# CBER SENTINEL SURVEILLANCE REPORT

# Application of TreeScan Data Mining Method to HPV9 (Gardasil 9) Vaccine

**Prepared by:**

Harvard Pilgrim Health Care, Boston, MA

Office of Biostatistics and Epidemiology, Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, MD

**September 3, 2019**

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.

# CBER Sentinel Surveillance Report
# Application of TreeScan Data Mining Method to HPV9 (Gardasil 9) Vaccine

**Table of Contents**

# I. ABSTRACT

## A. BACKGROUND

Tree-based scan statistics ("TreeScan") is an investigational method for vaccine and drug safety surveillance that involves application of diagnostic codes organized into a hierarchical "tree" structure. The method scans deidentified patient data and detects any statistically higher than expected clustering of cases within the hierarchically organized diagnoses that are within the post-exposure follow-up period, while adjusting for multiple testing. A strength of this approach compared to other approaches that pre-specify the health outcomes of interest, is that unexpected adverse events may be detected. However, a limitation of the method is the increased probability of false-positive results, particularly for chronic conditions, due in part to the short look-back period for pre-existing conditions and due to lack of adjusting for time-varying confounding. Use of the Treescan method to assess claims data in the CBER Sentinel Program was first explored for the Gardasil ("HPV4") vaccine. Expected adverse events (e.g., cellulitis) were found which did not require further investigation because they were known adverse events associated with the vaccine, and no false-positive results were generated. The purpose of the current study was to further explore the utilization of the TreeScan method for vaccines, and Gardasil 9 ("HPV9") was selected for this purpose.

## B. METHODS

Five Sentinel Data Partners provided data. The study population consisted of males and females 9-26 years of age who were vaccinated with HPV9 during mid-2015 to mid-2016. Eligible patient-doses were those preceded by at least 6 months of enrolled time (the look-back period) and followed by at least 56 days of enrolled time (the follow-up period). Hundreds of potential risk intervals between Days 1 and 42 post-vaccination were evaluated. The hierarchical reference tree of diagnoses used was the ICD-9-code-based Multi-Level Clinical Classification Software (MLCCS) of the Agency for Healthcare Research and Quality's Healthcare Cost and Utilization Project. Some of the diagnoses deemed very unlikely to be caused by vaccination (e.g., congenital conditions such as sickle cell disease) were excluded from the tree prior to analysis. A patient's diagnosis code was included only if it was observed in the inpatient or emergency department setting during the 56-day follow-up period and if there was no similar diagnosis for that patient in any setting during the 183 days (6 months) prior to the diagnosis in question. This incidence or "look-back" parameter setting of 6 months was chosen to address two competing priorities: to decrease the chances of including repeat visits for a single episode of illness (for which *more* required pre-exposure enrolled time is needed), and to increase sample size and statistical power (for which *less* required pre-exposure enrolled time is needed). Overall, the selection of the short (6-month) look-back period maximizes the sample size by making a larger population eligible faster and expedites the execution of the study., However, a short look-back period increases the probability of detecting many pre-existing conditions (particularly chronic conditions) as incident adverse events after vaccine exposure, thereby increasing the probability of a false-positive result for an adverse event.

The specific statistical method employed was the conditional self-controlled tree-temporal scan statistic, which looks for clustering of instances of the same or similar health outcomes (ICD codes) within the hierarchical tree of diagnosis codes as well as for temporal clustering of such cases within the follow-up period. The method adjusts for multiple testing by Monte Carlo simulation.

### C. RESULTS

The study performed one primary and two secondary analyses. The primary analysis consisted of all doses of HPV9 regardless of prior exposure to HPV2 or HPV4 vaccine and included 371,992 doses. The first secondary analysis consisted of all HPV9 doses without prior exposure to HPV2 or HPV4 vaccine and included 230,256 doses. The second secondary analysis consisting of only the first dose of HPV9 vaccine without prior exposure to HPV2 or HPV4 vaccine included 163,572 doses.

The primary analysis showed a total of 17 instances (16 patients) of "autistic disorder" in the Days 1-56 post-vaccination follow-up period. Of the 17 instances, there was a cluster of 10 cases in Days 27-32 which led to a statistically significant result (p=0.044). Upon examination of claims data going back further than the six months prior to the diagnosis date or review of medical records, it was determined 8 out of the 10 instances in the cluster and 6 of 7 instances outside the cluster (a total of 14 cases) had a pre-existing diagnosis for autism or autism-spectrum disorder prior to the HPV9 vaccination date. Medical records for the other 3 cases were not obtainable. Therefore, the statistically significant finding was determined to be a false-positive result, likely due to chance and the selected short look-back period in the study. Neither of the secondary analyses yielded a statistically significant result. In addition, the expected adverse event, namely, discomfort or irritation at the injection site did not show a statistically significant result possibly due to the limited sample size.

### D. CONCLUSION

This study explored the application of the TreeScan data-mining method to vaccines, using the HPV9 vaccine as a test case. The study incorporated approximately 372,000 doses of HPV9 and found one false positive alert for autism. This alert was ultimately determined to be a false alert through claims and medical record review. The alert resulted from a chance temporal clustering of cases of autism originally diagnosed years prior to HPV9 vaccination. The short "look-back" period contributed to this alert because pre-existing diagnoses were incorrectly detected as incident cases.

The study results suggest that application of the method may be limited for routine vaccine safety surveillance because of certain limitations. First, the false positive result seen here illustrates that with the TreeScan screening method, no conclusion about causality can be drawn from statistically significant results. Further investigation, including validation of cases, is required. Such validation requires access to patients' claims profiles and medical records which may be difficult to obtain in some settings. Second, the method does not allow customization of look-back periods that vary in duration for different outcomes, resulting in incorrect classification of pre-existing diagnoses as incident and an increased potential for false-positive findings for some outcomes. Third, the lack of adjustment for time-varying confounding may be associated with bias. Fourth, the method requires a large sample size to have sufficient statistical power to evaluate potential rare adverse events associated with vaccines and to assess a large number of adverse events simultaneously. Finally, exclusion of outpatient care setting from the study observation period could result in missing adverse reactions that are primarily treated in an outpatient setting.

## II. INTRODUCTION

Tree-based scan statistics (hereafter referred to as 'TreeScan') are a data-mining method that can potentially be used for vaccine and drug safety surveillance to look for a wide range of medically attended adverse events; the method is based on scanning deidentified patient data for thousands of diagnosis codes as a proxy for health outcomes.[1-4] A strength of this approach compared to vaccine and

drug safety studies that pre-specify health outcomes of interest is that unknown or unexpected adverse events may be detected. A limitation of the TreeScan method compared to conventional studies that focus on one or a few health outcomes of interest is that it is not possible to adjust for all possible confounders, as they vary by outcome. Additionally, the analysis parameters must be set the same way for all outcomes because many outcomes are evaluated. For example, an "incident diagnosis" is defined uniformly as the first to occur within a set period of time specified by the investigator, such as 6 months or 1 year. The period selected may be too short for certain chronic conditions, the first diagnosis of which may have occurred years prior, but requiring a long look-back period in which to check for prior occurrences of diagnoses has the drawback of reducing sample size and statistical power. Thus, if there is a finding of a higher number of cases of an outcome than expected (a "statistically significant result"), the finding must be carefully evaluated by reviewing claims profiles and medical charts to determine whether the condition was truly present and, if so, whether the case was truly of new onset as opposed to chronic. If it cannot be established that the statistically significant result is a false alert, then further evaluation using epidemiological study designs that can control for potential confounders pertinent to the specific outcome in question may be needed to confirm the result as a real association. Without such further evaluation, no conclusion about causality can be drawn; in effect, the TreeScan method is an investigational screening tool aimed at identifying possible adverse events that merit further careful pharmacoepidemiological investigation.

Previously in a pilot project, the specific variant of the TreeScan method known as the conditional self-controlled tree-temporal scan statistic was applied to Sentinel/PRISM's electronic health care data[5, 6] in order to evaluate its utility to assess for any short-term adverse events after immunization with quadrivalent human papillomavirus vaccine (Gardasil (Merck); "HPV4").[7] This vaccine was indicated for use in females and males 9-26 years of age as a 3-dose series, on a schedule of 0, 2, and 6 months.[8] The tree-temporal method not only simultaneously evaluates up to several thousand diagnosis codes but also simultaneously evaluates a large number of potential risk windows, adjusting for the multiple testing inherent in the many health outcomes and risk windows examined.

In the pilot, which included approximately 1.9 million doses of HPV4, two classes of statistically significant results appeared. One was in the category of "cellulitis and abscess of the arm." Cellulitis is a known adverse event listed in the HPV4 package insert,[8] and therefore the result was not investigated further. The other class of statistically significant results was in the more general category of "other complications of surgical and medical procedures." On investigation, it was determined that most of these "other complications" cases had claims for conditions already known to be associated with the vaccine and that the remainder had either non-serious or non-vaccine-related conditions. Thus, neither set of significant results represented a true new safety signal.

The nine-valent human papillomavirus vaccine (Gardasil 9 (Merck); "HPV9") was approved in December 2014 for females aged 9-26 years and males aged 9-15 years; the approved age range for use was extended to males aged 16-26 years in December 2015. During the period of available HPV9 data for this study, there was one FDA-approved dose schedule for HPV9: a series of 3 doses administered at 0, 2, and 6 months. In this study, we aimed to further examine the feasibility of application of the TreeScan method to vaccines by using the HPV9 vaccine as another test case in females and males 9-26 years of age.

# III. METHODS

## A. STUDY POPULATION AND FOLLOW-UP PERIOD

Claims and administrative data in the Sentinel Distributed Database were obtained from Aetna, Harvard Pilgrim Health Care, HealthCore, Humana, and Optum. Two sites contributed data from 1/1/2007, one site from 6/1/2007, and 2 sites from 1/1/2008 to the study. The data prior to 12/10/2014, which was the date of licensing of HPV9 vaccine, were included only to determine prior exposure of patients to bi-valent (Cervarix (GlaxoSmithKline); "HPV2") or quadrivalent HPV vaccines for the secondary analyses. The covered period for exposure to HPV9 vaccine in this study is shown in **Table 1**.

**Table 1. Date ranges of HPV9 available data, by Data Partner (masked and in a different order than above)**

| Data Partner | |
|---|---|
| A | Summer 2015 - 6/30/2016 |
| B | Summer 2015 - 7/31/2016 |
| C | Summer 2015 - 4/30/2016 |
| D | Summer 2015 - 3/31/2016 |
| E | Summer 2015 - 6/30/2016 |

The study population was females and males 9-26.99 years of age as of HPV9 vaccination date. The follow-up period was Days 1-56 after vaccination. The unit of analysis in this study is the individual doses of each HPV9 vaccine administered and not the individual patients who received the vaccines. Due to the HPV9 vaccination schedule, many patients contributed more than one unit (dose) to the analysis because they had received more than one dose of vaccine. For doses to be eligible for analysis, at least 183 days of pre-vaccination enrolled time and at least 56 days of post-vaccination enrolled time were required. Doses and patients that did not meet these criteria were excluded from the study.

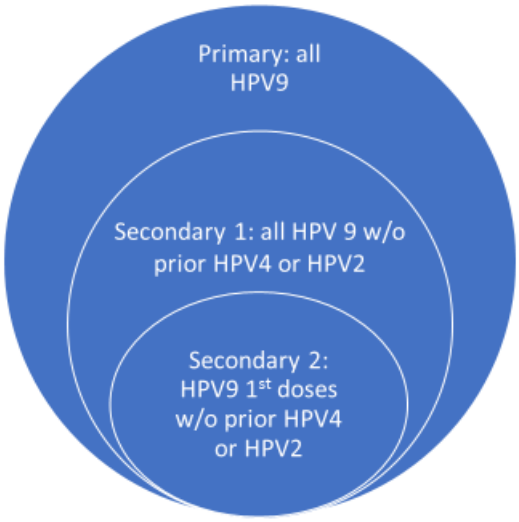
## B. PRIMARY AND SECONDARY ANALYSES

The primary analysis considered *all* doses of HPV9 without distinguishing among them and regardless of previous exposure to HPV2 or HPV4 vaccines. Doses of HPV9 occurring within 42 days of a previous dose of HPV9 were excluded. This was done to avoid overlapping risk windows and is presumed to have had minimal impact on the number of doses included, considering that descriptive statistics from an earlier PRISM HPV4 study[9] showed that only 1.5% of HPV4 Dose 2 were given within 42 days of Dose 1.

There were two secondary analyses, comprising subsets of the doses included in the primary analysis, as shown in **Figure 1**.

1) Including *all* doses of HPV9 that were *not preceded by either HPV4 or HPV2*. In other words, doses and patients who had prior exposure to HPV4 or HPV2 were excluded from the analysis. As in the primary analysis, we did not distinguish among Doses 1, 2, or 3 of HPV9, and doses of HPV9 occurring within 42 days of a previous dose of HPV9 were excluded.
2) Including only the *first* dose of HPV9 that was not preceded by either HPV4 or HPV2.

**Figure 1. Schematic representation of the relationships among the primary and secondary analyses**




To check for prior HPV vaccination in both secondary analyses, we looked back from the HPV9 dose in question through the maximum amount of available enrolled time. Ascertainment of previous HPV2 or HPV4 may not have been complete.

## C. STUDY VACCINE

HPV9 vaccination was identified by means of CPT code 90651. For the secondary analyses, we also made use of the CPT codes 90649 (HPV4) and 90650 (HPV2), in order to identify and exclude HPV9 doses that were preceded by these HPV vaccines. Because vaccinations are sometimes entered into claims data as NDC codes, we used those codes as well; they are listed in **Table A1** of the appendix.

## D. RISK AND COMPARISON WINDOWS

We evaluated all potential risk windows that were at least 2 days long, were at most 28 days long, started between 1 and 28 days after vaccination, and ended between 2 and 42 days after vaccination. The comparison or control period consisted of those days within the Days 1-56 follow-up period that were not in the risk window being evaluated.

## E. HIERARCHICAL DIAGNOSIS TREE AND MAPPING ICD-10-CM TO ICD-9-CM CODES

All ICD-9-CM diagnoses are represented in 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 ). It is a hierarchical system with four diagnosis levels, although on some branches there may only be two or three levels. The first and broadest level identifies 18 body systems, while the entries at the finest level contain one or multiple ICD-9-CM codes. For example, convulsions is a third level classification without a fourth level and for which there are five different ICD-9-CM codes, as shown in **Table 2**.

**Table 2. Excerpt from MLCCS hierarchical tree: Diagnoses related to convulsions**

| Node | 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 |

We used the MLCCS tree, although, as was the case for the HPV4 pilot, we excluded some ICD-9-CM/ICD-10-CM codes from it and therefore from the analysis, for example, those representing:

- Some 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 were dealing with in this study, such as cancer
- Most infectious diseases with an identified organism (e.g., typhoid fever, tuberculosis, shigellosis)
- 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.

Outcomes were identified by their ICD-9-CM or ICD-10-CM code. In order to use the MLCCS tree, we mapped ICD-10-CM codes to ICD-9-CM codes, using the Centers for Medicare & Medicaid Services' General Equivalence Mappings (GEMs). The mapping scheme was finalized prior to extracting the data and conducting the analyses.

## F. INCIDENT DIAGNOSES OF INTEREST

Incidence criteria were applied to minimize inclusion of repeat visits for a single episode of illness or for ongoing chronic conditions. A diagnosis was included if it was observed in the inpatient or emergency department 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 183 days (6 months) prior to the diagnosis in question. We included only outcomes recorded at inpatient stays or emergency department visits in order to capture only the most serious types of outcomes, aiming to avoid a potentially large number of statistically significant results for minor conditions that would require further investigation.

## G. TREE-TEMPORAL SCAN STATISTIC

With the tree-temporal scan statistic, multiple temporal scan statistics are performed, one for each of the many clinical outcomes and groups of related clinical outcomes (i.e., leaves and branches of the tree). Multiple potential risk windows are evaluated, comparing the number of events within the risk window with what would be expected by chance if they were randomly distributed over time. Under the null hypothesis, there is no unusual clustering of events within any branch or time interval. Under the alternative hypothesis, there is at least one branch of the tree for which there is a temporal cluster of events during some time interval, above and beyond any *general* temporal patterns in healthcare-seeking behavior. In a *conditional* analysis, such as the one we undertook for the current study, we condition not only on the number of events observed in each node of the tree during the whole follow-

up period but also on the total number of events 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.

The method adjusts for the multiple testing entailed in evaluating the many branches and time intervals. Each time interval is evaluated on each of the branches, and with the approximately 7300 nodes (i.e., outcome categories, whether first, second, third, fourth, or fifth level, which include, for example, the codes listed in **Table 2**) on the tree and the 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 relatively modest.

To implement the conditional tree-temporal scan statistic, we calculate a Poisson generalized log likelihood ratio test statistic for each tree node and time interval. Let n be the number of events in the node, let c be the number of those node events that are also in the time interval, let z be the number of events in the time interval summed over the whole tree, and let C be the total number of events in the tree. The number of events in the cluster, c, is then contrasted with the expected number of events in the cluster under the null hypothesis, which is u=nz/C. When u>0, the test statistic is calculated as

$$T = \{c \times \ln[c/u]\} + \{(C\text{-}c) \times \ln[(C\text{-}c)/(C\text{-}u)] \times I(c>u)\}$$

where I() is the indicator function. I(c>u) is 1 when there are more events 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 events, that is, the cluster that is least likely to have occurred by chance.

The distribution of the test statistic is not known analytically, and thus 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 events as the real dataset, and each day after vaccination has the same number of events when summed over all nodes. The only thing that varies is the pairing of the nodes and times, which is randomized using a permutation approach. 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 the 99,999 random datasets, and we note its rank. For example, if it has the fifth highest test statistic, its rank, R, is 5. Note that the most likely cut will be on a different branch in the different datasets, and 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 at the alpha=0.05 level. If the null hypothesis is true, we have a 5% probability of falsely rejecting the null and a 95% probability of not having any alert anywhere on the tree.

# IV. RESULTS

The numbers of doses of HPV9 captured in the primary and two secondary analyses are shown by age group in **Table 3**. In the primary analysis, which included all HPV9 doses, 371,992 doses of HPV9 were captured. The secondary 1 and 2 analyses included 230,256 and 163,572 doses of HPV9 vaccine, respectively.

**Table 3. Number of doses of HPV9 vaccine in the various analyses, by age group**

| | Primary, including all HPV9 doses regardless of prior HPV4 or HPV2 | | Secondary 1, all HPV9 doses without prior HPV4 or HPV2 | | Secondary 2, HPV9 first dose without prior HPV4 or HPV2 | |
|---|---|---|---|---|---|---|
| Age group | Count | Proportion | Count | Proportion | Count | Proportion |
| 9-11 | 64,462 | 17.3% | 51,532 | 22.4% | 37,833 | 23.1% |
| 12-14 | 169,558 | 45.6% | 103,635 | 45.0% | 71,505 | 43.7% |
| 15-17 | 93,043 | 25.0% | 50,567 | 22.0% | 36,090 | 22.1% |
| 18-20 | 27,969 | 7.5% | 13,737 | 6.0% | 10,135 | 6.2% |
| 21-23 | 9,474 | 2.5% | 5,747 | 2.5% | 4,380 | 2.7% |
| 24-26 | 7,486 | 2.0% | 5,038 | 2.2% | 3,629 | 2.2% |
| All | 371,992 | 100.0% | 230,256 | 100.0% | 163,572 | 100.0% |

In the primary, most inclusive analysis, one statistically significant cluster was detected (**Table 4**). It was for the diagnosis of "autistic disorder—current," corresponding to ICD-9-CM code 299.00 and ICD-10-CM code F84.0, within the "pervasive developmental disorders" branch of the tree. There was a total of 17 cases that had "incident" (first-in-183-days) diagnosis codes of 299.00 or F84.0 in the full Days 1-56 follow-up period, 2 of which happened to be in the same patient after different doses of HPV9 (longer than 183 days apart). Ten of the 17 cases contributed to the statistically significant result in that the diagnoses codes were clustered within Days 27-32, the period of increased risk detected by TreeScan within the 56-day follow-up period. The p-value for this significant result was 0.044. No statistically significant results were seen for the higher-level "pervasive developmental disorders" branch, which, in addition to the 17 cases, included 6 cases with a 299.xx/F84.x diagnosis code other than 299.00/F84.0 in the 56-day follow-up period. (All the codes belonging to that branch are in **Table A2** of the appendix.) None of these 6 additional cases had these diagnoses codes recorded during the Days 27-32 period.

**Table 4. Excerpt of tree-temporal scan statistical analysis results for all eligible HPV9 doses (primary analysis), showing the statistically significant cluster and other relevant results**

| Node | Node Name | Tree Level | Node Cases | Risk Window Start | Risk Window End | Cases in Risk Window | Expec-ted Cases | Test Statistic | P-value |
|---|---|---|---|---|---|---|---|---|---|
| 05.06.03.00 | Pervasive developmental disorders | 4 | 23 | 27 | 32 | 10 | 2.33 | 6.89 | 0.604 |
| ...29900 | Autistic Disorder-Current | 5 | 17 | 27 | 32 | 10 | 1.72 | 9.30 | 0.044 |

As it is expected for TreeScan activities as a screening tool and data mining method, the study had to further investigate whether the statistically significant cluster was a real alert or a false-positive result due to the study specifications, analytic errors, or data anomalies. The following steps were taken to investigate the validity of the observed statistically significant cluster with respect to the outcome autism.

After executing the data extraction program at the Data Partners and running the TreeScan analysis at the Sentinel Coordinating Center, we ran a program at the Data Partners in order to freeze (save a copy ) data behind their firewalls for the 17 cases with a 299.00 or F84.0 diagnosis in the 56 days after HPV9.[10] Using the saved data, the observed result was explored by generating a claims profile (a list of all available medical claims) for the period from 56 days before through 56 days after HPV9 vaccination for each of the 17 cases (16 patients) and conducting a review in which a physician on the team participated. This relatively short period for patients' claims data review was selected in conformity with the principle of obtaining the minimum necessary patient-level information. Medical records were sought for 7 of 17 cases, of which 4 were obtained and 3 were not obtainable.

A list of the 17 cases (16 patients) in the full 56-day follow-up period, with findings from claims profiles and/or medical records, is presented in **Table A3**. Of the 10 cases contributing to the statistically significant result (i.e., within Days 27-32, indicated by a red box in **Table A3**), 7 (Cases 5, 6, 7, 9, 10, 12, 13) had a code for pervasive developmental disorders—either "autistic disorder-current" or "other pervasive developmental disorder-current"—more than 183 days prior to their alerting 299.00/F84.0 code and therefore prior to HPV9 vaccination. Of the remaining 3 cases, 1 (Case 8) had a mention of autism prior to HPV9 exposure in the medical record. No medical records could be obtained for the other 2 (Cases 11 and 14), although one patient's claims data suggested that the reason for the medical encounter on the day of the alerting 299.00/F84.0 code involved an unrelated acute condition. Overall, it was determined that 8 of 10 cases contributing to the observed statistically significant cluster had been diagnosed with autism prior to HPV9 vaccine exposure. For the 2 remaining cases within the cluster, it remained unknown whether the autism was diagnosed prior to HPV9 exposure.

Regarding the 7 cases within the 56-day follow-up period but outside of the cluster period of Days 27-32, 3 (Cases 3, 4, 16) had a code for pervasive developmental disorders—either "autistic disorder-current" or "other pervasive developmental disorder-current"—more than 183 days prior to their alerting 299.00/F84.0 code in their claims history and therefore prior to HPV9 vaccination. Three others (Cases 1, 2, 15) had "autism spectrum disorder" or "autism" or "past psychiatric history of pervasive developmental disorder and autism" noted in their medical charts. The chart for the last of the 7 cases outside of the cluster period (Case 17) was not obtainable. Thus, it was determined that 6 of 7 cases outside of the cluster had been diagnosed with the outcome in question prior to HPV9 vaccination.

In summary, further investigation into the statistically significant cluster within the "pervasive developmental disorders" branch of the tree showed that all but 2 cases had received the diagnosis prior to the HPV9 vaccine exposure. Medical records could not be obtained to investigate whether autism was diagnosed prior to HPV9 vaccine exposure for these remaining 2 cases. Although the cluster included incident cases as defined by the protocol, further investigation showed that most cases were in fact not incident. Hence, the statistically significant cluster was resolved as a false-positive alert in this study.

In the primary analysis, no statistically significant results appeared for "cellulitis and abscess of arm" or "other complications of surgical and medical procedures," conditions for which there were statistically significant results in the earlier TreeScan analysis of 1.9 million first doses of HPV4.[7] The result for "other complications of surgical and medical procedures" was close to statistically significant (p=0.059), however.

This study included two secondary analyses, one of which analyzed patients who received any dose of HPV9 vaccine without a prior exposure to HPV4 or HPV2 vaccine. The other secondary analysis included only exposure to the first dose of HPV9 vaccine without a prior exposure to HPV4 or HPV2 vaccine. There were no statistically significant results in either of the secondary analyses.

In summary, the tree-temporal scan statistical analysis conducted for the approximately 372,000 eligible doses of HPV9 found one statistically significant cluster, for "autistic disorder-current" in Days 27-32 after vaccination. This was determined to be a false positive result, a consequence of (a) the inclusion of first-in-6-months diagnoses rather than only first-ever diagnoses and (b) a chance temporal clustering of such cases. Of the 10 cases in the cluster, 8 had pervasive developmental disorders diagnoses in their claims history or medical record prior to HPV9 vaccination. Medical records could not be obtained to investigate whether autism was diagnosed prior to HPV9 vaccine exposure for the remaining 2 cases.

# V. DISCUSSION

The TreeScan method is an investigational data-mining and screening tool to evaluate possible associations between exposure to medical products and a large number of unspecified adverse events (diagnosis codes) without adjusting for all the potential confounders. This study was conducted to investigate the feasibility of application of TreeScan method to vaccines, and HPV9 vaccine served as a test case in the study. In the primary analysis of the current study, a statistically significant result arose concerning the "autistic disorder" outcome. On examination of claims data and medical records, the result was determined to be a false alert related to using a parameter (look-back period) setting that defined (for all outcomes) "incident" diagnoses as the first such diagnosis in the prior 6 months. Defining incidence in this way, instead of as first ever, led to cases that had a post-HPV9 vaccine autism diagnosis with no such diagnosis in the prior 6 months being labeled as "incident," when there was an autism diagnosis longer than 6 months prior to the vaccine exposure. The two secondary analyses did not generate any statistically significant results.

The primary analysis of this data-mining study incorporated 372,000 doses of HPV9 vaccine and evaluated over 6,000 diagnosis codes. The level of statistical power was different for different diagnosis codes depending on the incidence of the outcome in the population. Overall, it had the statistical power to observe large increases in risk for many outcomes, if such risks had existed, and to rule out safety problems with large effect sizes. However, the study did not have sufficient power to observe modest increases in risk for many outcomes.

The statistically significant results that appeared in the TreeScan analysis of HPV4[7], namely "cellulitis and abscess of arm" and "other complications of surgical and medical procedures" did not appear in the primary or secondary TreeScan analyses of HPV9. However, it cannot be concluded from comparing the HPV4 and HPV9 TreeScan analyses that these two HPV vaccines differ with respect to their risk of adverse events. The current primary analysis included 372,000 doses of HPV9, constituting only one-fifth the sample size of the HPV4 analysis (1.9 million doses). The secondary HPV9 analyses included fewer doses. Thus, the statistical power of the HPV4 analysis and its ability to detect increased risks of potential adverse events was considerably greater.

The TreeScan method has strengths and limitations. One of its strengths is its design to examine the potential association between an exposure and a large number of outcomes in the form of diagnosis codes including unexpected outcomes while adjusting for multiple testing. Also, the capability to scan a large number of potential risk windows and risk window lengths is an important strength.

The most important limitation of this method is that it is a screening method and thus requires validation of any positive results. While self-controlled, the method does not control for time-varying confounding and does not allow for customization of some parameter settings according to diagnosis. If a statistically significant result emerges from a TreeScan analysis, it must be further investigated to confirm its validity and rule out confounding, prevalent instead of incident cases, data anomalies, analytical errors, or unexpected interactions; the statistically significant results do not imply or provide any evidence of either association or causality between the exposure and outcome. Validation of an alert, depending on what the outcome is, would require access to medical history of patients in claims data (claims profiles) as well as their medical charts in different health care facilities. Different patients have different lengths of claims coverage in the available databases, (sometimes quite short) and access to claims profiles is not readily available, requiring a long time and large budget. Access to medical charts in turn is even more difficult, time-consuming, and expensive. Even when a medical chart can be obtained, the records may not be available for a long enough period of the patient's medical history to validate the presence or absence of certain outcomes, particularly chronic outcomes. As a result, the validation of statistically significant results from TreeScan analysis is very challenging, and the validation cannot be conducted in a timely manner to reasonably contribute to the routine assessment of the safety profile of vaccines.

TreeScan scans data on more than 6,000 diagnosis codes and uses the same parameter settings for all outcomes, rather than customizing the settings for each. The "incidence" criterion parameter was set at first-in-6-months in order to increase the statistical power of the method to detect a smaller level of increased risk or increased risks of rare outcomes. However, this selection in turn increases the chances of false positive results resulting from misclassification of pre-existing cases as incident cases. If the method instead selected incidence as first-in-1-year or first-in-5-years, it would require the corresponding lengths of prior enrollment and therefore would reduce the sample size, statistical power, and the ability to detect lower risk for all outcomes. Selection of a first-in-6-months incidence criterion (look-back period) is arbitrary, and it impacts the level of false-positive and false-negative results obtained with the TreeScan method. .

Additionally, there are some limitations of the TreeScan method inherent to the tree-based scan statistic or related to the way in which the statistic is implemented. For instance, we considered only risk windows beginning during Days 1-28 post vaccination and ending during Days 2-42 post vaccination. Thus, we did not evaluate the risk of adverse reactions occurring on the actual day of vaccination, Day 0, so we would have missed most cases of vaccine-associated anaphylaxis or syncope if any had occurred. The reason for excluding Day 0 was that diagnosis codes for conditions that precede a vaccination visit are often entered on the day of the visit and can produce spurious alerts. The choice of Day 42 as the

last possible day of any risk window evaluated means we could only detect adverse reactions manifesting themselves within 6 weeks after vaccination, i.e., relatively acute outcomes. Other reasons for not considering longer risk windows was to avoid overlapping risk windows of consecutive doses and to minimize time-varying confounding.

Although the tree-temporal scan statistic is a self-controlled method and thus automatically adjusts for all time-invariant confounders, it does not adjust for time-varying confounders. For example, the uptake of all HPV vaccines 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 theoretically could have produced a bias either toward or away from alerting, depending on the outcome in question.

Although relying on electronic healthcare databases, claims and administrative data in the current study, has key advantages such as the efficient capture of the healthcare experiences of a large patient population, there are fundamental drawbacks to using administrative claims data for vaccine safety surveillance, such as variability in coding practices for one or more of the many outcomes evaluated and the fact that a diagnosis date is not necessarily the same as the date of onset of symptoms.

Exclusion of outpatient care setting from the study observation period could result in missing adverse reactions that are primarily treated in an outpatient setting.

Since TreeScan aims to examine the association between one medical product and many outcomes, and the incidence of each outcome in the population under study may be different, regardless of the sample size of a study, the method provides a different statistical power for detection of different levels of increased risk for each outcome. The rarer the outcome, the less power this method has to detect a lower level of risk for the outcome of interest. Many potential adverse events associated with vaccines are rare. Hence, it is prudent to reserve any potential use of TreeScan for evaluating vaccines to instances in which a very large amount of exposure will have accrued and the statistical power will be higher for more outcomes.

Currently, the TreeScan method uses the same tree (inclusion of the same diagnosis codes) for all vaccines. However, since the safety profile of different vaccines is likely to be different, it may be helpful to assemble and prune the tree specifically for the vaccine being evaluated, taking into consideration such factors as the known safety profile and the age groups in which it is indicated.

# VI. CONCLUSION

The TreeScan method was utilized in this data mining study to investigate its feasibility to the study of vaccines, with the HPV9 vaccine as a test case to screen for unexpected adverse events in the inpatient and ED care settings. The study did not ultimately indicate an increased risk of any adverse events, although one finding required additional investigation and validation. Since this data mining method is utilized as a large-scale screening tool for detection of potential and unexpected adverse events, at least one of the parameters of the software (i.e., definition of incident diagnosis) was set to improve sensitivity, which in turn resulted in a higher probability of detecting a false-positive result, as was demonstrated in this study. The false-positive result detected in the analysis arose from a chance temporal clustering of non-incident diagnosis codes for autism. Most cases that contributed to the false-positive result were determined to have been first diagnosed prior to HPV9 vaccination. Due to the small sample size, this study had limited statistical power to detect small increased risk for many rare adverse

events. However, the study provides some additional evidence for the continued safety of the HPV9 vaccine in the post-market setting.

The finding of the false-positive result in this study sheds light on the limitations of this method specifically when applied to evaluating the safety profile of vaccines. Challenges associated with access to patients' claims profiles and medical charts to validate results from this method, as well as limited statistical power for the detection of a small increased risk for potential rare adverse events, among other limitations, cast doubt on the feasibility of this method for vaccine safety surveillance.

# VII. REFERENCES


1. 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.
2. Kulldorff M, Dashevsky I, Avery TR, et al. Drug safety data mining with a tree-based scan statistic. Pharmacoepidemiol Drug Saf 2013;22:517-23.
3. Kulldorff M, Fang Z, Walsh SJ. A tree-based scan statistic for database disease surveillance. Biometrics 2003;59:323-31.
4. Kulldorff M. TreeScan User Guide, Version 1.1. 2015. Available at http://www.treescan.org
5. Nguyen M, Ball R, Midthun K, et al. The Food and Drug Administration's Post-Licensure Rapid Immunization Safety Monitoring program: strengthening the federal vaccine safety enterprise. Pharmacoepidemiol Drug Saf 2012;21(Suppl 1):291-297.
6. Baker MA, Nguyen M, Cole DV, et al. Post-licensure Rapid Immunization Safety Monitoring program (PRISM) data characterization. Vaccine 2013;31(Suppl 10):K98-K112.
7. Yih WK,* Maro JC,* Nguyen M, et al. (*co-primary). Assessment of quadrivalent human papillomavirus vaccine safety using the self-controlled tree-temporal scan statistic signal-detection method in the Sentinel System. Am J Epidemiol 2018. https://doi.org/10.1093/aje/kwy023.
8. Merck. Gardasil package insert. http://www.fda.gov/downloads/BiologicsBloodVaccines/Vaccines/ApprovedProducts/UCM111263.pdf. Revised April 2015. Accessed June 6, 2017.
9. 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.
10. Cole DV, Kulldorff M, Baker MA, et al. Infrastructure for evaluation of statistical alerts arising from vaccine safety data-mining activities in Mini-Sentinel. Final Report to FDA, publicly posted July 22, 2016. Available at: https://www.sentinelinitiative.org/sites/default/files/Methods/Mini-Sentinel_PRISM_Data-Mining-Infrastructure_Report_0.pdf.

# VIII. APPENDIX: SUPPLEMENTARY MATERIAL

**Table A1. NDC codes used to identify HPV vaccination**

| VacCode | VacCodeDesc |
|---|---|
| 00006404500 | HUMAN PAPILLOMAVIRUS VACCINE, QUADRIVALENT/PF |
| 00006404501 | HUMAN PAPILLOMAVIRUS VACCINE, QUADRIVALENT/PF |
| 00006404541 | HUMAN PAPILLOMAVIRUS VACCINE, QUADRIVALENT/PF |
| 00006410901 | HUMAN PAPILLOMAVIRUS VACCINE, QUADRIVALENT/PF |
| 00006410906 | HUMAN PAPILLOMAVIRUS VACCINE, QUADRIVALENT/PF |
| 00006410909 | HUMAN PAPILLOMAVIRUS VACCINE, QUADRIVALENT/PF |
| 00006410931 | HUMAN PAPILLOMAVIRUS VACCINE, QUADRIVALENT/PF |
| 58160083011 | HUMAN PAPILLOMAVIRUS VACCINE, BIVALENT/PF |
| 58160083032 | HUMAN PAPILLOMAVIRUS VACCINE, BIVALENT/PF |
| 58160083046 | HUMAN PAPILLOMAVIRUS VACCINE, BIVALENT/PF |
| 58160083052 | HUMAN PAPILLOMAVIRUS VACCINE, BIVALENT/PF |
| 00006410902 | HUMAN PAPILLOMAVIRUS VACCINE, QUADRIVALENT/PF |
| 00006411903 | HUMAN PAPILLOMAVIRUS VACCINE, 9-VALENT/PF |
| 00006412102 | HUMAN PAPILLOMAVIRUS VACCINE, 9-VALENT/PF |
| 00006411901 | HUMAN PAPILLOMAVIRUS VACCINE, 9-VALENT/PF |
| 00006411902 | HUMAN PAPILLOMAVIRUS VACCINE, 9-VALENT/PF |
| 00006412101 | HUMAN PAPILLOMAVIRUS VACCINE, 9-VALENT/PF |

**Table A2. Full set of diagnosis codes belonging to or mapping to the 05.06.03.00, "pervasive developmental disorders" branch of the MLCCS tree**

| ICD-9-CM Code | ICD-9-CM Diagnosis Name | ICD-10-CM Code | ICD-10-CM Diagnosis Name |
|---|---|---|---|
| 29900 | AUTISTIC DISORD-CURRENT | F84.0 | Autistic disorder |
| 29901 | AUTISTIC DISORD-RESIDUAL | | |
| 29910 | CHILD DISINTEGR DIS-CUR | F84.3 | Other childhood disintegrative disorder |
| 29911 | CHLD DISINTEGR DIS-RESID | | |
| 29980 | OTH PERVAS DEVEL DIS-CUR | F84.5 or F84.8 | Asperger's syndrome or Other pervasive developmental disorders |
| 29981 | OTH PERV DEVEL DIS-RESID | | |
| 29990 | PERVAS DEVEL DIS NOS-CUR | F84.9 | Pervasive developmental disorder, unspecified |
| 29991 | PERV DEVEL DIS NOS-RESID | | |

**Table A3. Some features of the 17 cases of first-in-183-days ICD code 299.00/F84.0 in the 56 days after HPV9 vaccination, including findings from claims profiles and/or medical records. The 10 cases in the red box, representing those with the 299.00/F84.0 alerting code during Days 27-32 after HPV9 vaccination, are the ones that contributed to the statistically significant result.**

| Case[1] | Age in Years at Event[2,3] | Time-to-Event in Days[4] | Nearest Prior Related ICD Code[5,6] | Medical Record Finding for Cases with No Prior Related ICD Code[5] | Pre-HPV9 Autism-Consistent Codes in Claims | Continuous Enrollment Length in Months Through the Event[2,7] | Number of Instances of Prior Related ICD Codes[5] |
|---|---|---|---|---|---|---|---|
| 1 | 17 | 8 | None | "Autism spectrum disorder" noted in Patient Active Problem List on day of HPV9 | No | 7 | N/A |
| 2 | 12 | 16 | None | History of autism noted | Yes | 8 | N/A |
| 3 | 15 | 17 | 299.00 | N/A | N/A | 13 | 4 |
| 4 | 13 | 22 | 299.00 | N/A | N/A | 97 | 6 |
| 5 | 17 | 27 | 299.00 | N/A | N/A | 21 | 27 |
| 6 | 15 | 27 | 299.00 | N/A | N/A | 89 | 20 |
| 7 | 15 | 28 | 299.00 | N/A | N/A | 6 | 2 |
| 8 | 11 | 28 | None | History of autism noted | No | 10 | N/A |
| 9 | 15 | 29 | 299.80 | N/A | N/A | 91 | 17 |
| 10 | 15 | 29 | 299.80 | N/A | N/A | 22 | 1 |
| 11 | 14 | 29 | None | Chart not obtainable | No | 92 | N/A |
| 12 | 17 | 30 | 299.00 | N/A | N/A | 46 | 2 |
| 13 | 15 | 32 | 299.00 | N/A | N/A | 101 | 5 |
| 14 | 17 | 32 | None | Chart not obtainable | Yes | 93 | N/A |
| 15 | 16 | 39 | None | Past psychiatric history of pervasive developmental disorder and autism noted | Yes | 8 | N/A |
| 16 | 12 | 55 | 299.80 | N/A | N/A | 100 | 1 |
| 17 | 12 | 55 | None | Chart not obtainable | Yes | 93 | N/A |

[1]There are 17 cases and 16 unique patients--2 of the cases, namely #3 and #7, occurred in the same patient, after different doses.
[2]"Event" refers to the day of the patient's alerting 299.00/F84.0 code (after vaccination with HPV9).
[3]Age is recorded as a truncated integer. For example, a patient who is 15.01 years old and a patient who is 15.99 years old would both be listed as 15 years old.
[4]"Time-to-event in days" refers to the number of days between the HPV9 vaccination and the day of the patient's alerting 299.00/F84.0 code.
[5]"Prior related ICD code" is defined as an ICD diagnosis code in the category of "pervasive developmental disorders" that appears in the claims data prior to HPV9 vaccination.
[6]Only the prior related ICD code closest to HPV9 vaccination is recorded in this list. It must have occurred at least 6 months prior to the patient's alerting 299.00/F84.0 code, given the incidence criteria.
[7]The length of months is measured by the 1st of each month. For example, 01 Dec 2010 to 01 Jan 2011 and 01 Dec 2010 to 31 Jan 2011 are both calculated as 1 month. 01 Dec 2010 to 01 Feb 2011 is calculated as 2 months.