SENTINEL CBER/PRISM METHODS

PILOT OF SELF-CONTROLLED TREE-TEMPORAL SCAN ANALYSIS FOR GARDASIL VACCINE

Prepared by: W. Katherine Yih*,[1] Judith C. Maro*,[1] Michael Nguyen,[2] Meghan A. Baker,[1] Carolyn Balsbaugh,[1] David V. Cole,[1] Inna Dashevsky,[1] Adamma Mba-Jonas,[2] Martin Kulldorff[3]

Author Affiliations: 1. Department of Population Medicine, Harvard Pilgrim Health Care Institute and Harvard Medical School, Boston, MA; 2. Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, MD; 3. Brigham and Women's Hospital and Harvard Medical School, Boston, MA

September 30, 2016

The Sentinel System is sponsored by the U.S. Food and Drug Administration (FDA) to proactively monitor the safety of FDA-regulated medical products and complements other existing FDA safety surveillance capabilities. The Sentinel System is one piece of FDA's Sentinel Initiative, a long-term, multi-faceted effort to develop a national electronic system. Sentinel Collaborators include Data and Academic Partners that provide access to healthcare data and ongoing scientific, technical, methodological, and organizational expertise. The Sentinel Coordinating Center is funded by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223201400030I. This project was funded by the FDA through HHS Mini-Sentinel contract number HHSF223200910006I.

Sentinel CBER/PRISM Methods Protocol

Pilot Of Self-Controlled Tree-Temporal Scan Analysis For Gardasil Vaccine

Table of Contents

I. INTRODUCTION

In the U.S., observational epidemiologic studies of vaccine safety after licensure have generally entailed either analyses of disproportionate reporting of adverse events to the Vaccine Adverse Event Reporting System (VAERS)[1] or analyses of the association between a vaccine and one or more pre-specified adverse health outcomes, such as the studies typically conducted by the CDC-sponsored Vaccine Safety Datalink (VSD)[2] and the FDA-sponsored Post-licensure Rapid Immunization Safety Monitoring (PRISM) system[3,4]. VAERS-based analyses, while useful for identifying previously unsuspected possible adverse reactions to vaccination, have a number of limitations shared by passive reporting systems[1,5,6]. VSD and PRISM studies, while population-based and generally well-designed and rigorously conducted, have generally not addressed identification of previously unsuspected possible adverse reactions.

The TreeScan method allows a wide range of unsuspected but potential adverse reactions to be simultaneously evaluated[7-10]. The main advantage is that otherwise unknown adverse reactions may be found. The main disadvantage is that it is not possible to adjust for all possible confounders. Indeed, no conclusion about causality should be based on TreeScan analyses alone. In effect, the TreeScan method serves as a tool for identifying adverse events that may merit a careful pharmacoepidemiologic investigation.

The purpose of the current project was to develop and test unconditional and conditional variants of the TreeScan self-controlled tree-temporal scan statistic, both with and without day-of-week adjustment, for vaccine safety surveillance, using automated electronic health insurance claims data from the Sentinel system. With the tree-temporal version of TreeScan, the risk window is not pre-specified. The method simultaneously evaluates several thousand potential adverse events and groups of related adverse events, while simultaneously evaluating a large number of potential risk windows, adjusting for the multiple testing inherent in the many types of adverse events and risk windows evaluated. The project was intended as a pilot to prepare for the future use of the TreeScan method as part of FDA's 18-month post-licensure safety review of vaccines and assess the viability of the method for monitoring the safety of vaccines given in adolescence in particular, with Human Papillomavirus Quadrivalent (Types 6, 11, 16, 18) Vaccine (Gardasil®; Merck & Co.) as the test vaccine. (For convenience, we refer to this vaccine as "HPV4" in this study report.)

This is the third PRISM TreeScan project, a continuation of prior methodological work to develop tree-based scan statistics for post-market vaccine safety surveillance. In the two prior PRISM TreeScan projects, we evaluated conditional and unconditional versions of the *Poisson* based tree scan statistic for vaccine cohort data, conditional and unconditional versions of the *Bernoulli* based tree scan statistic for self-control data, and a self-controlled *unconditional tree-temporal* scan statistic[10], the last of which was also used in a secondary analysis in the current project. The various TreeScan methods are compared in summary form in Appendix 1. Based on findings from the prior work, we expected the hitherto undeveloped *conditional tree-temporal* scan statistic to perform quite well for study populations beyond early childhood.

The eight aims of the original protocol[11] are listed in Table 1 below. This report addresses Aims 3, 4, and 7. The other aims, too, have been largely achieved, and that work will be presented in other papers.

Table 1. Specific aims of original protocol, with corresponding reports/papers (identified and distinguished from each other by letters A-D) and first authors.

#	Aim	Report/manuscript addressing aim	First author
1	Develop and evaluate a conditional version of the tree-temporal scan statistic	Methods paper (A)	M. Kulldorff
2	Develop and evaluate a day-of-week adjusted version of the tree-temporal scan statistic (for both unconditional and conditional variants)	Methods paper (A)	M. Kulldorff
3	Evaluate the tree-temporal scan statistic for use with adolescent vaccines	This report (B)	K. Yih & J. Maro
4	Provide comparative results in order to inform the choice of what type of tree-temporal scan statistic to use for adolescent vaccines (e.g., conditional vs. unconditional, with vs. without day-of-week adjustment) and what parameter settings to use	This report (B)	K. Yih & J. Maro
5	Enhance the TreeScan software to perform power evaluation for the tree-temporal scan statistic	TreeScan software enhancement, not a report	TreeScan programmer
6	Evaluate power for the tree-temporal scan statistic when used for adolescent vaccines, considering different sample sizes, outcomes, and relative risks	Paper on power (C)	J. Maro
7	Explore and document practices for first-line follow-up of TreeScan-generated statistical alerts	General paper (D) This report (B)	J. Maro (D) K. Yih & J. Maro (B)
8	Explore and document approaches to assess bias and time varying confounding in TreeScan statistical alerts	General paper (D)	J. Maro

II. METHODS

A. OVERVIEW OF ANALYSES

The primary analysis was the conditional self-controlled tree-temporal scan statistic, without day-of-week adjustment. In secondary analyses, we included a day-of-week adjustment, to account for the fact that all vaccines and some outcomes have an uneven weekly pattern, with, for example, more observations on weekdays than on weekends. We also conducted unconditional self-controlled tree-temporal analyses, with and without day-of-week adjustment. Table 2 summarizes these analyses.

Table 2. Tree-temporal analyses conducted (post-vaccination observation period = Days 1–56).

#	1°/2°	Conditional/unconditional	Adjusted for outcome-specific day-of-week effects	Alpha level for alerting
1	1°	Conditional	No	0.05
2	2°	Conditional	Yes	0.01
3	2°	Unconditional	No	0.01
4	2°	Unconditional	Yes	0.01

As mentioned, the method formally adjusts for the multiple testing generated by the many groupings of outcomes and many potential risk windows that are evaluated in a single analysis for a specific exposure. The secondary analyses represented additional multiple testing. We adjusted for this additional multiple testing informally by pre-specifying an alpha level of *0.01* to reject the null hypothesis in these secondary analyses. These methods are explained in Section II.H.

B. STUDY POPULATION, ENROLLMENT CRITERIA, AND FOLLOW-UP PERIOD

Data in Sentinel Common Data Model-format were obtained for 6/1/2006-12/31/2014 or the maximum available date range within that period from five Sentinel/PRISM sites ("Data Partners"): Aetna, Harvard Pilgrim Health Care, HealthCore, Humana, and Optum. For some Data Partners, there were restrictions on the use of even de-identified patient-level data for certain members, so those members were excluded.

We included both female and male vaccinees who received the HPV4 vaccine on or after their 9th birthday and before their 27th birthday during the available date range. Only members enrolled in the participating health plans for at least 183 consecutive days prior to vaccination were included in the base study population to allow incident (first-in-6-months) diagnoses to be ascertained. The follow-up period during which we identified cases of health outcomes of interest was Days 1-56 after vaccination. Thus, members also had to have been enrolled for ≥ 56 days after vaccination in order to be included. Enrollment gaps of ≤ 45 days were treated as continuously enrolled time.[1]

C. EXPOSURE

HPV4 vaccination was identified using CPT code 90649.

Only health outcomes in Days 1-56 after the *first* apparent dose were included in analysis. We considered an HPV4 dose to be a first dose if there was no prior record of an HPV4 dose for that patient, going back the maximum amount of available time, but no earlier than his/her 9th birthday. All subsequent doses were ignored, regardless of the timing of their occurrence. Descriptive statistics assembled in preparation for the PRISM study of Gardasil and venous thromboembolism[12] indicated that only 5.7% of Dose 2s were given within 56 days of Dose 1, hence the 56 days of follow-up time used for all first doses were likely largely unaffected by a subsequent dose.

[1] Apparent gaps in enrollment can occur due to administrative glitches during annual renewals or switches from one plan to another within the same health insurance company. They may or may not reflect true lapses in coverage and may or may not cause health events during the gaps to be missed. The allowance of apparent gaps in enrollment of up to 45 days is standard practice in Sentinel studies and is applied uniformly regardless of whether the gap is in a pre-exposure or a post-exposure period. While it is possible that apparent gaps in enrollment might lead to health events (such as health outcomes after HPV4) not being ascertained, we think the possibility of the TreeScan results being noticeably affected by allowing apparent enrollment gaps of up to 45 days is remote. This sense is reinforced by the fact that most HPV vaccination occurs in August and most reenrollment in health insurance companies occurs several months later, in December-January.

D. HIERARCHICAL DIAGNOSIS TREE

Outcomes were identified and defined using ICD-9 codes and a classification of all ICD-9 codes into a hierarchical tree structure defined by the Multi-Level Clinical Classification Software (MLCCS). The MLCCS is a product of the Agency for Healthcare Research and Quality's Healthcare Cost and Utilization Project (http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp). The tree has five diagnosis levels, although some branches extend only to the second or third level. The first and broadest level identifies 18 body systems, while the entries at the finest level contain one or multiple ICD-9 codes. **Figure 1** is a heuristic representation of the tree.

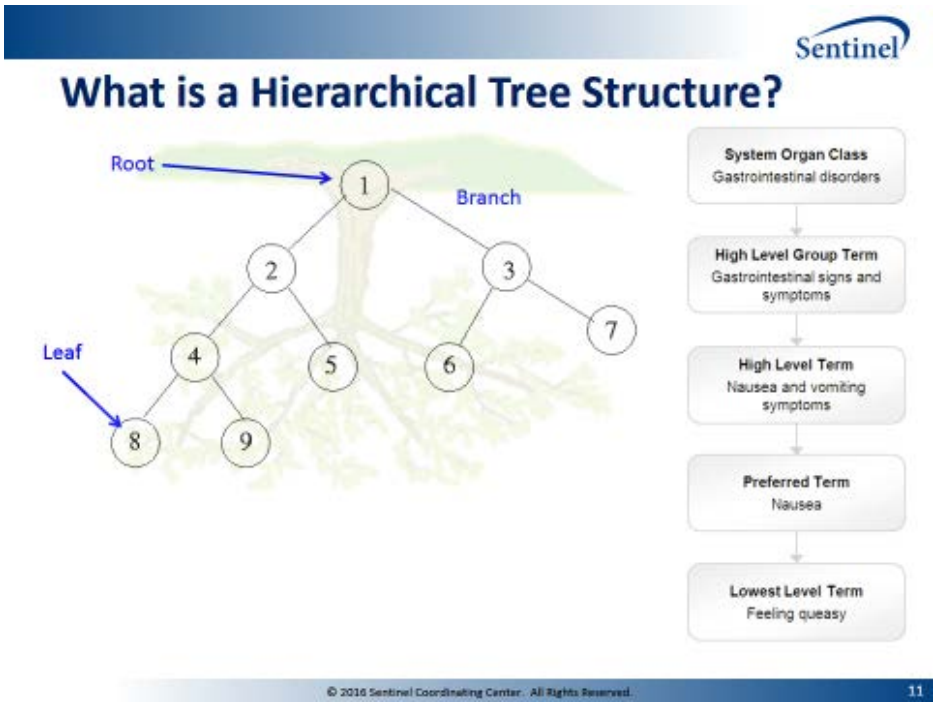


Figure 1. Heuristic diagram of a hierarchical tree of diagnoses. The MLCCS tree has 18 possible values at the root level, for the 18 body systems. The subsequent levels become more and more specific, ending with ICD-9 codes at the finest, "leaf" level. (The diagram lacks one level compared to the actual MLCCS tree, which has four levels before the leaf level.) (This figure was excerpted from a presentation to FDA by Judith Maro on March 31, 2016.)

As an example, "convulsions" is a third-level classification without a fourth level and corresponds to five different ICD-9 codes:

Table 3. Example of MLCCS hierarchical classification scheme.

ICD9 Code	Description
06	Diseases Of The Nervous System And Sense Organs
06.04	..Epilepsy; convulsions
06.04.02	....Convulsions
780.3	……Convulsions
780.31	……Febrile convulsions
780.32	……Complex febrile convulsions
780.33	……Post traumatic seizures
780.39	……Other convulsions

In the hierarchical tree we used, we filled in all levels out to the fourth level for all outcomes. For example, ICD-9 code 729.5, "pain in soft tissues of the limb," corresponds to a second-level outcome on the tree, 13.08, in which the first level 13 is "diseases of the musculoskeletal system and connective tissue" and the second level 08 is "other connective tissue disease." There is no finer differentiation of this outcome, but we created additional levels for it by adding zeroes—the third level is 13.08.00 and the fourth level is 13.08.00.00.

Some ICD-9 codes were excluded from the tree and therefore from the analysis, for example, those representing:

- Outcomes that are very unlikely to be caused by vaccination, such as well-care visits, delivery of a baby, vitamin deficiencies, or fracture of a lower limb
- Some conditions unlikely to manifest themselves within the short follow-up time we are dealing with, such as cancer
- Most infectious diseases with an identified organism (e.g., typhoid fever, tuberculosis, shigella)
- Congenital conditions (e.g., sickle cell disease, congenital heart disease)
- Outcomes that are common and of an unspecific or less serious nature, such as fever, croup, and acute pharyngitis.

E. INCIDENT DIAGNOSES OF INTEREST

The study focused on incident diagnoses observed during the 56-day follow-up period, since a repeat diagnosis may have been due to a follow-up visit for an earlier episode of illness and less likely due to the vaccine. A diagnosis was considered an incident diagnosis if it was observed in the inpatient or emergency department (ED) setting during the follow-up period and if there was no other diagnosis for that patient in the same third-level branch of the MLCCS diagnosis tree in any setting during the prior 183 days. This means that, even if it was a never-before-seen ICD-9 code, it was not counted if a different ICD-9 code belonging to the same third-level branch was observed during the prior 183 days. Based on results of testing, the third level was chosen for determining incidence in order to avoid double-counting and overestimation of incidence which may occur when physicians classify the same episode of illness in two slightly different ways (e.g., convulsions and febrile convulsions) in separate patient visits.

We allowed each patient to contribute multiple incident diagnoses during his/her follow-up period, as long as they were not part of the same third-level branch of the MLCCS tree. In the unusual situation

where a patient had multiple incident diagnoses on the same third-level branch on the same day, the program selected the rarest incident outcome, using an outcome frequency list based on emergency department and inpatient data for 9-26.99 year olds in Harvard Pilgrim Health Care.

Appendix 2 illustrates these rules by means of two fictional patients.

F. RISK AND COMPARISON WINDOWS

As mentioned in Section II.B., we evaluated outcomes occurring 1-56 days after vaccination. The day of vaccination (Day 0) was not included since (a) a preventive care visit at which vaccines were given could have generated diagnosis codes (outcomes) unrelated to vaccination, such as problems found during an eye examination, and (b) HPV4 may have been given during a health care visit that happened due to an illness or other health concern. We evaluated all temporal risk windows that were at least 2 days long, were at most 28 days long, started sometime between 1 and 28 days after vaccination, and ended sometime between 2 and 42 days after vaccination. The comparison period consisted of the days within the 56-day follow-up period that were not in the risk window being evaluated.

G. DATA FORMATS

The data provided by the Data Partners were in strata that included:

- ICD-9 code entered by clinician or coder
- Number of days between exposure and outcome (all values in the range of 1-56)
- Number of cases

The rules for determining incident diagnoses were applied at the sites, using the MLCCS tree, as this process requires access to patient-level data.

In addition, we obtained the total number of HPV4 first doses given in the age groups of interest. These were used to calculate attributable risks.

For each tree-temporal scan analysis there was a separate analysis data set, in the same format as above but including data from all the Data Partners.

H. TREE-TEMPORAL SCAN STATISTIC, UNCONDITIONAL AND CONDITIONAL

With the tree-temporal scan statistic, one performs multiple temporal scan statistics, one for each of the many overlapping branches of the tree, adjusting for the multiple testing stemming both from the many branches and from the many time intervals evaluated. Each time interval is evaluated on each of the branches, so with our approximately 7300 nodes (i.e., outcome categories, whether first, second, third, fourth, or fifth level, which include, for example, the codes listed in Table 3) on the tree and our 665 potential time intervals, there were more than 4.8 million potential clusters to evaluate and for which we needed to adjust for multiple testing. If these had been 4.8 million independent tests with non-overlapping data, there would have been a huge loss in power when adjusting for all the multiple testing. With scan statistics, such a large loss in power does not happen, since many of the potential clusters (4.8 million, in our case) are highly overlapping with each other. Hence, the penalty for adjusting for the multiple testing is more modest. Furthermore, no power is lost (i.e., no alpha is spent) in scanning nodes where the observed number of events in the follow-up period is less than 2.

Considering the thousands of overlapping disease outcome categories evaluated, adjustment for multiple testing is critical. This is accomplished through the simulation component of the method. The likelihood ratio test statistic from the most likely cut in the real dataset is compared with the likelihood ratio test statistics from the most likely cuts in each of, say, 99,999 random datasets, and we note its rank. For example, if it has the fifth highest test statistic, its rank R = 5. Note that the most likely cut will be on a different branch in the different datasets, so we are not comparing the likelihood ratios for the same cut, but rather, comparing the maxima of the likelihood ratios obtained over all possible cuts. Since the random datasets were all generated under the null hypothesis, if the null hypothesis is true in the real dataset, then the test statistics come from exactly the same probability distribution. This means that, if the null hypothesis is true, the rank test statistic from the real dataset will range uniformly from 1 to 100,000, and the probability of having a rank in the top 5% is exactly 5%. If the test statistic from the real dataset is in the top 5%, we will reject the null hypothesis, and we have a 5% probability of falsely rejecting the null. We generated 99,999 random datasets for each of our four analyses (the primary and the three secondary analyses).

The tree-temporal scan statistic conditions the analysis on the number of cases observed in each node of the tree. This means that, unlike the standard tree-based scan statistic, there is no probability distribution to model the number of cases in each node, but rather, it is deterministic. What is probabilistic is the timing of each case. In an analysis that is unconditional with respect to time, under the null hypothesis, the cases are assumed to be uniform across the follow-up period. In a conditional analysis, we not only condition on the number of cases observed in each node of the tree but also on the total number of cases occurring on the first day after vaccination, on the second day after vaccination, etc. This adjusts for the type of temporal confounding that would occur if there were some temporal differences in the general healthcare-seeking behavior shortly after compared to longer after the vaccination date. Under the alternative hypothesis, there is at least one branch of the tree for which there is a temporal cluster of cases during some time interval.

The tree-temporal scan statistic can be applied with various analysis parameter settings. We set our risk window parameters as stated in Section II.F.

1. Unconditional Tree-Temporal Scan Statistic

With the unconditional tree-temporal scan statistic and under the null hypothesis, any outcome is equally likely to occur on any of the days following the initial drug/vaccine exposure. For each tree node and time interval, we calculate the log likelihood ratio (LLR) test statistic:

$$LLR = \ln \frac{(\frac{c}{n})^c (\frac{n-c}{n})^{n-c}}{(\frac{w}{T})^c (\frac{T-w}{T})^{n-c}} \; I(c/n > w/T)$$

where n is the number of cases in the node, c is the number of those node cases that are also in the time interval, w is the length of the time interval, and T is the total length of the follow-up period (56 days, in this study). I() is the indication function, which is 1 when there are more cases in the time interval than expected under the null, and it is included to ensure that we are looking for an excess risk

of having the outcome rather than a protective decreased risk. Note that T is a constant that is the same for every node and every time interval.

For each node on the tree, the LLR is calculated for each time interval under consideration. The node-interval combination with the maximum LLR is the most likely cluster of cases, that is, the cluster that is least likely to have occurred by chance. Regardless of the data, there is always a most likely cluster, so that in itself does not mean that there is a true cluster.

Branches with zero events do not contribute to the analysis. Also, if there is only 1 case at the fifth level, that is, only one case with a specific ICD-9 code, no signal is possible for that specific ICD-9 code, although that case *can* contribute to a signal on one of the higher level branches.

The distribution of the test statistic is not known analytically, so there is no simple mathematical formula that can be used to obtain a p-value for the detected cluster. To evaluate whether the most likely cluster is statistically significant, after adjusting for the multiple testing inherent in the many node-interval combinations considered, Monte Carlo hypothesis testing is used in a one-sided test that looks for excess risk at the alpha=0.05 level. This is done by generating, say, 99,999 random replicates of the data. In each random data set, each node has exactly the same number of cases as the real data set, but the post-exposure timing of those cases varies, with each one generated from a uniform distribution independent of each other case. For each random data set, generated under the null hypothesis, we find the most likely cluster in the same way as we did for the real data set, and we note the maximum LLR of that data set. Note that the node and time interval for the most likely cluster will typically be different in each of the random data sets and also different from the real data. If the null hypothesis is true, then the maximum LLR from the real data set has a 5% chance of being among the 5000 highest maximum LLRs from the real and random data sets, so if that is the case, we can reject the null hypothesis at the alpha=0.05 level. If R is the rank of the maximum LLR from the real data set so that there are exactly R-1 random data sets with a higher maximum, the Monte Carlo based p-value is R/(S+1), where S is the number random data sets used, or 99,999 in our case. Adjustment for multiple testing is assured since we are comparing the maximum from the real data set with the maxima from the random data sets.

2. Conditional Tree-Temporal Scan Statistic

For each tree node and time interval, we calculate a Poisson generalized log likelihood ratio test statistic. Let n be the number of cases in the node, let c be the number of those node cases that are also in the time interval, let z be the number of cases in the time interval summed over the whole tree, and let C be the total number of cases in the tree. The number of cases in the cluster, c, is then contrasted with the expected number of cases in the cluster under the null hypothesis, which is u=nz/C. The test statistic is then

$$T = c \ln[c/u] + (C-c) \ln [(C-c)/(C-u)] \, I(c>u)$$

where I() is the indication function. I() is 1 when there are more cases than expected in the cluster, and 0 otherwise, and it is included to ensure that we are looking for an excess risk of having the outcome rather than a protective decreased risk.

For each node on the tree, the test statistic is calculated for each time interval under consideration. The node-interval combination with the maximum test statistic is the most likely cluster of cases, that is, the

cluster that is least likely to have occurred by chance. Regardless of the data, there is always a most likely cluster, so that in itself does not mean that there is a true cluster.

The distribution of the test statistic is not known analytically, so there is no simple mathematical formula that can be used to obtain a p-value for the detected cluster. To evaluate whether the most likely cluster is statistically significant, after adjusting for the multiple testing inherent in the many node-interval combinations considered, Monte Carlo hypothesis testing is used. We do this by generating 99,999 random replicates of the data. In each random data set, each node has exactly the same number of cases as the real data set, and each day after vaccination has the same number of cases when summed over all nodes (a key difference from the simulations done for the unconditional tree-temporal scan statistic). The only thing that varies is the pairing of the nodes and times, which is randomized using a permutation approach. For each random data set, generated under the null hypothesis, we find the most likely cluster in the same way as we did for the real data set, and we note the maximum test statistic of that data set. Note that the node and time interval for the most likely cluster will typically be different in each of the random data sets and also different from the real data. If the null hypothesis is true, then the maximum test statistic from the real data set has a 5% chance of being among the 5000 highest maximum test statistics from the one real and 99,999 random data sets, so if that is the case, we can reject the null hypothesis at the alpha=0.05 level. If R is the rank of the maximum test statistic from the real data set so that there are exactly R-1 random data sets with a higher maximum, the Monte Carlo based p-value is R/(99,999+1), when 99,999 random data sets are used. Adjustment for multiple testing is assured since we are comparing the maximum from the real data set with the maxima from the random data sets.

3. Day of Week Adjustment

Under the null hypothesis, the unconditional tree-temporal scan statistic assumes that all outcomes are equally likely during the 1-56 days after vaccination. This assumption may be violated for some outcomes that, for example, are more commonly diagnosed on a weekday than a weekend. This day-of-week effect can be adjusted for by assuming equal probability of the outcome on Days 1, 8, 15, … 50 and so on, equal probability on Days 7, 14, 21, … 56 and so on, and likewise for the rest of the seven sets of days, but *different* probabilities *among* those sets of days. The randomization in the Monte Carlo step is then done stratified on those seven groupings.

The conditional tree-temporal scan statistic automatically adjusts for any day-of-week effect that is common to all the outcomes. To adjust for any day-of-week effect that is different for different outcomes, it is necessary to do the randomization stratified on the above groupings. The kinds of day-of-week adjustment implemented in the various analyses are shown in Table 4.

Table 4. Tree-temporal analyses conducted (like Table 2 but with detail about day-of-week adjustment).

#	1°/2°	Conditional/ unconditional	Adjusted for outcome-specific day-of-week effects	Adjusted for day-of-week effects common to all outcomes	Alpha level for alerting
1	1°	Conditional	No	Yes, inherently	0.05
2	2°	Conditional	Yes	Yes, inherently	0.01
3	2°	Unconditional	No	No	0.01
4	2°	Unconditional	Yes	Yes, by virtue of adjustment for outcome-specific day-of-week effects	0.01

I. STATISTICAL ALERT FOLLOW-UP

Data related to statistical alerts were frozen soon after the TreeScan analysis was conducted. From the frozen data for one selected alert, a TreeScan Vaccine Episode Report (TVER) was generated that listed all the procedures, drug dispensings, and diagnoses captured in the claims data during the period -56 days through +84 days of the HPV4 vaccination for each patient with an incident diagnosis that contributed to the alert. The procedures for generating TVERs, which involve the Patient Episode Profile Retrieval (PEPR) system, are explained in detail in the report "Infrastructure for Evaluation of Statistical Alerts Arising from Vaccine Safety Data-mining Activities in Mini-Sentinel"[13].

Two members of the work group, including an internal medicine physician, reviewed and interpreted these claims profiles.

III. RESULTS

A total of 1,903,697 first doses of HPV4 vaccine were included in analysis.

The analysis results are shown in Table 5. All alerts with $p < 0.05$ are shown. In addition, in the case of statistically significant results at the fourth level (e.g., 12.01.01.03), results are shown for the fourth-level's corresponding first, second, and third levels (e.g., 12, 12.01, and 12.01.01), regardless of statistical significance; these are related and should be interpreted together. Similarly, results for ICD-9 codes within 16.10.02.07 for which there were any cases during Days 1-56 are included for context, regardless of statistical significance. There were no alerts with p just slightly greater than 0.05—the lowest p value for diagnoses not included in the table was 0.25 for the primary analysis and 0.19 for the secondary analyses.

Table 5. Details of statistical alerts from primary and secondary tree-temporal scan statistical analyses of HPV4 vaccine. Some related diagnoses for which there were no alerts are included for context. RW = risk window, Obs = number of cases observed in RW, AR = attributable risk in terms of number of excess cases per 100,000 first-dose vaccinees.

			Primary analysis, criterion for statistical significance pre-specified as p < 0.05				Secondary analyses, criterion for statistical significance pre-specified as p < 0.01											
			Conditional				Conditional with day-of-week adjustment				Unconditional				Unconditional with day-of-week adjustment			
Row #	Node code	Node text	RW	Obs	AR/ 100K	p	RW	Obs	AR/ 100K	p	RW	Obs	AR/ 100K	p	RW	Obs	AR/ 100K	p
1	12	Diseases of the skin and subcutaneous tissue	2-4	214	3.8	0.0019	1-4	266	4.0	0.16	2-4	214	3.8	0.0089	1-4	266	4.0	0.00065
2	12.01	. Skin and subcutaneous tissue infections	2-4	111	2.3	0.042	2-4	111	2.3	0.92	2-4	111	2.3	0.19	2-4	111	2.3	0.021
3	12.01.01	. . Cellulitis and abscess	2-4	93	2.0	0.20	1-4	115	2.2	0.99	2-4	93	2.0	0.66	1-4	115	2.2	0.082
4	12.01.01.03	. . . Cellulitis and abscess of arm (only 682.3)	2-3	31	1.3	0.00001	2-4	38	1.5	0.00002	2-3	31	1.3	0.00001	2-3	31	1.3	0.00001
5	682.3	. . . . Cellulitis and abscess of upper arm and forearm	2-3	31	1.3	0.00001	2-4	38	1.5	0.00002	2-3	31	1.3	0.00001	2-3	31	1.3	0.00001
6	12.02	. Other inflammatory condition of skin	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
7	695.9	. . . . Unspecified erythematous condition	2-3	13	0.5	0.25	2-3	13	0.5	0.94	2-3	13	0.5	0.81	2-3	13	0.5	0.20
8	16	Injury and poisoning	1-3	48	2.2	0.00001	1-3	48	2.2	0.00001	1-3	48	2.2	0.00001	1-2	40	1.9	0.00001
9	16.10	. Complications	1-3	36	1.8	0.00001	1-6	42	2.1	0.00001	1-3	36	1.8	0.00001	1-3	36	1.8	0.00001
10	16.10.02	. . Complications of surgical procedures or medical care	1-3	36	1.8	0.00001	1-6	42	2.1	0.00001	1-3	36	1.8	0.00001	1-3	36	1.8	0.00001
11	16.10.02.07	. . . Other complications of surgical and medical procedures	1-3	36	1.8	0.00001	1-6	42	2.1	0.00001	1-3	36	1.8	0.00001	1-3	36	1.8	0.00001
12	780.63	. . . . Post-vaccination fever	1-2	4	0.2	0.31	1-4	5	0.3	0.95	1-4	5	0.3	0.041	1-2	4	0.2	0.24
13	999.0	. . . . Generalized vaccinia	1-3	3	0.2	> 0.99	..	..	..	..	1-3	3	0.2	0.96	1-3	3	0.2	> 0.99
14	999.4	. . . . Anaphylactic reaction due to serum	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
15	999.42	. . . . Anaphylactic reaction due to vaccination	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
16	999.5	. . . . Other serum reaction not elsewhere classified	1-3	7	0.4	0.011	1-3	7	0.4	0.66	1-3	7	0.4	0.0099	1-3	7	0.4	0.011
17	999.52	. . . . Other serum reaction due to vaccination	1-2	11	0.6	0.00001	1-2	11	0.6	0.00044	1-2	11	0.6	0.00001	1-2	11	0.6	0.00001
18	999.59	. . . . Other serum reaction	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
19	999.9	. . . . Other and unspecified complications of medical care	1-6	12	0.6	0.0018	1-6	12	0.6	0.0027	1-6	12	0.6	0.00025	1-6	12	0.6	0.0022

There were two sets of alerts, whose primary analysis results are described below. We will call these *sets* of alerts "signals," to distinguish them from their specific constituent "alerts." Secondary analysis results are discussed at the end of the Results section.

A. CELLULITIS AND ABSCESS OF ARM (NODE 12.01.01.03)

Within Branch 12, "diseases of the skin and subcutaneous tissue," there were alerts at four levels (Table 5, Rows 1, 2, 4, 5). The highest statistical significance was seen at the fourth and fifth levels (Rows 4 and 5), for cellulitis and abscess of the arm, with a risk window of Days 2-3, 31 cases, an attributable risk (AR) of 1.3/100,000 first doses administered, and p=0.00001. (ICD-9 code 682.3 is the only one stemming from 12.01.01.03, so the results are identical for Rows 4 and 5.) Considering especially the statistical significance, these 31 cases appear to be driving the Branch 12 alert (Row 1). The 13 cases with ICD-9 code 695.9, "unspecified erythematous condition" (within Node 12.02, "other inflammatory condition of skin"), on Days 2-3 (Row 7) could have contributed to the alert at Branch 12, too, although there was no alert for the specific 695.9 code itself.

Since cellulitis is a known adverse reaction to HPV4 vaccination and is listed as such in the package insert[14], no further investigation of this signal was conducted.

B. OTHER COMPLICATIONS OF SURGICAL AND MEDICAL PROCEDURES (NODE 16.10.02.07)

There were alerts at five levels within Branch 16, "injury and poisoning," with risk windows all within 6 days after vaccination (Table 5, Rows 8-11, 16, 17, 19). At the first level (Row 8), there was a risk window of Days 1-3, 48 cases, an AR of 2.2/100,000 doses, and p=0.00001. The second through fourth levels had the same Days 1-3 risk window, 36 cases, an AR of 1.8/100,000 doses, and p=0.00001. (The results for those Rows 9-11 are identical, because there is no branching between 16.10 and 16.10.02.07. This is due to branches within the "injury and poisoning" category in the full MLCCS tree, such as "complication of device, implant, or graft" and "postoperative infection," having been excluded from the tree that we used—see Section II.D.) These 36 cases appear to be driving the signal, as there are no alerts in other "injury and poisoning" second-level branches, namely 16.11 ("poisoning") or 16.12 ("other injuries and conditions due to external causes"). There were three alerts at the fifth level within Node 16.10.02.07 ("other complications of surgical and medical procedures"), including one for "other serum reaction due to vaccination" (Row 17), with a risk window of Days 1-2, 11 cases, AR=0.6/100,000 doses, and p=0.00001.

Fifty-eight patients had incident diagnoses at Node 16.10.02.07 ("other complications of surgical and medical procedures")—36 with their diagnoses in the Days 1-3 risk window (as shown in Table 5, Row 11, the same cases as in Rows 9 and 10) and 22 with their diagnoses during the Days 4-56 control window. On examination of the claims profiles of the patients, we determined that 31 (86%) of the 36 cases in the risk window and 11 (50%) of the 22 cases in the control window had received at least one other vaccine on the same day as HPV4. (Concomitant vaccines included tetanus-diphtheria-acellular pertussis, meningococcal conjugate, varicella, pneumococcal conjugate, hepatitis A, hepatitis B, inactivated influenza, live attenuated influenza, rabies, typhoid, polio, and meningococcal polysaccharide vaccines.) Nine (25%) of the 36 cases in the risk window and 6 (27%) of the 22 cases outside of the risk window received another dose of HPV vaccine within 84 days of the first (84 days was the post-vaccination period specified to be displayed in the claims profiles).

Focusing now on the 36 cases whose index code fell in the Days 1-3 risk window of the main signal, let us first review the adverse events that are known and listed in the package insert: "Headache, fever, nausea, and dizziness; and local injection site reactions (pain, swelling, erythema, pruritus, and bruising) occurred after administration with GARDASIL … Anaphylaxis has been reported following vaccination with GARDASIL." As mentioned in connection with the other signal, cellulitis is also listed in the package insert[14].

Considering first the cases in the risk window that had the more specific incident diagnoses of post-vaccination fever, generalized vaccinia, or anaphylaxis, we see that there were 4, 3, and 2 cases with these incident diagnoses, respectively (Table 6). All 4 cases of post-vaccination fever (ICD-9 code 780.63) had claims for one or more additional vaccines on Day 0. Of the 3 cases with the generalized vaccinia ICD-9 code 999.0 as the incident diagnosis, 2 had codes for pain in or swelling of the limb, and 1 had codes for allergic urticaria and unspecified urticaria, and in all 3 cases, those symptoms were the only additional ones specified. These 3 "vaccinia" patients had claims for at least two additional vaccines on Day 0, including varicella in all 3 cases. (Varicella is relevant because generalized varicella-like rash, which could plausibly be incorrectly coded as "generalized vaccinia," has been documented on Days 0-23 after varicella Dose 2 in adolescents[15].) Regarding the 2 cases of anaphylaxis (ICD-9 codes 999.4 and 999.42), it is unclear whether either was truly anaphylaxis related to HPV vaccination—neither case had claims for epinephrine, 1 case had received meningococcal conjugate vaccine on the same day as HPV4, and the other had the anaphylaxis code on Day 2, which seems somewhat late to have been true anaphylaxis related to Day 0 vaccination.

Table 6. Distribution of the 58 cases with incident diagnoses at Node 16.10.02.07 ("other complications of surgical and medical procedures") by index ICD-9 code and timing after HPV4 vaccination.

		Timing of case (index code)		
ICD-9	**ICD-9 description**	**Days 1-3**	**Days 4-56**	**Total**
780.63	POSTVACCINATION FEVER	4	1	5
999.0	GENERALIZED VACCINIA	3	0	3
999.4	ANAPHYLACTIC SHOCK-SERUM	0	1	1
999.42	ANAPHYLACT REACTION D/T VACCINATION	2	3	5
999.5	SERUM REACTION NEC	7	3	10
999.52	OTH SERUM REACTION D/T VACCINATION	11	2	13
999.59	OTHER SERUM REACTION	2	3	5
999.9	COMP MED CARE NEC & NOS	7	9	16
Total		36	22	58

The other four ICD-9 codes belonging to the "other complications of surgical or medical procedures" node and with non-zero case counts are less specific, and we considered them together. There were 27 cases in the risk window with one of these non-specific codes. We assigned each case to a category of

symptoms inferred from the patients' claims profiles. The distribution of cases among the categories, as well as with respect to concomitant vaccination status, is shown in Table 7.

Table 7. Distribution of the 27 cases in the risk window that had non-specific index codes (i.e., the codes in Table 6 that were not for post-vaccination fever, generalized vaccinia, or anaphylaxis) among symptom and concomitant vaccination categories.

Symptoms according to claims profiles	HPV4 alone	HPV4 with ≥ 1 other vaccine	Total
1. Pain in and/or swelling of limb	1	7	8
2. Local skin reactions and/or unspecified allergic reactions	1	4	5
3. Cellulitis	0	1	1
4. Diverse systemic symptoms[a]	1	5	6
5. Unspecified[b]	0	3	3
6. Other[c]	1	3	4
Total	4	23	27

[a] The diagnosis codes included in this category were for nausea and/or vomiting, fever, viral exanthem, dizziness and giddiness, headache, and unspecified myalgia and myositis. The subsequent pattern of medical visits suggests that the conditions were not severe.
[b] In all 3 cases with unspecified symptoms, the next visit was not until ≥ 60 d later, suggesting that the patient's condition did not require further urgent medical attention.
[c] The "other" category includes 3 cases with complicated and diverse medical conditions and 1 case of coded acute pharyngitis.

The clinical characteristics of 26 of the 36 cases contributing to the signal (4 fever + 2 anaphylaxis + 20 from Groups 1-4 in Table 7) thus appear to conform to what was already known about HPV4 adverse events (although HPV4 is not necessarily the cause—it must be kept in mind, for instance, that 86% of the 36 cases received one or more other vaccines along with HPV4). An additional 3 cases, coded as "generalized vaccinia," had received varicella vaccine at the same time as HPV4 and may have had varicella vaccine-associated varicella-like rash in addition to their claims consistent with known HPV4 adverse events, namely, for pain in or swelling of the limb (2 patients) and for allergic urticaria and unspecified urticaria (1 patient). The 3 cases with no symptoms specified (Group 5 in Table 7) had no subsequent visit for medical care until at least 60 days later. The other 4 cases (Group 6 in Table 7) had claims for a variety of conditions, which differed from one case to the next.

The results of the secondary statistical analyses, where the cut-off for statistical significance had been pre-specified as $p = 0.01$ to informally deal with the additional multiple testing, were similar to those of the primary analysis (Table 5) and identified the same two signals with essentially the same level of statistical significance. There were some minor differences in the risk windows identified by the four analyses; for example, in the conditional analysis with day-of-week adjustment, risk windows for eight nodes or outcomes were somewhat longer than in the primary analysis (Rows 1, 3-5, 9-12). For two outcomes (Rows 12 and 16), p-values varied from fairly small ($p < 0.05$) to fairly large probabilities across

the four analyses. However, the two main signals that emerged in the primary analysis were robust to variations in the analysis method, remaining highly statistically significant throughout.

IV. DISCUSSION

In our TreeScan analyses of more than 1.9 million recipients of HPV4 Dose 1, we found two signals of adverse events within 42 days of vaccination. One was in the category of "cellulitis and abscess of the arm." Cellulitis is a known adverse event and is listed in the HPV4 package insert; thus, this was not investigated further. The other signal was in the more general category of "other complications of surgical and medical procedures." Approximately 90% of the 36 cases contributing to that signal appeared to have either conditions already identified as possible vaccine-associated adverse events or (in 3 cases) no specified symptoms but also no further medical visits until at least 60 days after the visit in which the incident diagnosis was identified. The other ~10% (4) of the cases had diverse symptoms, different in each case.

The fact that only two signals were found from more than 7,000 leaves and branches of the hierarchical tree, neither one of which was unexpected, provides reassurance about both the vaccine (with respect to potential adverse reactions of acute onset within 42 days of vaccination) and the TreeScan conditional temporal-tree scan method. The method had good statistical power, detecting an AR of 4/1,000,000 doses for "other serum reaction not elsewhere classified."

The results of the secondary analyses, namely the unconditional version and the day-of-week adjustment of both the conditional and unconditional versions, were very similar to those of the primary analysis. For those considering using TreeScan, we would recommend the use of the conditional tree-temporal scan statistic without the day-of-week adjustment (our primary method) for most purposes. The reason we prefer conditional to unconditional versions relates to our prior TreeScan work in adults, in which an overall excess count of events of around 10% was observed when comparing the 1-28 vs. 29-56 day post vaccination windows across a wide range of different diagnoses. This was likely due to a tendency for adult preventive care visits to entail discoveries and testing leading to follow-up visits in the subsequent few weeks. As a result of the excess events, the unconditional self-control version of TreeScan generated several alerts for high level branches of the tree with modest relative risks of around 1.1 or 1.2. Conditional TreeScan methods adjust for this phenomenon. The reason we prefer no day-of-week adjustment is that no such adjustment is needed or appropriate for outcomes with no day-of-week effect. Thus, it seems preferable to conduct the initial analysis without it, repeating the analysis using the adjustment only if unexpected signals arise in the unadjusted analysis.

There are some limitations of the methods, which are either inherent to the tree-based scan statistic or related to the way it was implemented.

First, in this pilot project, we were able to evaluate only the first dose of the HPV4 vaccine series. This was a result of the way the data-extraction program was constructed. Building the ability to evaluate subsequent doses is a future planned enhancement.

Second, the analysis was done using one particular hierarchical tree, which included over 6,000 ICD-9 codes. We expect that most other trees developed with clinical expertise would generate similar

results, although we do not know this for certain—it is possible for the pattern of alerts generated by alternative trees to differ.

Third, we considered only risk windows that began between 1 and 28 days post vaccination and ended between 2 and 42 days post vaccination. We excluded events on the day of vaccination (Day 0) from the analysis to prevent capturing antecedent conditions that were present at the time of vaccination; thus, we could have missed outcomes with elevated risks within a few hours of vaccination. (One such outcome is anaphylaxis, but, as mentioned, this is already known to be associated with vaccination, albeit only rarely, and is listed in the package insert.) Beyond Day 0, we could detect only adverse reactions that manifested themselves within a few weeks of vaccination, i.e., outcomes of relatively acute onset. In order to look for possible adverse reactions occurring several months or years after vaccination, it would be necessary to use a longer follow-up period. While the tree-temporal scan statistic can in principle be used for longer follow-up periods, it is as yet untested for such applications.

Fourth, to concentrate on detecting more serious kinds of outcomes, we used only outcomes recorded at emergency department visits and/or inpatient hospital stays, while outcomes in outpatient settings were ignored (except in post-signal examination of patients' claim profiles). Thus, adverse reactions that are primarily treated in an outpatient setting could have been missed. It is possible, for example, that patients with conditions reported after HPV vaccination, such as new-onset autoimmune diseases, complex regional pain syndrome (CRPS), or postural orthostatic tachycardia syndrome (POTS), would have been seen first in an outpatient setting.

Fifth, multiple outcomes from the same patient were allowed, as long as they were not on the same third level branch of the MLCCS tree. In theory, this could create problems when evaluating higher levels of a tree. For example, since myocardial infarction (07.02.03) is often preceded by chest pain (07.02.05) a few days before, a patient could be counted twice at the 07 and 07.02 level analyses, and the tree-based scan statistic would erroneously ignore the dependence of these events. The possibility of such a problem should be kept in mind when evaluating statistical alerts at the two highest levels of the tree in the future. However, in our HPV4 case, all higher-level alerts were driven by alerts at more specific levels, so this problem was not in evidence.

Sixth, while the self-controlled versions of the tree-based scan statistic, including the tree-temporal scan statistic used here, automatically adjust for all time-invariant confounders, they do not adjust for time-varying confounders. For example, HPV vaccine uptake has a demonstrated pattern of seasonality, with the greatest uptake occurring in August prior to the start of the school year. Thus, in the case of outcomes that also are seasonal in nature, seasonality is a potential source of confounding and could produce a bias either toward or away from alerting, depending on the outcome in question. However, we have no reason to believe that confounding by seasonality produced a noticeable bias toward signaling in our study, as the signals we saw seemed real, and it is difficult to imagine a plausible scenario that would have produced a significant bias away from signaling.

V. CONCLUSION

In summary, when tree-temporal scan statistics were applied to 1.9 million recipients of HPV4 Dose 1 in the Sentinel system, only two signals of adverse events within 42 days of vaccination emerged, and both

were easily interpretable and explained. A notable limitation of the study is that only relatively early-onset outcomes were evaluated, i.e., only risk windows beginning on Days 1-28 post vaccination and ending on Days 2-42 post vaccination were considered; also, the day of vaccination (Day 0) was not included in analysis.

The major strengths of the tree-temporal scan statistical method are that it does not require specific potential adverse events or risk windows to be specified by the investigator and that it adjusts for the multiple testing inherent in the simultaneous evaluation of several thousand potential adverse events and a large number of potential risk windows. Further, our analysis had good statistical power. This is evident from the fact that three of the statistically significant alerts detected had an estimated attributable risk of only 4 to 6 excess cases per 1 million vaccinations. On the basis of the results reported here, we consider the conditional tree-temporal scan statistical method a versatile and powerful tool for assessing vaccine safety in adolescent and young adult populations.

VI. ACKNOWLEDGMENTS

The authors would like to gratefully acknowledge the contributions of Dr. Laura Polakowski for her clinical consultation.

VII. REFERENCES

1. Shimabukuro TT, Nguyen M, Martin D, DeStefano F. Safety monitoring in the Vaccine Adverse Event Reporting System (VAERS). Vaccine 2015;33:4398-405.
2. McNeil MM, Gee J, Weintraub ES, et al. The Vaccine Safety Datalink: successes and challenges monitoring vaccine safety. Vaccine 2014;32:5390-8.
3. Nguyen M, Ball R, Midthun K, Lieu TA. The Food and Drug Administration's Post-Licensure Rapid Immunization Safety Monitoring program: strengthening the federal vaccine safety enterprise. Pharmacoepidemiology and drug safety 2012;21 Suppl 1:291-7.
4. Baker MA, Nguyen M, Cole DV, Lee GM, Lieu TA. Post-licensure rapid immunization safety monitoring program (PRISM) data characterization. Vaccine 2013;31 Suppl 10:K98-112.
5. Varricchio F, Iskander J, Destefano F, et al. Understanding vaccine safety information from the Vaccine Adverse Event Reporting System. The Pediatric infectious disease journal 2004;23:287-94.
6. Iskander JK, Miller ER, Chen RT. The role of the Vaccine Adverse Event Reporting system (VAERS) in monitoring vaccine safety. Pediatric annals 2004;33:599-606.
7. Brown JS, Petronis KR, Bate A, et al. Drug Adverse Event Detection in Health Plan Data Using the Gamma Poisson Shrinker and Comparison to the Tree-based Scan Statistic. Pharmaceutics 2013;5:179-200.
8. Kulldorff M, Fang Z, Walsh SJ. A tree-based scan statistic for database disease surveillance. Biometrics 2003;59:323-31.
9. Kulldorff M, Dashevsky I, Avery TR, et al. Drug safety data mining with a tree-based scan statistic. Pharmacoepidemiology and drug safety 2013;22:517-23.
10. Kulldorff M, et al. A tree-temporal scan statistic for drug and vaccine safety data mining. Manuscript, 2014.
11. Yih WK, Nguyen M, Maro JC, et al. Mini-Sentinel CBER/PRISM methods protocol: Pilot of self-controlled tree-temporal scan analysis for Gardasil vaccine. 2015.
12. Yih WK, Greene SK, Zichittella L, et al. Evaluation of the risk of venous thromboembolism after quadrivalent human papillomavirus vaccination among US females. Vaccine 2016;34:172-8.
13. Cole DV, Kulldorff M, Baker M, et al. Infrastructure for evaluation of statistical alerts arising from vaccine safety data-mining activities in Mini-Sentinel. 2016.
14. Gardasil package insert. 2015. at http://www.fda.gov/downloads/BiologicsBloodVaccines/Vaccines/ApprovedProducts/UCM111263.pdf.)
15. Varivax package insert. 2012. at http://www.fda.gov/downloads/BiologicsBloodVaccines/Vaccines/ApprovedProducts/UCM142812.pdf.)
16. Kulldorff M. TreeScan User Guide [software manual]. 2015.

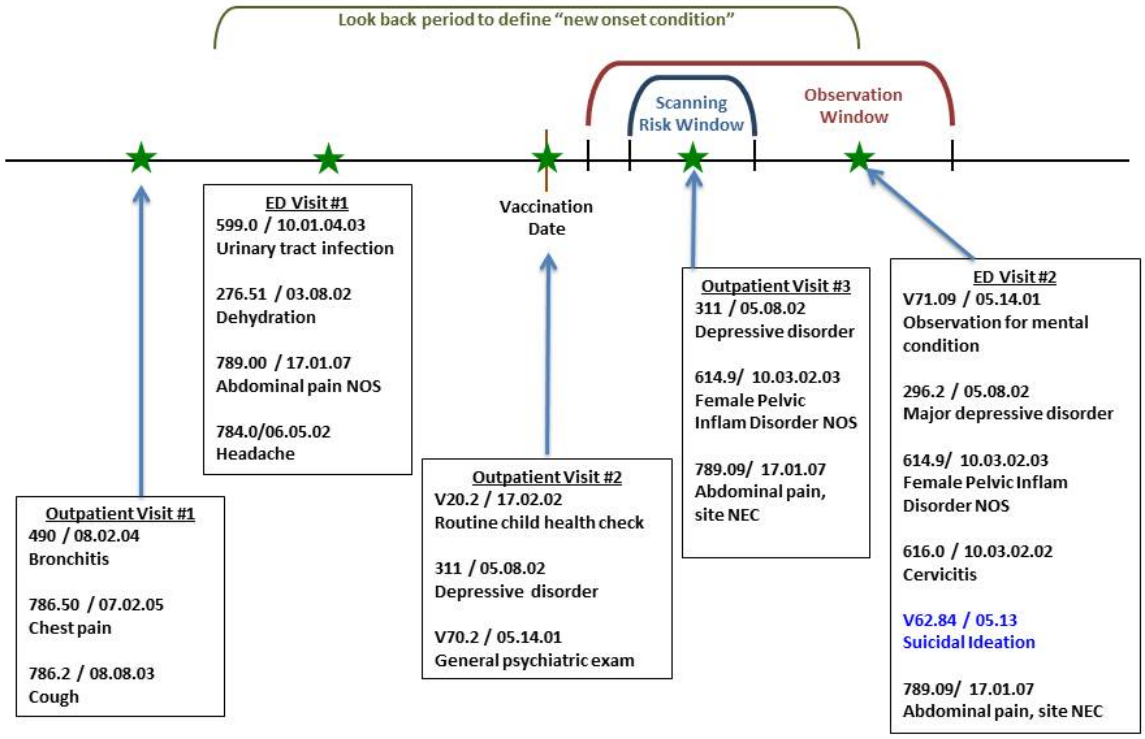
VIII. APPENDIX 1: COMPARISON OF TREESCAN METHODS

	Tree Scan Statistic				Tree-Temporal Scan	
	Cohort / Poisson Model		Self-Control / Bernoulli Model		Self-Control	
	Unconditional	Conditional	Unconditional	Conditional	Unconditional	Conditional
DATA NEEDS						
Exposure definition	One vaccine or multiple vaccines with any AND, OR, NOT logical operators. May differentiate between doses of the same vaccine.					
Outcome definition	Incident diagnosis, i.e., a diagnosis for which there was not the same or similar diagnosis in prior X days. Similar is defined as not being on the same 2nd, 3rd or 4th level of the tree.					
Adverse events (AEs) in risk interval	AE count in risk interval of, e.g., 1-2 or 1-28 days post vaccination				AE count in follow-up period of, e.g., 1-56 days post-vaccination, with information about the exact number of days post-vaccination	
AEs in comparison group	Many options, e.g., AEs in unexposed pre-vaccination time, after risk window, and among non-vaccinated. Used to generate age-adjusted expected counts.		AE count in control interval of, e.g., 29-56 days post vaccination			
TREESCAN INPUT						
Tree structure	A set of ICD9 codes and MLCCS codes, with information about the parent of each one					
AEs in risk interval	For each ICD9 code, number of AEs in risk interval				For each ICD9 code, number of AEs by days post-vaccination	
AEs in control interval	N/a		For each ICD9 code, number of AEs in control interval			
Expected count	For each ICD9 code, expected AEs under the null		N/a		N/a	
NULL HYPOTHESIS	AEs are generated from the expected counts.	The relative risk of different AEs is determined by the relative magnitude of the expected counts, but the total number of AEs is fixed and non-random.	The probability of an event occuring in the risk versus control interval is proportional to the lengths of those intervals. For each ICD9 code, total number of AEs is fixed and non-random.	The probability of an event occuring in the risk versus control interval is proportional to the total number of AEs in those intervals summed over all ICD9 codes. For each ICD9 code, total number of AEs is fixed and non-random.	An AE occurs uniformly over the follow-up period, with equal probability on each day.	Irrespective of the ICD9 code, all AEs have the same probability of occurring on a specific day. The probability is equal to the total number of AEs on that day divided by the total number of AEs in the follow-up period.
ALTERNATIVE HYPOTHESIS	There is at least one leaf or one branch on the tree where there are more expected AEs than what is defined under the null hypothesis.					
TREESCAN ANALYSIS OPTIONS						
Common parameters	Input file names; Type of Scan (tree only or tree-time); Probability Model; Conditional Analysis; Number of Monte Carlo replications; Output options					
Model-specific parameters	None		Risk Interval Probability	None	Temporal window start and end times; Maximum and minimum temporal window length; Data time range (length of follow-up)	
Power evaluation	Yes, available			No, not yet available		
TREESCAN EXECUTION						
Scan the tree	Scan the tree to find most likely and secondary cuts/clusters.				Perform the temporal scan on each potential cut on the tree, to find most likely and secondary clusters.	
Random Monte Carlo data sets	Generated from expected counts, using Poisson distribution. Random data sets may have more or less total AEs than the real data.	Conditioned on total number of AEs , so that each random data set has exactly the same number of total AEs as the real data set. The random number of AEs for a particular ICD9 code is binomially distributed as Bin(n,p), where n is the total number of AEs and p is the expected count in the leaf divided by the total expected count summed over the whole tree.	Generated using the specified risk interval probability. For each ICD9 code, the sum of the AEs in the risk and control intervals will be the same in each random data set and the real data set. The total number of AEs in the risk interval, summed over all ICD9 codes, may be different in the random and real data sets.	For each ICD9 code, the sum of the AEs in the risk and control intervals will be the same in each random data set and the real data set. The total number of AEs in the risk interval summed over all ICD9 codes, C, is the same in each random data set and the real data. Randomization is conducted by randomly picking C of the AEs to be assigned to the risk interval.	For each ICD9 code, the sum of the AEs taken over all days will be the same in each random data set and the real data set. Independently of every other AE, an AE is assigned to a day using a uniform distribution where each day in the follow-up period is equally likely to be chosen.	For each ICD9 code, the sum of the AEs taken over all days will be the same in each random data set and the real data set. For each day, the total number of AEs on day X, summed over all ICD9 codes, is the same in each random data set and the real data. Randomization is conducted by randomly permuting the days and the ICD9 code pairings, keeping the marginals fixed.
INTERPRETATION						
Self-controlled	No		Yes			
Pre-defined risk interval	Yes				No	
Adjusts for multiple testing	Yes					
Adjusts for temporal variation common to all ICD9 codes	No	Yes	No	Yes	No	Yes

IX. APPENDIX 2: EXAMPLES OF FICTIONAL PATIENTS

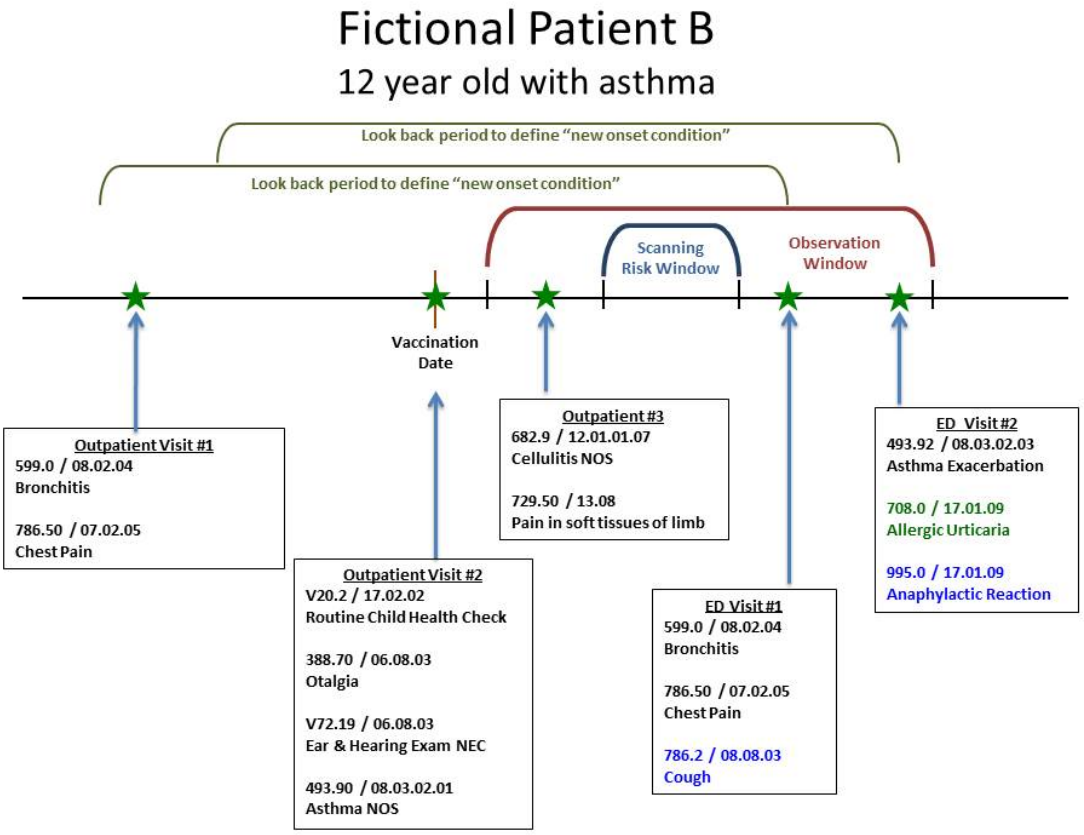
Fictional Patient A

18 year old with depression and pelvic inflammatory disease



This patient has 5 total healthcare visits in the period of analysis, 2 of which occur after vaccination (Outpatient visit #3 and ED visit #2). Of the 9 potential diagnosis codes post-vaccination, **only 1 will enter the Tree-Scan analysis (suicidal ideation from ED2).** All 3 diagnoses in OV3 are excluded because the analysis is restricted to the inpatient or ED settings. The 5 diagnoses from ED2 were excluded based upon several criteria designed to distinguish new onset conditions from pre-existing conditions, or acute exacerbations of pre-existing conditions.

Excluded Condition	Shares 3rd Level MLCCS with prior diagnosis in	Prior Diagnosis in Look back Period in
Observation for mental condition	OV2	
Major depressive disorder	OV3, OV2	
Female pelvic inflammatory disorder		OV3
Cervicitis	OV3	
Abdominal pain	ED1	OV3



This patient has 5 total healthcare visits in the period of analysis, 3 of which occur post-vaccination (OV3, ED1, ED2). Of the 8 potential diagnosis codes post-vaccination, **only 2 will enter the Tree-Scan analysis (cough from ED1 and anaphylactic reaction from ED2).** Both diagnoses in OV3 are excluded because the analysis is restricted to the inpatient or ED settings, even though "cellulitis" and "pain in soft tissues of limb" may represent actual vaccine-related events. The other 4 diagnoses were excluded as detailed below. Recall that each patient can contribute multiple diagnosis codes as long as they are on different 3rd level MLCCS branches. If two or more codes share the same 3rd level branch on the same visit, the program will select the rarest incident outcome. On ED2, "allergic urticaria" and "anaphylactic reaction" occupy the same 3rd level. Only anaphylactic reaction is selected because it is the rarer of these two diagnoses.

Excluded Condition	Shares 3rd Level MLCCS with prior diagnosis in	Shares 3rd Level MLCCS with diagnosis on same day	Prior Diagnosis in Look back Period in
Bronchitis			OV1
Chest pain			OV1
Asthma exacerbation	OV2		
Allergic urticaria		ED2	