Establishing Potency Specifications for Antigen Vaccines

Clinical Validation of Statistically Derived Release and Stability Specifications

Robert Capen, Mary L. Shank-Retzlaff, Heather L. Sings, Mark Esser, Carlos Sattler, Michael Washabaugh, and Robert Sitrin

Gardasil (a registered trademark of Merck and Co., Inc., www.merck.com) is the first vaccine approved for women aged 9–26 years old in prevention of cervical cancer and genital warts as well as vulvar and vaginal precancerous legions. The vaccine contains noninfectious virus-like particles (VLPs) corresponding to HPV types 6, 11, 16, and 18. It is produced by recombinantly expressing the major HPV capsid protein, L1, for each type in yeast (1, 2). The L1 monomers self-assemble to produce icosahedral VLPs that are structurally similar to the native virion in size, assembly, and immunological properties (3–5). The VLPs are purified chromatographically, diluted, and then adsorbed onto Merck’s aluminum adjuvant to produce monovalent bulks (1, 2). A single dose contains 20, 40, 40, and 20 µg of VLP types 6, 11, 16, and 18, respectively, formulated on 225 µg of the aluminum adjuvant and administered in a final volume of 0.5 mL. Each bulk lot is formulated based on protein concentration. Four monovalent bulks, one per HPV type, are subsequently diluted and blended to produce the final container material. The dose is fixed for each type and based on mass; thus, there is minimal lot-to-lot variation in protein concentration.

Because the vaccine’s potency depends on the specific activities (specific antigenicities) of each type-specific VLP source bulk, some variation in potency among final containers is anticipated.

An enzyme-linked immunosorbent assay, referred to as the in vitro relative potency assay (IVRP), is used as the potency test for the Gardasil vaccine (6). It measures the amount of antibodies bound to neutralizing epitopes for each HPV type (7–10). Results are reported relative to a Gardasil lot that was used in a phase 3 clinical trial. The assay, therefore, provides a direct comparison of the antigen content of each VLP type in a given test sample and the corresponding antigen content of a lot that has been shown to be efficacious in humans. IVRP results correlate with immunogenicity results obtained using a traditional mouse potency test and are considered predictive of immunogenicity in humans (6).

At present, there is no immune correlate of protection for HPV. In the phase 2/3 clinical trial program, prophylactic vaccination was highly efficacious in preventing infection and disease caused by HPV types 6, 11, 16, and 18 for at least five years (11). No breakthrough cases due to waning immunity have been described (12–14). Because there is no immune correlate, and only a limited number of final container lots were manufactured before licensure, a novel approach for establishing potency specifications was developed and applied to the Gardasil vaccine. Preliminary specifications were derived using a propagation-of-error calculation starting from the IVRP values of bulk production lots.
validated using a clinical potency-ranging study in which formulations simulating low-potency samples, or those with reduced IVRPs, were evaluated and the results compared with the proposed stability specification. Here we present the statistical model and results of the potency-ranging study.

**Materials and Methods**

**Study Vaccine:** Samples of the vaccine used in this study consisted of bulks and final container lots manufactured as described previously (1, 2). Briefly, type-specific HPV L1 proteins were independently expressed in Saccharomyces cerevisiae. The cells were harvested and lysed, and the L1 proteins and VLPs were purified chromatographically. For Types 6, 11, and 16, the purified monovalent VLPs were treated with dithiotheritol to disassemble the particles, which were then allowed to reassemble using a process similar to that described previously (15, 16). The purified VLPs were diluted and adsorbed onto the aluminum adjuvant at a concentration of 320 µg/mL protein to produce the monovalent bulks. To make the final product, four monovalent bulks, one per HPV type, were diluted and blended to give a final product containing HPV types 6, 11, 16, and 18 VLPs at concentrations of 40, 80, 80, and 40 µg/mL protein, respectively (100% dose formulation).

As we describe below under “Results and Discussion,” Gardasil samples are inherently stable. It was not feasible to generate samples formulated at the 100% dose level that would exhibit sufficiently reduced IVRP values (17). Because IVRP values are proportional to protein concentration, low-potency samples were simulated by preparing final container lots formulated at reduced protein concentrations. As Table 1 shows, lots were formulated at 20%, 40%, and 60% of the standard formulation by diluting the 100% dose formulation used in this study with varying amounts of aluminum adjuvant such that all final vaccine formulations (full-dose and partial-dose) contained 225 µg of aluminum adjuvant per 0.5 mL dose.

**IVRP Testing:** Both monovalent bulks and final container samples were tested for IVRP using a sandwich-type enzyme immunoassay that has been described in detail elsewhere (6). A 96-well microplate was coated with the capture antibody, one of H6.M48, K11.B2, H16.J4, or H18.J4, depending on the HPV type being tested. Unless otherwise noted, all antibodies were obtained from Dr. Neil Christensen (Penn State University) or produced in-house.

The plates were allowed to incubate overnight and were subsequently washed. The plates were then blocked with bovine serum albumin. These blocked, antibody-coated plates are referred to as assay plates. The samples and reference standard were diluted in assay diluent to a target starting concentration of 2 µg/mL.

From the initial dilution, 10 three-fold dilutions were prepared and transferred to the assay plate. To dissolve the aluminum-containing adjuvant, a citrate-phosphate dissolution buffer was added to the assay plate. The plates were allowed to incubate overnight at room temperature. They were again washed and the detection antibody added. The detection antibody was one of H6.B10.5, H11.B2, H16.V5 or H18. R5, depending on the HPV type being tested. The amount of detection antibody that bound to the plate was quantified using a goat anti-IgG 2b-horseradish peroxidase conjugate (Southern Biotechnology, www. southernbiotech.com) and tetramethylbenzidine (Sigma, www.sigmaaldrich.com), a colorimetric substrate.

The resulting optical densities were plotted relative to the dilution factor. The data were analyzed using a four-parameter logistic function. The final IVRP was calculated using the following equation:

$$\text{IVRP} = \frac{\text{ED}_{50,\text{sample}} \times \text{assigned reference standard potency}}{\text{ED}_{50,\text{standard}}}$$

where ED_{50,\text{sample}} and ED_{50,\text{standard}} are the ED_{50}s for the test sample and

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Expected Value for Type 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM_{MB}</td>
<td>Geometric mean IVRP for monovalent bulks used to formulate final container lots of Gardasil</td>
<td>311 Units/mL</td>
</tr>
<tr>
<td>f_{dilution}</td>
<td>Dilution factor used during formulation</td>
<td>0.25</td>
</tr>
<tr>
<td>f_{stability}</td>
<td>Mean stability loss (slope)</td>
<td>1.0 (0% loss)</td>
</tr>
<tr>
<td>GM_{RC}</td>
<td>Expected geometric mean IVRP for final container lots of Gardasil</td>
<td>78 Units/mL</td>
</tr>
<tr>
<td>RSD_{RC}</td>
<td>Variation in IVRP for final container lots of Gardasil at release</td>
<td>13.4%</td>
</tr>
<tr>
<td>RSD_{FC}</td>
<td>Variation in IVRP for final container lots of Gardasil at expiry</td>
<td>9.3%</td>
</tr>
<tr>
<td>IVRP_{MB}</td>
<td>Minimum monovalent bulk release limit</td>
<td>232 Units/mL</td>
</tr>
<tr>
<td>Final</td>
<td>Minimum final container release limit</td>
<td>≥53 Units/mL</td>
</tr>
<tr>
<td>Container_{min}</td>
<td>Minimum final container stability limit</td>
<td>≥44 Units/mL</td>
</tr>
</tbody>
</table>
standard, respectively (the theoretical dilution that produces a response halfway between the minimum and maximum responses).

The potency of the reference standard was assigned before testing and is equal to 40 units/mL for types 6 and 18 and 80 units/mL for types 11 and 16. Those values are based on the nominal protein concentration of the reference standard lot. All samples were tested in triplicate. Because IVRP values are proportional to protein concentration, final container lots with specific activities equal to the reference standard lot will exhibit IVRPs of 40 units/mL for type 6 and 18 and 80 units/mL for types 11 and 16. Monovalent bulk lots with specific activities equal to the reference standard lot will exhibit IVRPs of 320 units/mL.

**Derivation of Bulk Release Specification:** The statistically derived lower release specification for the bulks was established by calculating a three-sigma limit based on process capability limits were calculated in the log scale and then converted to the linear scale as follows:

\[ \ln(\text{Monovalent Bulk}_{\text{spec}}) = \bar{y} - 3s_{\text{total}} \]

\[ \text{Monovalent Bulk}_{\text{spec}} = \text{GM}_{\text{MB}} \times \left(1 + \frac{RSD_{\text{total}}}{100}\right) \]

where \( \bar{y} \) is the mean of the natural-logarithm-transformed IVRP values for the monovalent bulks, \( \text{GM}_{\text{MB}} \) is the corresponding geometric mean found by exponentiating \( \bar{y} \), and \( s_{\text{total}} \) and \( RSD_{\text{total}} \) are defined above.

**Derivation of Final Container Release Specification:** The final-container minimum release specification (Final Container min) was derived using a propagation-of-error calculation. This model mathematically represents the expectation that the IVRP (for a given HPV type) of a final container lot is affected by the monovalent bulk lot used for formulation, the stability of the bulk lot, other manufacturing aspects related to formulation and filling, and assay variability. It further

---

**Table 3:** Summary of subjects excluded from the per-protocol immunogenicity populations by vaccination group

<table>
<thead>
<tr>
<th>Quadrivalent HPV (Types 6,11,16,18) L1 VLP Vaccine</th>
<th>20% Formulation (N = 504)</th>
<th>40% Formulation (N = 514)</th>
<th>60% Formulation (N = 508)</th>
<th>100% Formulation (N = 1,019)</th>
<th>Total (N = 2,545)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects receiving ≥1 injection</td>
<td>503</td>
<td>513</td>
<td>508</td>
<td>1,017</td>
<td>2,541</td>
</tr>
<tr>
<td>Subjects excluded from PP population</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV 6/11</td>
<td>128</td>
<td>120</td>
<td>138</td>
<td>271</td>
<td>657</td>
</tr>
<tr>
<td>HPV 16</td>
<td>135</td>
<td>136</td>
<td>138</td>
<td>284</td>
<td>693</td>
</tr>
<tr>
<td>HPV 18</td>
<td>118</td>
<td>105</td>
<td>124</td>
<td>248</td>
<td>595</td>
</tr>
<tr>
<td>Reason for exclusion:*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General protocol violation</td>
<td>72</td>
<td>60</td>
<td>78</td>
<td>141</td>
<td>351</td>
</tr>
<tr>
<td>Day 1 serum or swab± sample results missing</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Missing Month 7 serum results</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>Month 7 serum sample out of day range</td>
<td>32</td>
<td>36</td>
<td>30</td>
<td>71</td>
<td>169</td>
</tr>
<tr>
<td>Missing Month 7 swab results‡</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>Positive to HPV 6 or 11†</td>
<td>28</td>
<td>28</td>
<td>31</td>
<td>61</td>
<td>148</td>
</tr>
<tr>
<td>Positive to HPV 16†</td>
<td>41</td>
<td>42</td>
<td>32</td>
<td>71</td>
<td>186</td>
</tr>
<tr>
<td>Positive to HPV 18†</td>
<td>15</td>
<td>9</td>
<td>12</td>
<td>33</td>
<td>69</td>
</tr>
</tbody>
</table>

*Subjects were counted once in each applicable exclusion category. A subject may appear in more than one category.
†Exclusions based on swab samples apply only to subjects ≥16 years of age.
‡Seropositive at Day 1 and/or (for subjects ≥16 years of age) PCR positive at or before Month 7 to the relevant HPV type(s) protocol population for the relevant HPV type(s) only.

N = Number of subjects randomized to the respective vaccination group.

HPV = Human papillomavirus; VLP = Virus-like particle; PCR = Polymerase chain reaction; PP = Per protocol.

---

34  BioProcess International  MAY 2007
Table 4: Per-protocol analysis of noninferiority comparing the proportions of subjects who seroconverted to vaccine HPV types at Month 7

<table>
<thead>
<tr>
<th>Group A vs. Group B</th>
<th>n</th>
<th>Response (%)</th>
<th>Group B</th>
<th>n</th>
<th>Response (%)</th>
<th>Percentage Point Difference Group A - Group B (95% CI)</th>
<th>p-Value for Noninferiority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-HPV 6</td>
<td>20% vs. 100%</td>
<td>375</td>
<td>100</td>
<td>746</td>
<td>100</td>
<td>0.0 (–1.0, 0.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>40% vs. 100%</td>
<td>393</td>
<td>100</td>
<td>746</td>
<td>100</td>
<td>0.0 (–1.0, 0.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>60% vs. 100%</td>
<td>370</td>
<td>99.7</td>
<td>746</td>
<td>100</td>
<td>–0.3 (–1.5, 0.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Anti-HPV 11</td>
<td>20% vs. 100%</td>
<td>375</td>
<td>100</td>
<td>746</td>
<td>100</td>
<td>0.0 (–1.0, 0.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>40% vs. 100%</td>
<td>393</td>
<td>100</td>
<td>746</td>
<td>100</td>
<td>0.0 (–1.0, 0.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>60% vs. 100%</td>
<td>370</td>
<td>99.7</td>
<td>746</td>
<td>100</td>
<td>–0.3 (–1.5, 0.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Anti-HPV 16</td>
<td>20% vs. 100%</td>
<td>368</td>
<td>100</td>
<td>733</td>
<td>100</td>
<td>0.0 (–1.0, 0.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>40% vs. 100%</td>
<td>377</td>
<td>100</td>
<td>733</td>
<td>100</td>
<td>0.0 (–1.0, 0.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>60% vs. 100%</td>
<td>370</td>
<td>99.7</td>
<td>733</td>
<td>100</td>
<td>–0.3 (–1.5, 0.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Anti-HPV 18</td>
<td>20% vs. 100%</td>
<td>385</td>
<td>99.7</td>
<td>769</td>
<td>99.6</td>
<td>0.1 (–1.1, 0.9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>40% vs. 100%</td>
<td>408</td>
<td>99.3</td>
<td>769</td>
<td>99.6</td>
<td>–0.3 (–1.8, 0.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>60% vs. 100%</td>
<td>384</td>
<td>99</td>
<td>769</td>
<td>99.6</td>
<td>–0.6 (–2.3, 0.3)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

CI = Confidence interval; mMU = Milli Merck units; HPV = Human papillomavirus

postulates that these influences affect the final container IVRP multiplicatively.

The following equation was used to calculate Final Container min for each HPV type:

\[
\text{Final Container}_{\text{min}} = \text{GM}_{\text{FC}} + (1 + \text{RSD}_{\text{FC}}/100)^3
\]

where \( \text{GM}_{\text{FC}} \) and \( \text{RSD}_{\text{FC}} \) are the geometric mean and relative standard deviation derived from the model for final container IVRP. The \( \text{GM}_{\text{FC}} \) for each HPV type is determined by multiplying the corresponding geometric mean for the monovalent bulk (\( \text{GM}_{\text{MB}} \)) by factors that affect the final container IVRP mean release value: stability loss of the monovalent bulk before final container formulation and the dilution factor used during formulation.

\[
\text{GM}_{\text{FC}} = \text{GM}_{\text{MB}} \times f_{\text{stability}} \times f_{\text{dilution}}
\]

The terms in that equation for the \( \text{GM}_{\text{FC}} \) of the vaccine are:

- \( \text{GM}_{\text{MB}} \): the geometric mean of the monovalent bulk IVRP values. This term was estimated as described above.
- \( f_{\text{stability}} \): the influence due to change in bulk IVRP over time. This term takes into account the expected loss during the maximum monovalent bulk hold time (currently 36 months). The term \( f_{\text{stability}} \) was estimated using all available stability data as described below. In calculation of the \( \text{GM}_{\text{FC}} \), \( f_{\text{stability}} \) is expressed as a fraction of the bulk release IVRP value and calculated by dividing the predicted IVRP at 36 months by that at release.
- \( f_{\text{dilution}} \): the dilution factor. This term is a fixed value for each HPV type based on the target protein concentration in the final container for each type. The dilution factor is 0.125 for types 6 and 18, which are diluted from 320 µg/mL protein to 40 µg/mL protein during formulation, and 0.25 for types 11 and 16, which are diluted from 320 µg/mL protein to 80 µg/mL protein.

The \( \text{RSD}_{\text{FC}} \) is derived based on a propagation-of-error calculation as shown in the Equation box below.

The terms in the equation for the RSD are described below. The observed final container variability primarily comes from the lot-to-lot variability of the monovalent bulks, which are diluted based on protein concentration rather than IVRP, and the assay variability, which is typical of an ELISA-type assay.

- \( \text{RSD}_{\text{dilution}} \): the estimated variability of the dilution step, which was determined based on experimental data from the manufacturing facility. This term is different from \( \text{RSD}_{\text{form/fill}} \) because it reflects errors encountered only during the weighing process. \( \text{RSD}_{\text{dilution}} \) was evaluated theoretically using a separate propagation-of-error calculation.
- \( \text{RSD}_{\text{stability}} \): the estimated variability in the estimate of the mean loss rate. The value was calculated as described below.
- \( \text{RSD}_{\text{slope}} \): the estimated lot-to-lot variability in the slopes (i.e., loss rates) of the stability profiles. The value was calculated as described below.

Equation: Propagation-of-error calculation

\[
\text{RSD}_{\text{FC}} = \sqrt{\text{RSD}_{\text{form/fill}}^2 + \text{RSD}_{\text{stability}}^2 + \text{RSD}_{\text{slope}}^2 + \text{RSD}_{\text{dilution}}^2 + \text{RSD}_{\text{assay}}^2}
\]

36 BioProcess International MAY 2007
As noted, the terms \( RSD_{\text{stability}} \) and \( RSD_{\text{slope}} \) were included in the model to account for the impact of stability loss during storage of monovalent bulks and the impact of this loss on the final container IVRP. These were evaluated on both the bulks and final container. The analysis strategy involved first assessing differences among the loss rates corresponding to the different types of images (market container/closure system), then characterizing the stability profiles of bulks and final container lots using a mixed model analysis (18). The mixed model analysis treats loss rate as a random variable and, therefore, besides providing estimates of the mean loss rate \( \bar{f}_{\text{loss}} \) and the variability in this estimate \( RSD_{\text{stability}} \), it also estimates the variability in the loss rates \( RSD_{\text{slope}} \) among the stability profiles.

**Stability Specifications:** As shown in Figure 1, stability specifications were subsequently derived from the release specifications limits using a similar model to that developed for the release limits. The terms for the mean loss rate and variability among the loss rates, discussed above, need to be included twice in the model to account for both bulk and final container loss in potency.

Calculation of the lower stability specification \( IVRP_{\text{stability}} \) is similar to that used to derive the lower bulk release specification and can be described mathematically as

\[
IVRP_{\text{stability}} = \text{Final Container}_{\text{min}} \times \left( 1 + \frac{RSD_{\text{FC stability}}^2}{100} \right)^{-1}
\]

where \( RSD_{\text{FC stability}} \) is the total variation in final container IVRP at expiry. Each lot placed on stability will be tested multiple times. As the number of tests on a single lot increases, the probability that a single reportable test result will fall below the stability limit due to assay variability alone also increases. This type of failing result is due to statistical multiplicity and does not reflect an unacceptable change in the product. For a lot that has a true IVRP at release equal to the lower specification \( IVRP_{\text{min}} \), the use of “2” in the exponent fixes the risk, at a given time point, of generating a failing reportable test result by chance variation alone to be approximately 2.5%.

\[ RSD_{\text{FC stability}} = \sqrt{RSD_{\text{stability}}^2 + RSD_{\text{slope}}^2 + RSD_{\text{assay}}^2} \]

The first two terms are associated with final container stability. Their numerical values are the same as those used in the derivation of \( RSD_{\text{FC}} \) because analysis of the stability data suggested that there was no statistically significant difference in stability profiles among the bulks and final container images. Because it was assumed that the final container lot starts with a true IVRP at the lower release limit, this equation did not include variability associated with bulk manufacture, formulation, fill, and storage. Table 2 shows numerical values associated with each of these terms using type 16 as an example.

**Clinical Study Design**

Protocol 016 (ClinicalTrials.gov number NCT00092495) was a randomized, multicenter dose-expiry study. Between 7 December 2002 and 15 July 2003, 2,594 subjects were enrolled from 61 centers located throughout 19 countries. The study enrolled nonpregnant, healthy female subjects, 10–15 years of age, who had never been and did not plan to become sexually active through the course of the study. It also enrolled nonpregnant, healthy women 16–23 years of age who reported no prior abnormal pap smears and a lifetime history of four or fewer male sex partners. An Institutional Review Board at each clinical site approved the study protocol. At enrollment, written consent was obtained from each participant or her legal guardian.

The full- and partial-dose formulations of the quadrivalent HPV vaccine (Table 1) were supplied in identical vials and were visually indistinguishable. Subjects were randomized in a 1:1:1:2 ratio to receive three intramuscular injections of a 20%, 40%, 60%, or 100% dose formulation of quadrivalent HPV vaccine at Day 1, Month 2, and Month 6.
Clinical Follow-Up: Blood samples were obtained from all subjects at Day 1, Month 3, and Month 7. Serum antibodies to HPV 6, 11, 16, and 18 were measured using a competitive Luminex immunoassay (cLIA) and reported in arbitrary units, milliMerck units per mL (a relative measure defined by comparison to an internal reference sample), or mMU/mL as described previously \((19, 20)\). The seropositive cutoffs were 20, 16, 20, and 24 mMU/mL for HPV 6, 11, 16, and 18, respectively \((19, 20)\). An audit conducted by Merck Research Laboratories between 1 May 2006 and 16 May 2006 concluded that there was a deviation from the standard operating procedure (SOP) for testing a subset of serum samples from the protocol. Approximately 15 Month 7 samples, distributed among the four quadrivalent vaccine formulations, were determined to have been tested outside of the SOP. Those samples were included in the analyses presented here.

Statistical Analyses of Clinical Data: The primary immunogenicity hypothesis of the study stated that at least one partial-dose formulation of the quadrivalent HPV vaccine (containing 20%, 40%, or 60% of each VLP component), given in a three-dose regimen, induces noninferior immune responses compared to administration of a three-dose regimen of full-dose quadrivalent HPV vaccine, as measured by the percentages of subjects who seroconvert for each of HPV types 6, 11, 16, and 18 by four weeks postdose three (Month 7). A subject seroconverted for a given HPV type if she achieved a Month 7 anti-HPV level greater than the seropositivity threshold value for that HPV type (defined as 20, 16, 20, and 24 mMU/mL for HPV types 6, 11, 16, and 18, respectively).

Noninferiority tests of proportions were conducted based on methods developed by Miettinen and Nurminen \((21)\), with stratification by geographic region. The three partial-dose formulations were compared with the full-dose formulation for the four HPV types following the same testing strategy used for GMTs. To reject the null hypothesis for a given HPV type, the lower bound of the 95% CI on the difference in percentages of seroconverters between the two comparison groups had to be greater than \(-0.05\).

Analyses were conducted in three per-protocol populations, one each for HPV 6/11, 16, and 18. These populations consisted of:

- subjects who were naïve to the relevant HPV type(s) (defined as seronegative to the relevant type(s) at Day 1, and for subjects ≥16 years of age, PCR negative to the relevant HPV type(s) from Day 1 through Month 7) based on assessment of serum and (for subjects ≥16 years of age) genital swab samples collected within protocol-specified time frames
- subjects who did not violate the protocol in ways that might have interfered with immune responses, as determined before unblinding
- subjects who had a Month 7 serum sample collected within the protocol-specified time frame.

Results and Discussion: A modified process capability model was used to derive release and stability limits. The clinical relevance of these limits was confirmed from data obtained from a clinical potency-ranging study. The purpose of that study was to ensure that the stability limit was set at a level well above any drop-off in the dose-response curve. When the propagation-of-error approach is combined with clinical results, the resulting release and stability limits ensure that the potency of future material will be consistent with material used in the pivotal clinical trials and provide reasonable confidence that the released material will remain effective over the full shelf-life of the vaccine.

The process used to derive the release and stability specification is described in detail in the “Materials and Methods” section and shown schematically in Figure 1. Briefly, the following three-step process was used:

- Release limits for monovalent bulks were calculated based on an assessment of bulk process capability.
- The lower release limit for final container material was calculated using a propagation-of-error calculation that accounts for each of the factors that impacts the mean final container IVRP and the variation in final container IVRP.
- The lower stability limit was derived similarly to the lower release limit assuming a final container lot started with a true IVRP equal to the lower release limit and accounting for statistical multiplicity.

A discussion of those steps is provided below using the results.
obtained for type 16 as an example. The process followed for the other types was the same.

**Derivation of Bulk Release Specification**

Data were available on a total of 25 HPV type 16 monovalent bulk lots, making it possible to derive a lower bulk release specification directly by calculating a three-sigma limit (Panel A, Figure 1). The derivation of the bulk release specification required the geometric mean IVRP and total variability (RSD), which is composed of both process variability and assay variability. For type 16, the geometric mean IVRP, total RSD and lower three-sigma limit were calculated to be 311 units/mL, 10.5% and 232 units/mL, respectively.

As described in “Materials and Methods,” IVRP results are reported relative to a final container lot used in the clinic that was assigned an IVRP value 80 units/mL. The IVRP for the standard was assigned based on the protein concentration and does not account for the specific activity or variation in the true protein concentration of the lot selected. As a result, bulk lots with specific activities or true protein concentrations greater than the reference standard exhibit IVRP values >320 units/mL, and bulk lots with specific activities lower than the reference standard lot exhibit IVRP values <320 units/mL. The observation that the geometric mean IVRP for type 16 bulks is 311 units/mL indicates that the type 16 VLPs contained in the reference standard lot have a slightly higher specific activity than the average lot.

**Final Container Release Specification:** As shown in Panel B of Figure 1, Final Container min was calculated using the data obtained for the monovalent bulks as a starting point. It was assumed that a randomly selected bulk lot was obtained and subjected to routine formulation, dilution and storage conditions before filling. Bulk lots can be stored before formulation, and loss during that time will directly affect the geometric mean final container IVRP. Therefore, the geometric mean bulk IVRP was adjusted for each type to account for potential stability loss.

In general, no statistically significant difference in loss rate was detected among the bulks and final container images. Further, the sample matrix used for the bulks is the same as that used for the final containers, and there is no evidence that the VLP stability profile is dependent on protein concentration. Therefore, data obtained on both bulk and final container lots were pooled to estimate

---

**Figure 2:** Per-protocol immunogenicity analysis of noninferiority and confidence intervals comparing Month 7 geometric mean antibody titers for full dose (100%) and partial dose (20%, 40%, 60%) formulations for HPV 6 (a), HPV 11 (b), HPV 16 (c) and HPV 18 (d).
For type 16, the loss rate was estimated to be –0.05% per year (a statistically significant but biologically meaningless rate). This was neither statistically significant nor biologically meaningful, and therefore a loss rate of 0% per year was assumed to be conservative.

During formulation, monovalent bulks are diluted to the appropriate concentration based on mass (i.e., protein concentration), a dilution fixed for each type. In the case of type 16, each bulk is diluted four-fold from a starting protein concentration of 320 µg/mL to a final protein concentration of 80 µg/mL. During this dilution process, weighing and transfer errors contribute to variability in the estimate of the mean slopes of the stability profiles, as well as the variability in the overall variation in final container process variation.

The variability contribution from each of those terms was combined to give an estimate of the overall final container variability. The major sources of final container variation results were lot-to-lot variation, variability in the slopes of the stability profiles, variability in the estimate of the mean loss rates, weighing and transfer errors during formulation and filling, and vial-to-vial heterogeneity.

The variability contribution from each of those terms was combined to give an estimate of the overall final container variability. The major sources of final container variation results were lot-to-lot variation, variability in the slopes of the stability profiles, variability in the estimate of the mean loss rates, weighing and transfer errors during formulation and filling, and vial-to-vial heterogeneity.

The procedure used to define the release and stability specifications for Gardasil offers several ADVANTAGES over existing models.
APPICIBILITY OF THE MODEL

Based on the observation that the clinical responses obtained at the 20% level were statistically noninferior to the 100% dose-level, it is reasonable to assume that the dose-response curve does not exhibit a sharp drop-off across the range of potencies that were simulated in this study. It should be noted that this study was not designed to evaluate long-term persistence of the antibody response, and it is not known whether the duration of the immune response is equivalent across all doses tested. Although the study did not evaluate the long-term persistence of the antibody response, the data provide a significant level of assurance that material provided to clinicians, corresponding to material formulated at the 60% dose level or above, will be immunogenic and remain efficacious across the full-shelf life of the vaccine.

The procedure used to define Gardasil’s release and stability specifications offers several advantages over existing models. In the absence of an immune correlate of protection, the model used for the vaccine links the statistically derived release and stability limits with clinical data providing a high level of confidence that the released material will remain efficacious throughout its shelf life. Additionally, the limits established here are based on bulk process capability data and will ensure that the process is well controlled and that the potency of future lots of the Gardasil vaccine will be consistent with materials used in the pivotal clinical trials. It is anticipated that this model for establishing specifications will be applicable to a wide-range of antigen-based vaccines when clinical efficacy data do not exist to establish a minimum effective dose and/or when there are only limited final container lots available to calculate the limits directly.

ACKNOWLEDGMENTS

The authors thank Drs. Pat Brill-Edwards, Anne Aunins, Yang Wang, David Wohlpart, James Klein, Ann Lee, and Kathryn Jansen for guidance and helpful discussions pertaining to this work. They also acknowledge Drs. Clint Potter and Bridgette Carragher (Scrpps) for their support of cryoEM studies on HPV virus-like particles.

REFERENCES


All authors are from Merck Research Laboratories. Robert Capen is associate director nonclinical statistics; corresponding author Mary L. Shank-Retzlaff is research fellow, bioprocess and bioanalytical research, Merck Research Laboratories, WP42A-30, PO Box 4, West Point, PA 19486; 1-215-652-3835, fax 1-215-993-4846, mary_retzlaff@merck.com; Heather L. Sings is associate director medical communications; Mark Esser is research fellow vaccine and biologics research; Charles Sattler is medical director policy, public health, and medical affairs; Michael Washabaugh is senior director bioprocess and bioanalytical research; and Robert Sitrin is executive director, bioprocess and bioanalytical research.