Practice of Epidemiology

Adapting Group Sequential Methods to Observational Postlicensure Vaccine Safety Surveillance: Results of a Pentavalent Combination DTaP-IPV-Hib Vaccine Safety Study


* Correspondence to Dr. Jennifer C. Nelson, Biostatistics Unit, Group Health Research Institute, 1730 Minor Avenue, Suite 1600, Seattle, WA 98101 (e-mail: nelson.jl@ghc.org).

Initially submitted September 7, 2011; accepted for publication March 29, 2012.

To address gaps in traditional postlicensure vaccine safety surveillance and to promote rapid signal identification, new prospective monitoring systems using large health-care database cohorts have been developed. We newly adapted clinical trial group sequential methods to this observational setting in an original safety study of a combination diphtheria and tetanus toxoids and acellular pertussis adsorbed (DTaP), inactivated poliovirus (IPV), and *Haemophilus influenzae* type b (Hib) conjugate vaccine (DTaP-IPV-Hib) among children within the Vaccine Safety Datalink population. For each prespecified outcome, we conducted 11 sequential Poisson-based likelihood ratio tests during September 2008–January 2011 to compare DTaP-IPV-Hib vaccinees with historical recipients of other DTaP-containing vaccines. No increased risk was detected among 149,337 DTaP-IPV-Hib vaccinees versus historical comparators for any outcome, including medically attended fever, seizure, meningitis/encephalitis/myelitis, nonanaphylactic serious allergic reaction, anaphylaxis, Guillain-Barré syndrome, or invasive Hib disease. In end-of-study prespecified subgroup analyses, risk of medically attended fever was elevated among 1- to 2-year-olds who received DTaP-IPV-Hib vaccine versus historical comparators (relative risk = 1.83, 95% confidence interval: 1.34, 2.50) but not among infants under 1 year old (relative risk = 0.83, 95% confidence interval: 0.73, 0.94). Findings were similar in analyses with concurrent comparators who received other DTaP-containing vaccines during the study period. Although lack of a controlled experiment presents numerous challenges, implementation of group sequential monitoring methods in observational safety surveillance studies is promising and warrants further investigation.

Abbreviations: CI, confidence interval; DTaP, diphtheria and tetanus toxoids and acellular pertussis adsorbed; Hib, *Haemophilus influenzae* type b; ICD-9-CM, *International Classification of Diseases, Ninth Revision, Clinical Modification*; IPV, inactivated poliovirus; MaxSPRT, maximized sequential probability ratio test; MCO, medical care organization; RR, relative risk; VSD, Vaccine Safety Datalink.

Gaps in medical product safety evidence have led to national calls for better and more rapid postlicensure vaccine surveillance (1, 2) and have spurred development of new systems that prospectively monitor large observational cohorts using computerized health-care databases. In its Vaccine Safety Datalink (VSD) project (3), the Centers for Disease Control and Prevention has created a population-based framework with which to conduct near real-time surveillance by weekly updating of vaccine and adverse event data on over 9 million members of 8 medical care organizations (MCOs) (4) and has used this framework to monitor the safety of many new vaccines since 2005 (5–10). One
such study found an increased risk of febrile seizure after receipt of a combination measles-mumps-rubella-varicella vaccine in comparison with separate injections of measles-mumps-rubella and varicella component vaccines among children aged 12–23 months (6); this and other evidence (11) led to national policy changes for use of measles-mumps-rubella-varicella combination vaccine (12). More broadly, the Food and Drug Administration is building a national electronic sentinel system to conduct surveillance for all of its regulated medical products (13).

One method that has been used in these new surveillance settings is sequential monitoring. It involves repeated analyses of accruing individuals who receive a new drug or vaccine while controlling the overall false-positive (type 1) error rate across the sequential tests. This approach is appealing because safety problems can be detected early, as soon as desired scientific and statistical criteria for an elevated adverse event risk are met. Several continuous sequential methods, which test after each new individual is exposed, have been proposed for use in observational safety surveillance, including cumulative sum charts (14), the original (15) and maximized (16) sequential probability ratio tests (MaxSPRT), and sequential generalized likelihood ratio tests (17). MaxSPRT, which has been shown to be a sequential generalized likelihood ratio test without a lower (i.e., futility) boundary (17), has been used routinely within the VSD (5–10). Continuous versus less frequent testing can reduce the expected time to signal detection, but highly frequent testing is not always practical, inherently reduces statistical power, and may yield false-positive findings at early study testing points when relatively few data have accumulated and test statistics are less stable (18).

Group sequential methods are designed for more periodic interim testing and are used extensively in clinical trials (19). Sequential designs customized to monitor efficacy outcomes in trials have been well researched, and key design questions (e.g., How frequently should tests be performed? What should the shape of the stopping boundary be over time? When should surveillance end?) have been addressed. Answers to these questions determine the design’s statistical performance characteristics (e.g., type 1 error, power, expected time to signal detection) and are thus generally guided by the study’s scientific goals and ethical concerns (19). However, in observational safety studies as compared with trials, the scientific question, the consequences of a signal, and the costs of false-positive and -negative errors are different, and much less attention has been paid to assessing what designs may be preferred given these differences. In fact, group sequential methods have not been widely considered at all for observational surveillance. Doing so could yield important improvements over existing continuous sequential designs and, in particular, increase power to detect serious rare adverse events.

In June 2008, a pentavalent combination diphtheria and tetanus toxoids and acellular pertussis adsorbed (DTaP), inactivated poliovirus (IPV), and Haemophilus influenzae type b (Hib) conjugate vaccine (trade name Pentacel; Sanofi Pasteur, Swiftwater, Pennsylvania) was approved for use in the United States among children aged 6 weeks–4 years (20, 21) as doses 1–4 of the DTaP, IPV, and Hib infant vaccine series, recommended for administration at 2, 4, 6, and 15–18 months of age, respectively (22). Use of combination versus individual vaccines is associated with increased rates of vaccine series completion (23) and may be preferred to reduce the number of injections and administrative costs (23–25).

The purposes of this work were 1) to adapt clinical trial group sequential methods to an observational vaccine safety setting and 2) to present results of an original DTaP-IPV-Hib vaccine safety study among children in the VSD population that represented a novel implementation of these methods. Thus, the novelties of this work are also 2-fold and include both the study methodology and the vaccine safety study results. The sequential design considerations and analysis approach are detailed, including how they address the many challenges that arise in observational healthcare databases as compared with a clinical trial setting. These challenges primarily derive from the lack of a controlled experiment and, most notably, include confounding.

MATERIALS AND METHODS

Study design

A prospective group sequential observational safety study of DTaP-IPV-Hib vaccine was conducted from September 2008, when vaccine uptake began in the VSD, to January 2011. The study population included children aged 6 weeks–2 years (i.e., ≤36 months) who received DTaP-IPV-Hib vaccine or another DTaP-containing comparator vaccine during the study period and were enrolled at one of 7 VSD MCOs: Group Health Cooperative (Seattle, Washington); Harvard Vanguard Medical Associates (Boston, Massachusetts); Harvard Pilgrim Health Care (Boston, Massachusetts); Kaiser Permanente of Colorado (Denver, Colorado); Kaiser Permanente of Northern California (Oakland, California); Kaiser Permanente of Southern California (Pasadena, California); Marshfield Clinic Research Foundation (Marshfield, Wisconsin); and Kaiser Permanente Northwest (Portland, Oregon). Databases that were updated weekly captured information on demographic factors, immunizations, and International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnoses assigned to inpatient, emergency department, and outpatient visits and were used to identify and follow vaccine recipients for the occurrence of specific ICD-9-CM coded adverse events. Each MCO’s institutional review board approved this study; informed consent was not required.

Adverse events and exposure groups

On the basis of prelicensure safety data (20, 26–30), postmarket spontaneous reports, expert input, and biologic plausibility, 7 adverse events were monitored in predefined postvaccination at-risk windows: medically attended fever, seizure, meningitis/encephalitis/myelitis, serious nonanaphylactic allergic reaction, anaphylaxis, Guillain-Barré syndrome, and invasive Hib disease. Invasive Hib disease was of concern because some prelicensure data suggested that
immunogenicity for the Hib component was lower among recipients of DTaP-IPV-Hib than among recipients of separately administered vaccine components (20). Concern was heightened because of the Hib vaccine shortage (31) and the occurrence of 5 cases of invasive Hib disease in Minnesota (32). Vaccinees were also followed for the occurrence of hospitalization due to any cause to capture serious events not suspected in advance.

Table 1 displays detailed information on each outcome definition, including the ICD-9-CM code(s), medical setting (outpatient, inpatient, emergency department), and at-risk window. Only inpatient and emergency department events were included for many outcomes in order to restrict the data to the most severe outcomes and reduce misclassification. Further, to remove events with potential onset prior to vaccination, only the first (i.e., incident) event recorded in a prespecified time period was included, and events occurring on the day of vaccination (day 0) were excluded for several outcomes.

The exposed group of interest was recipients of DTaP-IPV-Hib vaccine during the study period. The primary comparison group was children who had received other DTaP-containing vaccines prior to the approval of DTaP-IPV-Hib vaccine, either as an individual injection or as part of a combination vaccine other than DTaP-IPV-Hib. In both groups, children may have received other non-DTaP vaccines on the same day. To minimize bias due to secular trends in outcome coding, comparators from a narrow 2-year historical period (April 2007–March 2009) were used when possible. For very rare outcomes, a longer historical period (January 2005–March 2009) was needed to obtain reliable historical estimates. Secondary concurrent comparator analyses contrasted DTaP-IPV-Hib vaccinees with children who received other DTaP-containing vaccines during the study period.

### Table 1. Definitions of Adverse Events in a Study of DTaP-IPV-Hib Vaccine Safety Among Members of the Vaccine Safety Datalink Population, September 2008–April 2011

<table>
<thead>
<tr>
<th>Adverse Event Outcome Group</th>
<th>ICD-9-CM Code(s)</th>
<th>Medical Setting</th>
<th>Incident Event Inclusion Criteria: First in What Period and Medical Setting?</th>
<th>Postvaccination At-Risk Interval, days</th>
<th>Events Excluded if These Codes Occurred on Same Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medically attended fever</td>
<td>780.6</td>
<td>Outpatient, inpatient, ED</td>
<td>In past 5 days in outpatient, inpatient, and ED settings</td>
<td>1–5</td>
<td>None</td>
</tr>
<tr>
<td>Seizure</td>
<td>345, 780.3</td>
<td>Inpatient, ED</td>
<td>In past 180 days in outpatient, inpatient, and ED settings</td>
<td>0–7</td>
<td>None</td>
</tr>
<tr>
<td>Meningitis, encephalitis, and myelitis</td>
<td>047.8, 047.9, 049.9, 321.2, 322, 323, 348.3, 348.5</td>
<td>Inpatient, ED</td>
<td>In past 180 days in outpatient, inpatient, and ED settings</td>
<td>1–30</td>
<td>047.0–047.1, 048, 049.0–049.8, 053–056, 320</td>
</tr>
<tr>
<td>Nonanaphylactic serious allergic reaction</td>
<td>995.1, 995.2, 708.0, 708.1, 708.9</td>
<td>Inpatient, ED</td>
<td>In past 30 days in outpatient, inpatient, and ED settings</td>
<td>1–2</td>
<td>None</td>
</tr>
<tr>
<td>Anaphylaxis</td>
<td>995.0, 999.4</td>
<td>Inpatient, ED</td>
<td>In past 30 days in outpatient, inpatient, and ED settings</td>
<td>0–2</td>
<td>None</td>
</tr>
<tr>
<td>Guillain-Barré syndrome</td>
<td>357.0</td>
<td>Outpatient, inpatient, ED</td>
<td>None (both incident and recurrent events are included)</td>
<td>1–30</td>
<td>None</td>
</tr>
<tr>
<td>Invasive Hib disease</td>
<td>038.41, 320.0</td>
<td>Outpatient, inpatient, ED</td>
<td>None (both incident and recurrent events are included)</td>
<td>Any day postvaccination after the third dose</td>
<td>None</td>
</tr>
<tr>
<td>Any hospitalization</td>
<td>All codes</td>
<td>Inpatient</td>
<td>In past 30 days in outpatient, inpatient, and ED settings</td>
<td>0–7</td>
<td>None</td>
</tr>
</tbody>
</table>

Abbreviations: DTaP, diphtheria and tetanus toxoids and acellular pertussis adsorbed; ED, emergency department; Hib, Haemophilus influenzae type b; ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; IPV, inactivated poliovirus.

a Three-digit codes (e.g., 345) include those that start with 3 digits and contain any additional fourth or fifth digit (e.g., 345.11); 4-digit codes include those that start with 4 digits and contain any fifth digit.

As doses of DTaP-IPV-Hib were newly administered among study participants, we performed a sequential test of the 1-sided alternative hypothesis that the risk of each targeted adverse event was elevated for DTaP-IPV-Hib vaccinees versus historical comparators (H₀: relative risk (RR) = 1; H₁: RR > 1). In all, 12 sequential Poisson-based likelihood ratio tests that compared the observed number of events among accrued DTaP-IPV-Hib recipients with the expected number based on historical data were planned for each prespecified outcome. To control for confounding, historical event rates were computed by MCO site, gender, and age group and used to compute stratum-specific

expected counts based on the observed confounder distribution among DTaP-IPV-Hib recipients (see Web Appendix 1, available at http://aje.oxfordjournals.org/). Expected counts were then summed across strata and compared with the total number of observed counts. The first test was performed 1 year after DTaP-IPV-Hib uptake began; 33,308 doses had been received by this time. Thereafter, 11 tests were planned to be evenly spaced based on the number of newly accumulated doses. A Pocock-style stopping boundary (33) was computed via simulation (34, 35) (Web Appendix 2) and applied so that the null hypothesis was rejected at test \( t \) if the log likelihood ratio exceeded a preset
threshold designed to maintain the overall type 1 error at $\alpha = 0.05$. $P$ values that took into account the sequential testing were calculated using standard methods from clinical trials (36).

Given this planned testing schedule, boundary, and historical event rates, the between-test sample size (i.e., for tests after the first) required to achieve 80% power to detect a specific relative risk for each event was found via simulation (Web Appendix 3). To achieve at least 80% power to detect a relative risk of 2 for the more common events of medically attended fever and seizure, we planned tests 2–12 after each new 3,500 doses of DTaP-IPV-Hib accrued, yielding a maximum total planned sample size of 71,808. To ensure 80% power to detect relative risks of 2 and 2.5 for rarer events like meningitis/encephalitis/myelitis and nonanaphylactic allergic reaction, respectively, larger between-test ($N_{\text{bet}} = 10,500$) and total ($N = 148,808$) sample sizes were needed. For the rarest outcomes (Guillain-Barré syndrome, anaphylaxis, invasive Hib disease), occurring at a rate of less than 0.05 cases per 10,000 doses in historical data, event counts were tracked but were not formally tested because of a lack of power (<30%). The nonspecific outcome of any-cause hospitalization was analyzed as an end-of-study endpoint. For comparison, we also conducted analyses using Poisson-based MaxSPRT (16), a continuous sequential method that has been used in prior studies (5–10).

**Secondary end-of-study analyses**

Secondary analyses were performed using all accrued DTaP-IPV-Hib doses at the end of the study. To assess the robustness of the sequential results across subgroups, we examined relative risks comparing DTaP-IPV-Hib recipients with historical comparators by MCO site and by age group within site. In addition, we also performed sensitivity analyses comparing event risks between DTaP-IPV-Hib recipients and concurrent comparators by estimating adjusted odds ratios with logistic regression models that controlled for age group, gender, and MCO site. Odds ratios were also examined by MCO site and by age group within site. Concurrent comparisons were not made sequentially because of the expected replacement of separate injections of DTaP, Hib, and IPV by DTaP-IPV-Hib and an anticipated lack of overlap between those groups (and thus an inability to make comparisons concurrently) at most sites at any given sequential test.

**RESULTS**

**Vaccine uptake**

During the study, 149,337 doses of DTaP-IPV-Hib vaccine were administered. Uptake was initially slower (a mean of 857 doses per week (standard deviation, 418) in the first half of the study) and increased over time (a mean of 1,513 doses per week (standard deviation, 246) in the second half). Uptakes of DTaP-IPV-Hib vaccine were comparable over time by gender but differed by age group (Figure 1A), with receipt almost exclusively by infants under 1 year of age during the first year, representing the first 3 doses of the series. Consistent with recommendations, most older children did not receive the final booster dose until after the first year of the study. Across MCO sites, DTaP-IPV-Hib uptake ranged from rapid (sites 1–3) to gradual (site 4) to delayed (site 5) to virtually nonexistent (sites 6 and 7) (Figure 1B).

**Primary sequential analyses**

No increased risk was detected for any of the 4 sequentially monitored events (medically attended fever, seizure, meningitis/encephalitis/myelitis, or nonanaphylactic allergic reaction) among DTaP-IPV-Hib vaccinees versus historical comparators, as evidenced by the log likelihood ratio remaining below the stopping boundary at each test point (Figures 2 and 3). Among 72,651 DTaP-IPV-Hib recipients, 348 fevers were observed versus 317.7 expected based on historical data ($RR = 1.10$, $P = 0.06$), and 9 seizures occurred versus 8.6 expected ($RR = 1.04$, $P = 0.50$). Among 149,337 DTaP-IPV-Hib vaccinees, only 5 cases each of nonanaphylactic allergic reaction and meningitis/encephalitis/myelitis were observed versus 6.1 and 8.8 expected, respectively, suggesting no elevated risk (Table 2). Comparative analyses using MaxSPRT similarly did not detect any increased risks, which was not surprising since continuous testing necessitates a higher stopping boundary in order to control the same type 1 error and is thus less powerful. For example, for fever, the log likelihood ratio stopping boundary for MaxSPRT was 4.15 as compared with 2 for the primary group sequential design (Figure 2).

**Secondary end-of-study analyses**

*Subgroup analyses using historical comparators.* Subgroup analyses by age group and MCO site were conducted only for medically attended fever, as there were too few cases of other adverse events to do so. Among all 149,337 DTaP-IPV-Hib recipients at the end of the study, a modest decrease in risk of medically attended fever was suggested for DTaP-IPV-Hib vaccinees under 1 year of age versus historical comparators ($RR = 0.83$, 95% confidence interval (CI): 0.73, 0.94). Fever risk was elevated, however, among older children aged 1–2 years who received DTaP-IPV-Hib vaccine versus comparators ($RR = 1.83$, 95% CI: 1.34, 2.50), a trend that was observed at most sites (site-specific range of RRs: 1.02–2.12).

*Concurrent comparator analyses.* During the study period, 813,325 children received DTaP vaccine, either individually or as part of a combination vaccine other than DTaP-IPV-Hib, and thus represented concurrent comparators. Of these, 66.0% of DTaP vaccines were given concomitantly with Hib and IPV vaccines, 13.7% were given with either Hib or IPV, and 20.3% were not given with either. Vaccine use was similar by gender, but a greater proportion of DTaP-IPV-Hib vaccinees versus concurrent comparators were under 1 year of age (87.1% vs. 71.7%). After adjusting for age group, gender, and site, no evidence of an increased risk of medically attended fever, seizure, meningitis/encephalitis/myelitis, nonanaphylactic allergic reaction, or all-cause hospitalization was observed.
among DTaP-IPV-Hib vaccinees versus concurrent DTaP comparators (Table 3). In addition, no cases of Guillain-Barré syndrome, anaphylaxis, or invasive Hib disease occurred among DTaP-IPV-Hib recipients during the study period.

Among children aged 1–2 years, DTaP-IPV-Hib vaccinees had a higher fever risk than did concurrent comparators (odds ratio = 1.75, 95% CI: 1.38, 2.22), while DTaP-IPV-Hib recipients versus comparator recipients under 1 year of age

Figure 2. Risk of medically attended fever among DTaP-IPV-Hib vaccine recipients in the Vaccine Safety Datalink population as compared with historical recipients of comparator vaccines, September 2008–January 2011. A) Trajectory of the observed log likelihood ratio (LLR) test statistic during the study period and the sequential stopping boundaries for each of the 11 sequential tests performed. B) Trajectory of the corresponding observed relative risk (RR) and the sequential stopping boundaries on the scale of the RR for each of the 11 sequential tests performed. DTaP, diphtheria and tetanus toxoids and acellular pertussis adsorbed; Hib, Haemophilus influenzae type b; IPV, inactivated poliovirus.
had a slightly lower risk (odds ratio = 0.83, 95% CI: 0.71, 0.96), a result that was observed across several sites.

Among the 545 DTaP-IPV-Hib and 2,045 concurrent comparator vaccinees who experienced a medically attended fever, a smaller percentage of DTaP-IPV-Hib recipients (20.6%) versus comparators (25.8%) were ≥1 year of age. Most fevers were observed on the first day after receipt of DTaP-IPV-Hib (40.6%) or comparator vaccines.

Figure 3. Risk of seizure among DTaP-IPV-Hib vaccine recipients in the Vaccine Safety Datalink population as compared with historical recipients of comparator vaccines, September 2008–January 2011. A) Trajectory of the observed log likelihood ratio (LLR) test statistic during the study period and the sequential stopping boundaries for each of the 11 sequential tests performed. B) Trajectory of the corresponding observed relative risk (RR) and the sequential stopping boundaries on the scale of the RR for each of the 11 sequential tests performed. DTaP, diphtheria and tetanus toxoids and acellular pertussis adsorbed; Hib, Haemophilus influenzae type b; IPV, inactivated poliovirus.
DISCUSSION

We designed and implemented a group sequential approach to prospectively monitor the safety of a newly introduced DTaP-IPV-Hib vaccine during its uptake in a large cohort of infants within the VSD population. Prior studies have used near-continuous monitoring instead of interim monitoring (5–10) and, to our knowledge, this is one of the first attempts to adapt clinical trial group sequential methods to an observational setting where rare adverse events are of primary interest. No differences in risk were detected among DTaP-IPV-Hib vaccinees versus historical comparators who received other DTaP-containing vaccines prior to the licensure of DTaP-IPV-Hib for any sequentially monitored event (medically attended fever, seizure, meningitis/encephalitis/myelitis, and serious nonanaphylactic allergic reaction), a finding that was replicated in end-of-study sensitivity analyses using concurrent comparators who received other DTaP-containing vaccines during the study.

Table 2. Numbers and Rates (Per 10,000 Doses) of Adverse Events Among DTaP-IPV-Hib Vaccine Recipients From the Vaccine Safety Datalink Study Population Versus Rates Estimated From Age-, Gender-, and Site-Comparable Historical Comparators (Primary Sequential Analyses)\textsuperscript{a}

<table>
<thead>
<tr>
<th>Adverse Event Outcome Group</th>
<th>DTaP-IPV-Hib Recipients</th>
<th>Historical DTaP Comparator Group Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed No. of Events</td>
<td>Expected No. of Events</td>
</tr>
<tr>
<td></td>
<td>Observed No. of DTaP-IPV-Hib Doses</td>
<td>Rate Per 10,000 Doses</td>
</tr>
<tr>
<td>Medically attended fever</td>
<td>348</td>
<td>317.7</td>
</tr>
<tr>
<td>Seizure</td>
<td>9</td>
<td>8.6</td>
</tr>
<tr>
<td>Meningitis, encephalitis, myelitis</td>
<td>5</td>
<td>8.8</td>
</tr>
<tr>
<td>Nonanaphylactic serious allergic reaction</td>
<td>5</td>
<td>6.1</td>
</tr>
</tbody>
</table>

Abbreviations: DTaP, diphtheria and tetanus toxoids and acellular pertussis adsorbed; Hib, Haemophilus influenzae type b; IPV, inactivated poliovirus.

\textsuperscript{a} Sequential surveillance was conducted during September 2008–March 2010 for medically attended fever and seizure and during September 2008–January 2011 for meningitis/encephalitis/myelitis and nonanaphylactic serious allergic reaction.

Table 3. Numbers and Rates (Per 10,000 Doses) of Adverse Events Among DTaP-IPV-Hib Vaccinees and Concurrent Comparators Who Received DTaP as an Individual Injection in the Vaccine Safety Datalink Study Population (Secondary End-of-Study Concurrent Comparator Analyses), September 2008–January 2011\textsuperscript{a}

<table>
<thead>
<tr>
<th>Adverse Event Outcome Group</th>
<th>DTaP-IPV-Hib Recipients</th>
<th>Concurrent DTaP Recipients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of Events</td>
<td>No. of DTaP-IPV-Hib Doses</td>
</tr>
<tr>
<td>Medically attended fever</td>
<td>545</td>
<td>149,337</td>
</tr>
<tr>
<td>Fever, by age group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 weeks–11 months</td>
<td>433</td>
<td>130,073</td>
</tr>
<tr>
<td>12–35 months</td>
<td>112</td>
<td>19,264</td>
</tr>
<tr>
<td>Seizure</td>
<td>17</td>
<td>149,337</td>
</tr>
<tr>
<td>Meningitis, encephalitis, myelitis</td>
<td>5</td>
<td>149,337</td>
</tr>
<tr>
<td>Nonanaphylactic serious allergic reaction</td>
<td>5</td>
<td>149,337</td>
</tr>
<tr>
<td>Any hospitalization</td>
<td>192</td>
<td>149,337</td>
</tr>
</tbody>
</table>

Abbreviations: DTaP, diphtheria and tetanus toxoids and acellular pertussis adsorbed; Hib, Haemophilus influenzae type b; IPV, inactivated poliovirus; MCO, medical care organization.

\textsuperscript{a} End-of-study subgroup analyses are also shown for medically attended fever.

\textsuperscript{b} Adjusted for gender, age group, and MCO site and to reflect the average covariate distribution in the study sample.

\textsuperscript{c} Adjusted for MCO site only (due to the small number of events) and to reflect the average covariate distribution in the study sample.

period. No cases of Guillain-Barré syndrome, anaphylaxis, or invasive Hib disease occurred among DTaP-IPV-Hib vaccinees, and no elevated risk of hospitalization due to any cause was found, suggesting that DTaP-IPV-Hib vaccine is not associated with nonspecific serious adverse events not suspected in advance. These findings are consistent with prelicensure data, which also did not suggest increased risks (20, 26–30).

Although no overall increase in the risk of medically attended fever was detected for DTaP-IPV-Hib vaccinees versus historical comparators, the estimated relative risk of 1.10 nearly achieved statistical significance in sequential analyses. In secondary end-of-study prespecified subgroup analyses, this marginal overall increase was observed to be driven by 1- to 2-year-olds who received DTaP-IPV-Hib vaccine versus historical comparators (RR = 1.83), while the opposite trend was seen among infants under 1 year of age (RR = 0.83). Similar results were observed when comparing DTaP-IPV-Hib vaccinees with concurrent comparators. These differing age group trends were observed at some but not all MCO sites, and they are consistent with some prelicensure data (20). Although suggestive, the age group findings in the current study represent one of many prespecified subgroup analyses and are thus preliminary in nature. Additional study is needed to assess whether they are independently replicable before drawing stronger conclusions.

A major strength of this study was its novel use of group sequential monitoring to conduct safety surveillance for rare adverse events using observational health-care data. Prior studies (5–10) have utilized near-continuous sequential monitoring, which is inherently less powerful and can exhibit false-positive signals at early test points because of small-sample variability (18). A detailed review (37) and discussions of the trade-offs between differing sequential designs in the context of drug and vaccine safety surveillance (17, 38, 39) have been published elsewhere. Key features of the current design included 1) a delayed first sequential test followed by more frequent interim likelihood ratio testing, 2) a Pocock-style (approximately flat on the log likelihood ratio scale) signaling threshold, and 3) spacing between tests and length of surveillance based on accrued sample sizes designed to achieve specific power goals. The choice of a likelihood ratio test and (nearly) flat Pocock-style threshold were similar to previous sequential vaccine safety studies (5–10) and pilot drug safety studies (40). A Pocock-style threshold, which is lower at earlier testing points than, for example, the O’Brien-Fleming boundary that is commonly used in clinical trials (41), was preferred since the risk tolerance for elevated adverse events among healthy children receiving vaccines is very low. The testing frequency and surveillance duration plans differed from prior work, which has generally employed near-continuous (e.g., weekly) testing, sometimes for a preset study period based on calendar time (e.g., 2 years) (5).

The primary rationale for delaying the first test and reducing the planned total number of tests in the current study (to 12 tests in 2 or more years) versus previous studies (approximately 104 tests in 2 years) was to increase study power for rare events that were not feasible to detect in smaller prelicensure studies. Simulation studies that we have conducted have shown that power gains close to 20% can be achieved with this less frequent testing design versus near-continuous testing (42). Delaying the first test also prevents false-positive signals early in the study when sample sizes are smaller and test statistics are more variable and susceptible to influence by just a few events. Last, since the rate of new vaccine uptake in an observational setting over time is generally not predictable, we used accrued sample size instead of calendar time to plan the duration of surveillance in order to ensure that power requirements were met. This led to a more complicated design in which rarer events were monitored less frequently and for a longer period of time than more common events, but this was deemed important from a scientific perspective in order to allow more information to accumulate for rarer outcomes so that adequate power to address study questions could be achieved.

Applying clinical trial group sequential methods in an observational database safety setting raises several additional challenges, including confounding and unpredictable changes in the data over time. For example, in the current study, there was highly differential DTaP-IPV-Hib vaccine uptake by age group. In addition, an unexpectedly large number of new DTaP-IPV-Hib vaccinees was accrued from a single site after the fifth test because of a data quality issue that was identified and resolved, causing one planned sequential test to be skipped. Collectively, these factors affect outcome variability over time and thus the probability of committing a type 1 error, which investigators want to control. Planned sequential thresholds should ideally be adjusted at each test to account for these issues, and existing group sequential methods readily offer the flexibility for making such adaptations. A more in-depth discussion of these and other challenges can be found elsewhere (38).

This study shares the standard limitations of database studies that rely on electronic medical data collected during the course of routine health care and not for research purposes. The drawbacks of such data have been described (5–10) and include misclassification of adverse event and vaccination status and incomplete data capture due to delays in receipt of some insurance claims by MCO systems (43). An additional limitation of the current study was the primary use of historical comparators, a relatively weaker comparison group. Although a concurrent comparison group would have been preferred, as it would have inherently controlled for temporal bias associated with changes in outcome coding practices over time, DTaP-IPV-Hib vaccine was expected to quickly replace separate administration of component vaccines at most MCO sites because of the Hib vaccine shortage (31). Thus, we did not anticipate having sufficient overlap at most sites between DTaP-IPV-Hib recipients and children who received separate injections to make sequential comparisons throughout the study period. In addition, it would have been preferable to use a more narrowly defined comparator group of children who received separate injections of all 3 components of DTaP-IPV-Hib (i.e., DTaP, IPV, and Hib vaccines) on
the same day rather than those who received DTaP on the same day regardless of their IPV and Hib vaccine status. However, this group was too small in some age groups for sequential comparisons. Nevertheless, we did take advantage of this more specific comparator group in secondary end-of-study analyses and found results to be comparable to analyses with the broader comparator group.

Improving postlicensure drug and vaccine safety surveillance is a recognized national and international priority (1, 2, 13, 44). Sequential monitoring of health-care databases to address postlicensure safety questions offers great potential, and the use of group sequential methodology appears particularly promising. In the current study, novel implementation of a group sequential approach provided reassurance about the safety of a new DTaP-IPV-Hib vaccine. Further method development and evaluation is needed, however, to more fully assess the performance and determine the optimal implementation of group sequential methods in observational safety surveillance settings.

ACKNOWLEDGMENTS

Author affiliations: Group Health Research Institute, Seattle, Washington (Jennifer C. Nelson, Andrea J. Cook, Clara Dominguez, Lisa A. Jackson, Do Peterson, Onchee Yu); Department of Biostatistics, School of Public Health, University of Washington, Seattle, Washington (Jennifer C. Nelson, Andrea J. Cook, Clara Dominguez); Department of Population Medicine, Harvard Medical School, Boston, Massachusetts (Sharon K. Greene, Katherine Yih); Harvard Pilgrim Health Care Institute, Boston, Massachusetts (Sharon K. Greene, Katherine Yih); Kaiser Permanente Vaccine Study Center, Oakland, California (Nicola P. Klein); Immunization Safety Office, Division of Health Care Quality Promotion, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia (Karen R. Broder, Eric Weintraub); Institute for Health Research, Kaiser Permanente Colorado, Aurora, Colorado (Matthew F. Daley); Department of Pediatrics, School of Medicine, University of Colorado, Aurora, Colorado (Matthew F. Daley); and Department of Research and Evaluation, Kaiser Permanente of Southern California, Pasadena, California (Steven J. Jacobsen).

This work was supported by a subcontract with America’s Health Insurance Plans under contract 200-2002-00732 from the Centers for Disease Control and Prevention.

For their efforts related to study conception and oversight, data acquisition, and technical support, we thank several Vaccine Safety Datalink investigators (Drs. Ed Belongia, Martin Kulldorf, Tracy Liu, David L. McClure, and Allison Naleway) and staff members (Patti Benson, Nick Berger, Lora Bounds, Stephanie Irving, Ruihuа Yin).

Preliminary results of this work were presented orally at meetings of the Western North American Region of the International Biometric Society (Seattle, Washington, June 21, 2010) and the International Biometric Society (Florianopolis, Santa Catarina, Brazil, December 10, 2010).

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Dr. Nelson has served as a statistical consultant to Glaxo-Smith-Kline. Dr. Klein has received research support from Sanofi Pasteur, Glaxo-Smith-Kline, Pfizer, Novartis, and Merck. Dr. Jackson has received research funding from Sanofi Pasteur, Pfizer, and Novartis and has received travel support for meeting attendance from Pfizer.

REFERENCES


29. Gold R, Barreto L, Ferro S, et al. Safety and immunogenicity of a fully liquid vaccine containing five-component pertussis-