with the Use of an Automated Surveillance Tool

Michael E. Matheny, MD, MS[1,2,3], Nipun Arora, MD[4], Lucila Ohno-Machado, MD, PhD[1,3], Frederic S. Resnic, MD, MS [1,4]

[1] Decision Systems Group, Brigham & Women's Hospital, Boston, MA
[2] Division of General Medicine, Brigham & Women's Hospital, Boston, MA
[3] Harvard-MIT Division of Health Sciences & Technology, Cambridge, MA
[4] Division of Cardiovascular Medicine, Brigham & Women's Hospital, Boston, MA

Abstract

Prospective outcomes surveillance using population level data allows for statistical methodologies and confounder adjustment not supported by the FDA's current monitoring system. We explored propensity score matching integrated into an automated surveillance tool as a method for confounder adjustment in an observational cohort.

The application analyzed all patients undergoing PCI via femoral access route from 2002-2006. The rare outcome of interest was retroperitoneal hemorrhage (RPH) and the device was a vascular closure device (VCD). A propensity score model was developed to match VCD and non-VCD match patients.

Our tool was able to detect sustained elevations in RPH among those patients who received a VCD. A root cause analysis revealed an association between high femoral access and RPH which prompted an educational program to modify clinical practice. Our results suggest use of propensity score matching can play a useful role in computer-based surveillance of rare events in a prospective cohort.

Introduction

Product recalls of both medications and medical devices in recent years by the Food and Drug Administration (FDA) have highlighted the limitations present in the current adverse event reporting system. The FDA is well aware of these problems, and has recently solicited the Institute of Medicine as well as numerous industry and academic experts to provide insight into ways to improve the system. The current system is capable of detecting completely unexpected outcomes, but it lacks important clinical patient data, and does not capture population level (denominator) event rate data. The detection of rare events in particularly difficult, as there is no method for robust risk adjustment.

The increasing prevalence of electronic health records in institutions and health networks provides an opportunity to perform complementary types of adverse event monitoring on those data. Challenges remain to format routine inpatient and outpatient clinical data collection for this purpose. However, domain specific mandatory registries are becoming increasingly common, and provide excellent sources of data by requiring standard data collection for a given set of elements for all patients that meet the inclusion criteria. For example, patient registries for cardiac surgery and interventional cardiology have been mandated by some states (New York and Massachusetts) and are used to provide quality assessment of physicians and intuitions. These registries collect a rich set of data for all patients undergoing those surgeries or procedures that could be used for a wide variety of prospective, risk-adjusted outcomes monitoring in a relatively inexpensive way.

Prospective and exposure-adjusted automated adverse event monitoring based on data sent periodically to mandatory registries may provide complementary medical product safety monitoring to what is currently used by the FDA. The use of propensity scores (PS) to adjust for confounding in a prospective cohort in regular intervals has not been explored. To illustrate the potential usefulness of our tool, we utilized our local institutional interventional cardiology registry to evaluate whether retroperitoneal hemorrhage (RPH), a rare adverse event, was associated with the use of vascular closure devices (VCDs).

Background

Surveillance Tools

Serial evaluation of prospective outcomes within medicine is best described in randomized, controlled clinical trials. The data safety monitoring board reviews the data in intervals during the study and evaluates whether the event rates in each of the arms are significantly different. For outcomes considered adverse events, there is less concern with type II error associated with repeated analysis of outcomes in order to emphasize specificity over sensitivity, and the number of measurements in a trial can vary. The utility of post-marketing data used to monitor new

medications or medical devices relates directly to the breadth of data collected and the ability to control for confounding. Almost all applications that prospectively monitor outcomes were designed for randomized clinical trials, (e.g., Clinitrace, Phase Forward, Waltham, MA; Trialex, Fremont, CA) and there are very few examples of systems that perform registry monitoring.

Propensity Score Matching

In non-randomized studies, there are ongoing efforts to develop new methodologies to control for confounding, and propensity score matching has become one of the successful ways to do this, especially when the outcome of interest is rare. A propensity score (PS) is a conditional probability of exposure to a treatment given measured covariates. This method can be used to match scores in an observational patient cohort who did and did not receive the treatment, and removes measured confounding among those covariates.[3] Large numbers of covariates may be used for this purpose.[4] It has been reported that this method may outperform traditional logistic regression adjustment when the number of positive outcomes per covariate is seven or less.[5] A primary limitation of this method is the inability to adjust for confounding among unmeasured covariates.[6]

Methods

Application

Data Extraction and Longitudinal Time Analysis (DELTA) was developed to provide prospective outcomes monitoring for any clinical data source that could be imported in regular intervals. The tool uses a web-based graphical user interface developed in Microsoft.NET (Microsoft, Redmond, VA), and stores data and algorithms in a SQL 2000 server. (Microsoft, Redmond, VA) DELTA allows the user to specify a desired confidence interval or sigma to generate an alerting threshold, and to select the time interval for analysis. When the application detects an elevated outcome rate for a given exposure, alerts are generated and emailed to the designated researcher. Full details of the specifications and design of the application are available elsewhere.[2]

Monitoring Methodology

A statistical module was developed and added to DELTA in order to allow the system to generate propensity scores and perform matching between samples with and without the exposure of interest. Propensity score model development was performed

by SAS (Version 9.1, Cary, NC), and incorporated into DELTA. PS matching was enforced between the groups within a time interval. PS model quality was assessed by mean squared error (MSE) and area under the operating curve (AUC) characteristic.[4]

Matched observations without the exposure were placed in the 'control' group, and matched observations with the exposure were placed in the 'case' group. The cumulative number of events and observations per specified time period were used to calculate a difference of proportions by the Wilson method in the 'case' and 'control' groups.[7] Point estimates of the difference of proportions with confidence intervals (CI) were generated by these calculations, and if the CIs of an estimate did not cross 0, a statistically significant difference was detected between the groups for that time period.

Clinical Use Case

The interventional cardiology catheterization laboratory at Brigham & Women's Hospital (BWH), Boston, MA, maintains a detailed database of all patient cases requiring percutaneous coronary angioplasty. Data collection conforms to the American College of Cardiology National Data Repository standard data definitions,[1] and real-time acquisition is accomplished through a team of trained nurses, physicians, and technologists. All patients are followed prospectively by the team for any inpatient post-procedural vascular complications. This study was approved by the BWH Institutional Review Board. 8,324 patient cases received PCI by femoral access route between January 01, 2002 and December 31, 2006 at BWH and were included for this analysis. Time intervals of three months were evaluated, and alerting thresholds were generated using 95% confidence intervals.

The exposure of interest was a vascular closure device (VCD) which is used to stop bleeding from a femoral access site after removal of the arterial sheath upon completion of the procedure. If a VCD was not used, then manual compression (MC) of the arterial puncture site was used to stop the bleeding. The primary outcome was retroperitoneal hemorrhage (RPH) after percutaneous coronary intervention (PCI). RPH is diagnosed clinically by physicians caring for the patient on the basis of clinical signs, and confirmation of the diagnosis is made by computed tomography imaging.

Unadjusted (crude) event rate comparisons were performed between groups using all of the cases available in order to provide a baseline. This type of analysis does not adjust for confounding.
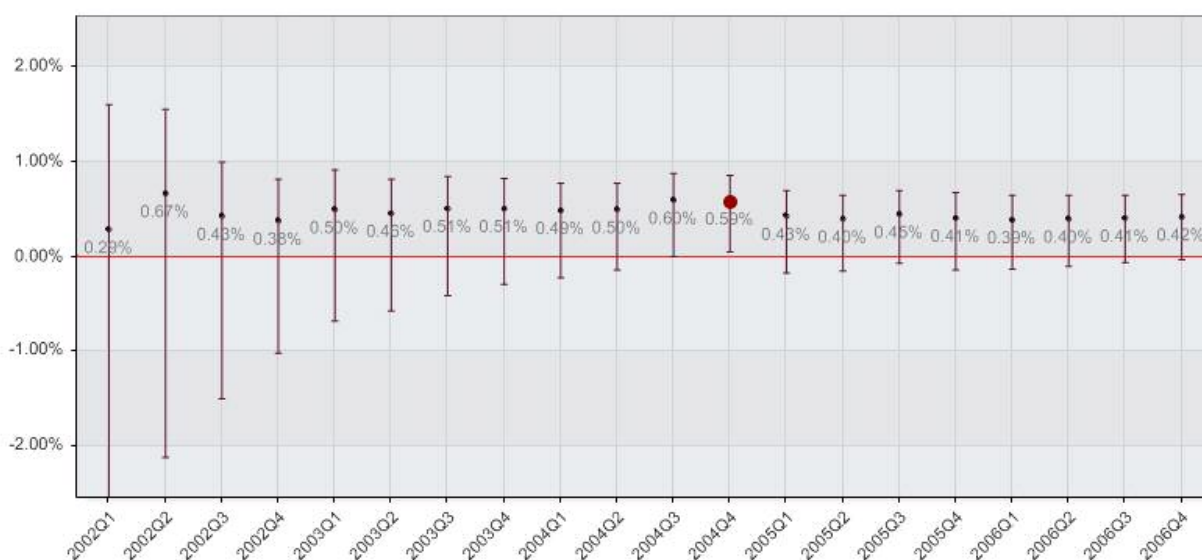
Figure 1: Results by calendar quarter between the use of a vascular closure device and manual compression in the unadjusted patient procedure cohort. The figure shows the cumulative proportional difference point estimate with 95% confidence intervals, with significant differences between the groups highlighted by the enlarged red dots.

A propensity score for VCD was calculated using 62 covariates modeling the probability of exposure to the device. The covariates were all clinical characteristics collected at the point of care, including age, sex, race, chronic medical conditions, and acute clinical findings available prior to the decision to use a VCD. VCD patients were then randomly matched 1:1 with MC patients within the same calendar quarter of their procedures to propensity scores of +/- 0.03.

A literature search was also conducted for possible confounders for RPH. One predictor of RPH (obtaining femoral arterial access above the inguinal ligament) was reported in a single study but this predictor is not collected in our registry nor is it part of the national data collection recommendations.[10] Subsequently, a root cause analysis was conducted, which included manual chart review to capture additional covariates in the sample, and further risk adjustment using the additional data.

Results

Among all 8,324 cases, 43 cases were complicated by RPH. There were 41 RPH events among 7,129 patient cases in which a VCD was used, and 2 RPH events among 1,238 patient cases in which MC was used. Crude analysis of RPH event rates between MC and VCD patient groups is shown in Figure 1, and an isolated alert was generated in the last quarter of 2004.

1,144 of the 1,238 possible cases were matched in each group, with 2 and 10 RPH events in the MC and VCD exposure groups, respectively. Comparison of patient demographics and common chronic clinical conditions between the two matched groups are displayed in Table 1. The propensity score model resulted in an AUC of 0.70, and an MSE of 0.114.

Cumulative proportional difference analysis between MC and VCD groups by calendar quarter resulted in significant differences detected in 9 out of

Characteristic	MC (%)	VCD (%)	p value
Age (mean)	69.2	69.2	0.90
Female	454 (39.7)	446 (39.0)	0.77
Race			
White	846 (74.0)	864 (75.5)	0.41
AA	28 (2.5)	34 (3.0)	0.52
Hispanic	23 (2.0)	26 (2.3)	0.77
Other	18 (1.6)	20 (1.8)	0.87
Unknown	229 (20.0)	200 (17.5)	0.13
Smoker	662 (57.9)	678 (59.3)	0.52
Hyperlipidemia	841 (73.5)	840 (73.4)	1.00
Hypertension	942 (82.3)	959 (83.8)	0.37
Diabetes	428 (37.4)	434 (37.9)	0.83

Table 1: MC = Manual Compression, VCD = Vascular Closure Device. AA = African American. 1144 cases in each group.
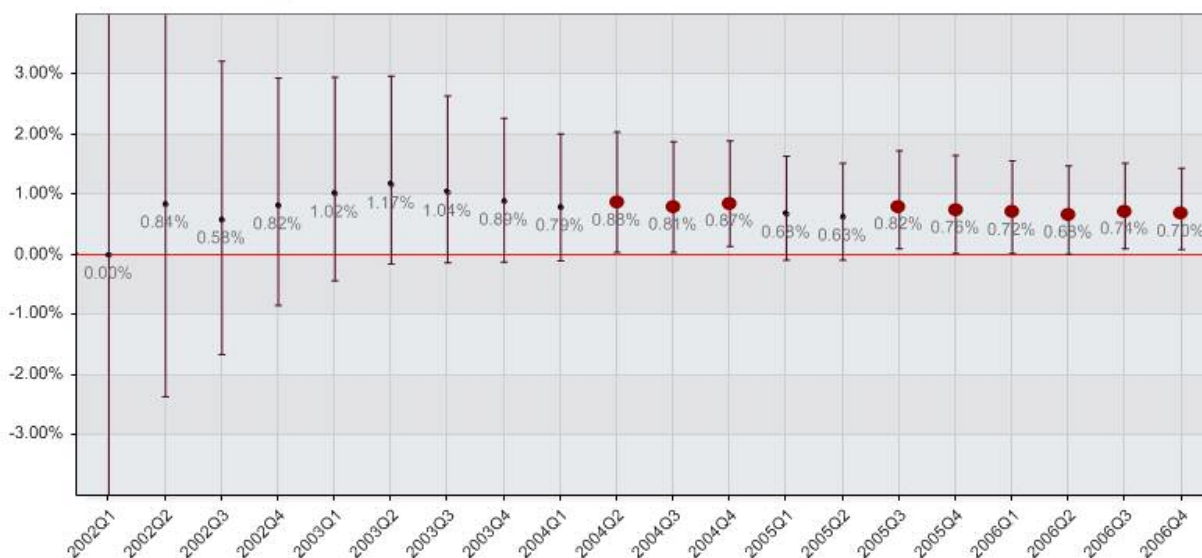
Figure 2: Results by calendar quarter between the use of a vascular closure device and manual compression in the PS matched patient procedure cohort. The figure shows the cumulative proportional difference point estimate with 95% confidence intervals, with significant differences between the groups highlighted by the enlarged red dots.

20 quarters. The first alert was generated by a difference of 0.87% (0.04% - 2.03%) in the second quarter of 2004, and continued to alert throughout the rest of 2004. The first two quarters of 2005 did not generate alerts, but from then on the difference regained statistical significance in the third quarter of 2005 with a value of 0.82% (0.10% - 1.73%). The alerts continued throughout the conclusion of the study with a final cumulative difference of 0.70% (0.09% - 1.43%). A summary graph of the analysis is shown in Figure 2.

In the root cause analysis, high femoral access was found to be associated with RPH. Including this variable in the risk adjustment model removed the association between VCD and RPH.

Discussion

The overall event rate in our local institution was comparable to that reported by other institutions.[8-10] An association between the use of vascular closure devices and retroperitoneal hemorrhage has been reported in two retrospective studies,[8, 9] and refuted in one retrospective study.[10]

In our institution, DELTA detected a significantly elevated rate of retroperitoneal hemorrhage in those patients who received a VCD when compared with those that received MC at the conclusion of a PCI procedure. This was first detected nine quarters after the initiation of the analysis, and remained significant at the end of the study. The delay in detecting an

event rate elevation was likely due to the rarity of the adverse event, which also dictated a relatively large time interval of analysis. The crude analysis detected a single significant elevation *six months after* the one detected by the PS-score adjusted system. The mild elevation (difference of 0.59%) detected in the non-adjusted system was not confirmed in the following quarter, and could have easily been ignored in practice.

The root cause analysis findings suggest that there was confounding by indication between high femoral access and the use of a VCD. As a result of this finding, an educational intervention is currently underway in the catheterization lab to reduce the incidence of high femoral access.

The primary limitation in this device safety methodology is that PS is unable to account for unmeasured confounding (i.e., variables that are not in the database, as was shown with the association with high femoral access and VCD). PS can also be limited when populations between matching arms are very disparate, such as when a device has a strong indication or contra-indication. This is reflected by a low matching rate, and conclusions cannot be drawn for population sub-groups not matched. This was not observed in our study (92.4% matched). Finally, repeated measurements increase the probability of false positives. DELTA is intended as an early screening tool with an emphasis on specificity over sensitivity, but determination of the acceptable balance between false negatives and false positives

needs to be studied before its practical application in a clinical environment.

In summary, this study highlights the usefulness of rare adverse event rate monitoring by an automated application that utilizes propensity scores to match cases. This automated monitoring system can alert researchers when a potential problem with the use of a medical device requires further investigation. Checking whether unmeasured confounding removes the association between the outcome and device of interest is then necessary, but automating the initial safety screening allows many devices and outcomes to be monitored simultaneously.

We illustrated one of several potential applications of this safety monitoring tool. In our example, the system was able to properly indicate an increased occurrence of RPH in patients receiving VCDs, prompting investigation of related causes and the development of an education program to modify medical practice

References

1. Cannon CP, Battler A, Brindis RG, et al. American College of Cardiology key data elements and definitions for measuring the clinical management and outcomes of patients with acute coronary syndromes. A report of the American College of Cardiology Task Force on Clinical Data Standards (Acute Coronary Syndromes Writing Committee). *J Am Coll Cardiol.* Dec 2001;38(7):2114-2130.
2. Matheny ME, Ohno-Machado L, Resnic FS. Monitoring device safety in interventional cardiology. *J Am Med Inform Assoc.* Mar-Apr 2006;13(2):180-187.
3. Rosenbaum P, Rubin D. The central role of the propensity score in observational studies for causal effects. *Biometrika.* 1983;70:41-55.
4. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Sturmer T. Variable selection for propensity score models. *Am J Epidemiol.* Jun 15 2006;163(12):1149-1156. Epub 2006 Apr 1119.
5. Cepeda MS, Boston R, Farrar JT, Strom BL. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *Am J Epidemiol.* Aug 1 2003;158(3):280-287.
6. Joffe MM, Rosenbaum PR. Invited commentary: propensity scores. *Am J Epidemiol.* Aug 15 1999;150(4):327-333.
7. Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. *Stat Med.* Apr 30 1998;17(8):873-890.
8. Ellis SG, Bhatt D, Kapadia S, Lee D, Yen M, Whitlow PL. Correlates and outcomes of retroperitoneal hemorrhage complicating percutaneous coronary intervention. *Catheter Cardiovasc Interv.* Apr 2006;67(4):541-545.
9. Cura FA, Kapadia SR, L'Allier PL, et al. Safety of femoral closure devices after percutaneous coronary interventions in the era of glycoprotein IIb/IIIa platelet blockade. *Am J Cardiol.* 2000 Oct 1;86(7):780-782.
10. Farouque HM, Tremmel JA, Raissi Shabari F, et al. Risk factors for the development of retroperitoneal hematoma after percutaneous coronary intervention in the era of glycoprotein IIb/IIIa inhibitors and vascular closure devices. *J Am Coll Cardiol.* Feb 1 2005;45(3):363-368.