Establishing Potency Specifications for Antigen Vaccines

Clinical Validation of Statistically Derived Release and Stability Specifications

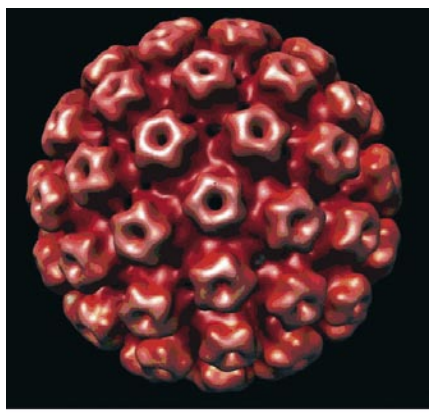
Robert Capen, Mary L. Shank-Retzlaff, Heather L. Sings, Mark Esser, Carlos Sattler, Michael Washabaugh, and Robert Sitrin

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 immunoassay, 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. The statistically derived specifications were



Three-dimensional reconstruction of a human papillomavirus (L1 type 11) virus-like particle, a component of Gardasil®.
COURTESY OF MERCK & CO., INC. AND THE NATIONAL RESOURCE FOR AUTOMATED MOLECULAR MICROSCOPY, THE SCRIPPS RESEARCH INSTITUTE

PRODUCT FOCUS: VACCINES

PROCESS FOCUS: DOWNSTREAM PROCESSING, FORMULATION DEVELOPMENT

WHO SHOULD READ: FORMULATION SCIENTISTS, STATISTICIANS, FILL AND FINISH

KEYWORDS: GARDASIL, SPECIFICATIONS, POTENCY, HUMAN PAPILLOMAVIRUS

LEVEL: ADVANCED

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 each 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 samples 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 serial dilutions were prepared and transferred to the assay plate. To dissolve the aluminum-containing adjuvant, a citrate-phosphate dissolution buffer was also 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:

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

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

Table 1: Formulations of quadrivalent HPV (6/11/16/18) L1 VLP vaccine used in the potency ranging clinical study

Quadrivalent HPV L1 VLP Vaccine Formulation	HPV 6 VLP (µg)	HPV 11 VLP (µg)	HPV 16 VLP (µg)	HPV 18 VLP (µg)	Aluminum (µg)
100%	20	40	40	20	225
60%	12	24	24	12	225
40%	8	16	16	8	225
20%	4	8	8	4	225

Table 2: Summary of terms used to calculate release and stability specifications, the associated numerical values used in the statistical calculations for type 16, and the resulting release and stability limits

Variable	Description	Expected Value for Type 16
GM_{MB}	Geometric mean IVRP for monovalent bulks used to formulate final container lots of Gardasil	311 Units/mL
$f_{dilution}$	Dilution factor used during formulation	0.25
$f_{stability}$	Mean stability loss (slope)	1.0 (0% loss)
GM_{FC}	Expected geometric mean IVRP for final container lots of Gardasil	78 Units/mL
RSD_{FC}	Variation in IVRP for final container lots of Gardasil at release	13.4%
$RSD_{FC\,stability}$	Variation in IVRP for final container lots of Gardasil at expiry	9.3%
$IVRP_{MB}$	Minimum monovalent bulk release limit	232 Units/mL
$Final\ Container_{min}$	Minimum final container release limit	≥53 Units/mL
$Final\ Container_{stability}$	Minimum final container stability limit	≥44 Units/mL

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 implementation 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. Panel A, Figure 1, shows this schematically. Data on all available bulk lots were analyzed to determine the geometric mean bulk IVRP, lot-to-lot variation, and assay variability. Lot-to-lot and assay variability were assessed using test results generated by the quality control laboratory in which all future release testing will be performed.

The estimates of lot-to-lot standard deviation (s_{MB}) and assay standard deviation (s_{assay}) were combined to give an estimate of the total standard deviation (s_{total}) using the following equation:

$$S_{total} = \sqrt{S_{MB}^2 + S_{Assay}^2}$$

The term for assay variability in the above equation reflects variability in the reportable value, which is the average of three independent IVRP determinations. The estimate of percent relative standard deviation (RSD) is calculated from the estimated standard deviation as follows:

$$RSD = 100\,(e^{S} - 1)$$

Here, s represents s_{MB}, s_{assay}, or s_{total}, depending on the context. To better approximate a normal distribution, the natural logarithm of each IVRP measurement (natural-logarithm-transformed IVRP value) was used in the statistical analysis that estimated these quantities.

From these estimates, lower process capability limits were calculated in the log scale and then converted to the linear scale as follows:

$$\ln(Monovalent\ Bulk_{min}) = \bar{y} - 3s_{total}$$

$$Monovalent\ Bulk_{min} = GM_{MB} \div (1 + RSD_{total}/100)^3$$

where $\bar{y}$ is the mean of the natural-logarithm-transformed IVRP values for the monovalent bulks, GM_{MB} is the corresponding geometric mean found by exponentiating $\bar{y}$, and s_{total} and RSD_{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

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

* 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

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

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

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:

$$Final\ Container_{min} = GM_{FC} \div (1 + RSD_{FC}/100)^3$$

where GM_{FC} and RSD_{FC} are the geometric mean and relative standard deviation derived from the model for final container IVRP. The GM_{FC} for each HPV type is determined by multiplying the corresponding geometric mean for the monovalent bulk (GM_{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.

$$GM_{FC} = GM_{MB} \times f_{stability} \times f_{dilution}$$

The terms in that equation for the GM_{FC} of the vaccine are

GM_{MB}: the geometric mean of the monovalent bulk IVRP values. This term was estimated as described above.

$f_{stability}$: 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 $f_{stability}$ was estimated using all available stability data as described below. In calculation of the GM_{FC}, $f_{stability}$ 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.

$f_{dilution}$: 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 RSD_{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.

RSD_{MB}: the estimated lot-to-lot variability of the monovalent bulks of each HPV type. This term incorporates only process variability.

$RSD_{form/fill}$: the estimated variability from the formulation and filling processes. Heterogeneity due to settling of VLP-adjuvant particles is the expected source of variability for the formulation and filling processes. This can be estimated from the variability of the aluminum concentration because VLPs are fully adsorbed to the aluminum adjuvant. Therefore, $RSD_{form/fill}$ was determined using data from the full-scale formulation and filling process validation lots for both vials and syringes. The $RSD_{form/fill}$ value was evaluated by measuring the variability of aluminum concentration among individual vials or syringes for each of six lots (three syringe lots and three vials lots) at multiple locations across each fill.

$RSD_{dilution}$: the estimated variability of the dilution step, which was determined based on experimental data from the manufacturing facility. This term is different from $RSD_{form/fill}$ because it reflects errors encountered only during the weighing process. $RSD_{dilution}$ was evaluated theoretically using a separate propagation-of-error calculation.

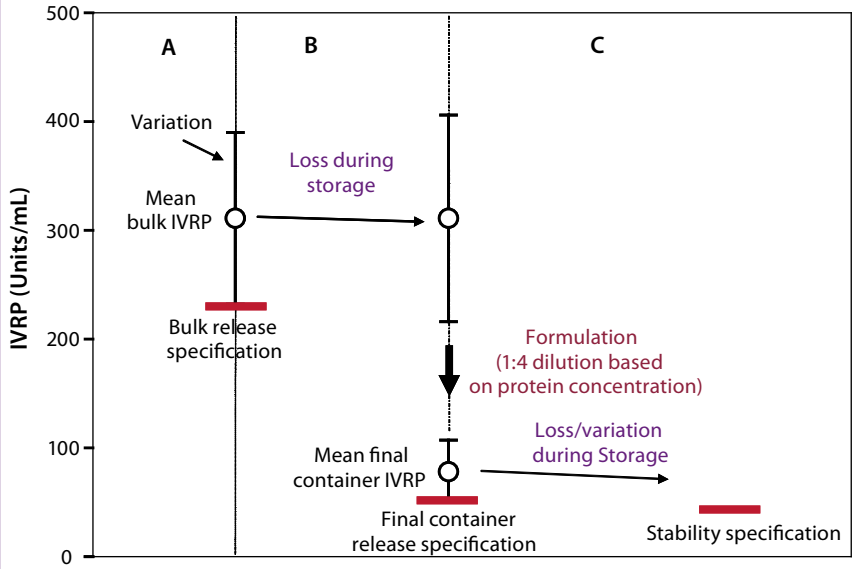
$RSD_{stability}$: the estimated variability in the estimate of the mean loss rate. The value was calculated as described below.

RSD_{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

$$RSD_{FC} = \sqrt{RSD_{MB}^2 + RSD_{stability}^2 + RSD_{slope}^2 + RSD_{dilution}^2 + RSD_{form/fill}^2 + RSD_{assay}^2}$$

Figure 1: Schematic diagram of the process used to derive release and stability specifications for final container lots of Gardasil; (A) the geometric mean bulk IVRP and the associated variation were determined by evaluating release results from 25 bulk lots; (B) the mean final container IVRP was estimated by multiplying the mean bulk IVRP by the dilution factor (0.25 for types 16 and 11, 0.125 for type 6 and 18) and accounting for the anticipated loss during storage of the bulks. The variation in final container IVRP values was estimated by accounting for the variation in bulk IVRPs, variation in the slopes of the stability profiles, variation due to formulation and filling and assay variability. (C) Stability specifications were subsequently calculated by accounting for loss during final container storage, variation in the final container slopes, assay variability and statistical multiplicity.



RSD_{assay}: the estimated assay variability.

As noted, the terms $f_{stability}$ and $RSD_{stability}$ 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 ($f_{stability}$) and the variability in this estimate ($RSD_{stability}$), it also estimates the variability in the loss rates (RSD_{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 ($IVRP_{stability}$) is similar to that used to derive the lower bulk release specification and can be described mathematically as

$$Final\ Container_{stability} = Final\ Container_{min} \div \left(1 + \frac{RSD_{FC\ stability}}{100}\right)^2$$

where $RSD_{FC\ stability}$ 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 ($Final\ Container_{min}$), 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%.

$RSD_{FC\ stability}$ was calculated using the following equation:

$$RSD_{FC\ stability} = \sqrt{RSD^2_{stability} + RSD^2_{slope} + RSD^2_{assay}}$$

The first two terms are associated with final container stability. Their numerical values are the same as those used in the derivation of RSD_{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 (**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 for each of HPV types 6, 11, 16, and 18, as measured by the geometric mean titers (GMTs) at four weeks postdose three (Month 7). For each pairwise comparison of formulations (partial-dose with full-dose), four tests of noninferiority (one per HPV type) were conducted at the 0.025 level.

A nested testing procedure of formulation comparisons was used to ensure an overall one-sided type-1 error rate ≤0.025. Analysis of variance (ANOVA) was used to model the natural log of the Month 7 anti-HPV cLIA antibody level as a function of age group (<16 compared with ≥16 years of age), geographic region, and vaccination group. The anti-log of the estimated vaccination group difference in the ANOVA model and the confidence interval (CI) associated with this difference were computed. The statistical criterion for noninferiority for a given HPV type required that the lower bound of the 95% CI on the ratio of GMTs between the two comparison groups exceed 0.5 for that HPV type. For a partial-dose formulation to be declared noninferior to the full-dose formulation, it was necessary for the partial-dose formulation to meet the statistical criterion for noninferiority for all four HPV types.

The secondary immunogenicity hypothesis of the clinical study was that at least one partial-dose formulation of the quadrivalent HPV vaccine, 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.

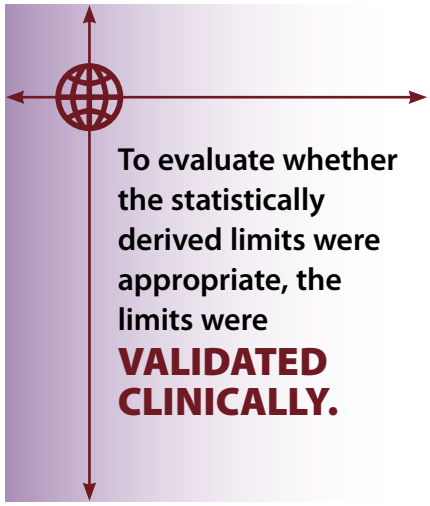
A discussion of those steps is provided below using the results

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.

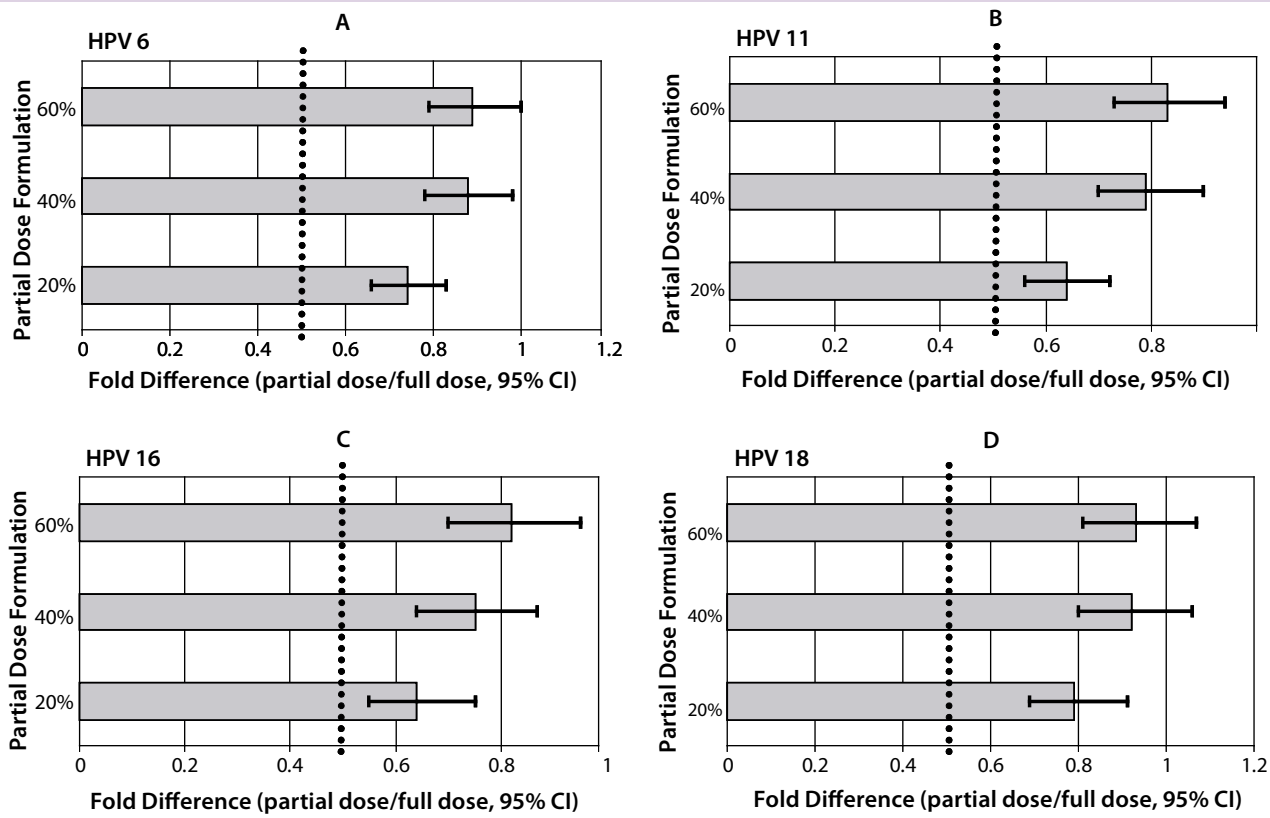
As described in "Materials and Methods," IVRP results are reported relative to a final container lot used in the clinic that was assigned an IVRP value 80 units/mL. The IVRP for the standard was assigned based on the protein concentration and does not account for the specific activity or variation in 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).

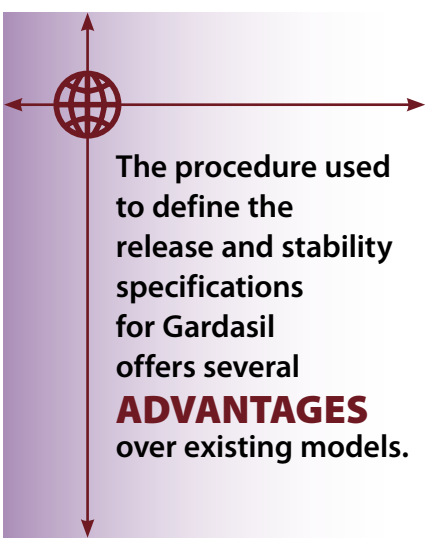


a common (mean) loss rate per type. For type 16, the loss rate was estimated to be –0.05% per year (a gain in potency). This was neither statistically significant nor biologically meaningful, and therefore a loss rate of 0% per year was assumed to be conservative.

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 protein to a final protein concentration of 80 µg/mL. Taking the dilution factor and the anticipated loss during bulk storage into account, the geometric mean final container IVRP was predicted to be 78 units/mL for this type.

Final container process variation was estimated by evaluating each of the factors that contribute to this variation. These factors include bulk 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 variability of the monovalent bulks, variation in the loss rates, and assay variability. Using the estimated geometric mean final-container IVRP (78 units/mL) and the overall variation in final container IVRP (13.4%), the lower release limit was determined to be 53 units/mL.



Stability Specifications: Stability specifications were subsequently derived from the release specification by accounting for the anticipated loss over the shelf life of the product and statistical multiplicity (Panel C, Figure 1). The total variation in final container IVRP at expiry was estimated to be 9.3% for type 16, yielding a stability limit of ≥44 units/mL.

Clinical Validation of Statistically Derived Specifications: To evaluate whether the statistically derived limits were appropriate, the limits were validated clinically. It was not feasible to generate low-potency samples formulated at the 100% dose level due to the inherent stability of the vaccine. It is estimated that the half-life of the vaccine is well over 100 months at 2–8 °C, and stress studies have shown that the only way to generate material with significantly reduced IVRP-to-protein ratios is to expose the material to extreme physical or chemical stress (**17**). When the stressed samples are removed from adjuvant for IVRP testing, the VLPs are prone to aggregation. Although the IVRP of those samples can be greatly reduced, aggregated samples are not considered representative of material that will be observed during routine manufacturing and storage. The stressed samples are considered unsuitable for use in humans for ethical reasons.

Therefore, low-potency samples were simulated by preparing final container samples formulated at reduced protein concentration. Lots were formulated at 20%, 40%, and 60% of the standard formulation by diluting the 100% dose formulation used in the same study with an excipient-balanced diluent containing the aluminum adjuvant. The statistically derived release specification (53 units/mL) roughly corresponds to material formulated at a 60% dose level. The stability specification corresponds to material formulated at roughly a 50% dose level.

Of the 2,541 subjects who received at least one injection, 74.1% were included in the anti-HPV 6/11 per protocol population, 72.7% were included in the anti-HPV 16 per protocol population, and 76.6% were included in the anti-HPV 18 per protocol population (Table 3). The proportions of subjects who met each reason for exclusion were generally comparable among the four vaccination groups. The most common reasons for exclusion from the per protocol population were baseline positivity for one or more vaccine HPV types and general protocol violations.

Figure 2 displays the statistical analysis of noninferiority comparing Month 7 anti-HPV GMTs between subjects who received the partial-dose formulations of quadrivalent vaccine and subjects who received the full-dose formulation of quadrivalent vaccine. The figure displays the estimated fold difference (partial dose/full dose) between the two groups and the corresponding 95% CI for the fold difference. The ratios of the partial-dose formulation to the full-dose formulation ranged from 0.64 (95% CI [0.56, 0.72]) to 0.93 (95% CI [0.81, 1.07]). For each HPV type, a dose response was observed, with the 20% formulation group having the lowest GMT and the 100% formulation having the highest GMT. However, all partial-dose formulations were found to be noninferior to the full-dose formulation according to the prespecified criteria ($p \leq 0.001$ for all comparisons).

Table 4 shows that at Month 7, for each of the 4 HPV types, the lower bound of the 95% CI for the difference in the proportions of subjects who became seropositive between the two groups (partial-dose minus full-dose) excluded a decrease of five percentage points or more for each HPV type ($p < 0.001$ for all comparisons). Therefore, all partial-dose formulations were found to be noninferior to the full-dose formulation with respect to the percentages of subjects who seroconverted for each of HPV types 6, 11, 16 and 18.

Applicability 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 (Scripps) for their support of cryoEM studies on HPV virus-like particles.

References

1 Cook JC, et al. Purification of Virus-Like Particles of Recombinant Human Papillomavirus Type 11 Major Capsid Protein L1 from *Saccharomyces Cerevisiae*. *Protein Expr. Purif.* 17(3) 1999: 477–484.

2 Lowe RS, et al. Human Papillomavirus Type 11 (HPV-11) Neutralizing Antibodies in the Serum and Genital Mucosal Secretions of African Green Monkeys Immunized with HPV-11 Virus-Like Particles Expressed in Yeast. *J. Infect. Dis.* 176(5) 1997: 1141–1145.

3 Kirnbauer R, et al. Papillomavirus L1 Major Capsid Protein Self-Assembles into Virus-Like Particles That Are Highly Immunogenic. *Proc. Natl. Acad. Sci*. 89(24) 1992: 12180–12184.

4 Hagensee ME, et al. Three-Dimensional Structure of Vaccinia Virus-Produced Human Papillomavirus Type-1 Capsids. *J. Virol.* 68(7) 1994: 4503–4505.

5 Chen XJS, et al. Structure of Small Virus-Like Particles Assembled from the L1 Protein of Human Papillomavirus 16. *Mol. Cell*; 5(3) 2000: 557–567.

6 Shank-Retzlaff ML, et al. Correlation Between Mouse Potency and In Vitro Relative Potency for Hum. Papillomavirus Type 16 Virus-Like Particles and Gardasil Vaccine Samples. *Human Vaccin*. 1(5) 2005: 191–197.

7 Christensen ND, et al. Monoclonal Antibody-Mediated Neutralization of Infectious Human Papillomavirus Type 11. *J. Virol.* 64(11) 1990: 5678–5681.

8 Wang ZH, et al. A Monoclonal Antibody Against Intact Human Papillomavirus Type 16 Capsids Blocks the Serological Reactivity of Most Human Sera. *J. Gen. Virol.* 78(9) 1997: 2209–2215.

9 Yeager MD, et al. Neutralization of Human Papillomavirus (HPV) Pseudovirions: A Novel and Efficient Approach to Detect and Characterize HPV Neutralizing Antibodies. *Virology* 278(2) 2000: 570–577.

10 Roden RBS, et al. Characterization of a Human Papillomavirus Type 16 Variant-Dependent Neutralizing Epitope. *J. Virol.* 71(8) 1997: 6247–6252.

11 Villa LL, et al. Prophylactic Quadrivalent Human Papillomavirus (Types 6, 11, 16, and 18) L1 Virus-Like Particle Vaccine in Young Women: A Randomised Double-Blind Placebo-Controlled Multicentre Phase II Efficacy Trial. *Lancet Oncol.* 6(5) 2005: 271–278.

12 Villa LL, et al. High Sustained Efficacy of a Prophylactic Quadrivalent Human Papillomavirus (HPV) Types 6/11/16/18 L1 Virus-Like Particle Vaccine Through 5 Years of Follow-Up. *Br. J. Cancer.* 95(11) 2006: 1459–1466.

13 Sattler C and Investigators for the Futures I Investigators. Efficacy of a Prophylactic Quadrivalent Human Papillomavirus (HPV) (Types 6, 11, 16, 18) L1 Virus-Like Particle (VLP) Vaccine for Prevention of Cervical Dysplasia and External Genital Lesions. *Presented at the 45th Interscience Conference on Antimicrobial Agents and Chemotherapy,* 16–19 December 2005, Washington, D.C; American Society for Microbiology: Washington, D.C.

14 Skjeldestad FE and Committee for the Futures II Steering Committee. Prophylactic Quadrivalent Human Papillomavirus (HPV) (Types 6, 11, 16, 18) L1 Virus-like Particle (VLP) Vaccine (Gardasil) Reduces Cervical Intraepithelial Neoplasia (CIN) 2/3 Risk. Presented at the *43rd Infectious Diseases Society of America*, 6–9 October 2005, Alexandria, VA; Infectious Diseases Society of America: San Francisco, CA.

15 Mach H, et al.. Disassembly and Reassembly of Yeast-Derived Recombinant Human Papillomavirus Virus-Like Particles (HPV VLPs). *J. Pharm. Sci.* 95(10) 2006; 2195–2206.

16 McCarthy MP, et al. Quantitative Disassembly and Reassembly of Human Papillomavirus Type 11 Viruslike Particles In Vitro. *J. Virol.* 72(1) 1998: 32–41.

17 Shank-Retzlaff ML, et al. Evaluation of the Thermal Stability of Gardasil. *Hum.Vaccin.* 2(4) 2006: 147–154.

18 Little RC, et al. *SAS System for Mixed Models.* SAS Institute: Cary, NC, 1997.

19 Opalka D, et al. Simultaneous Quantitation of Antibodies to Neutralizing Epitopes on Virus-Like Particles for Human Papillomavirus Types 6, 11, 16 and 18 By a Multiplexed Luminex Assay. *Clin. Diagn. Lab. Immunol.* 10(1) 2003: 108–115.

20 Dias D, et al. Optimization and Validation of a Multiplexed Luminex Assay To Quantify Antibodies to Neutralizing Epitopes on Human Papillomavirus 6, 11, 16 and 18. *Clin. Diagn. Lab. Immunol.* 12(8) 2005: 959–969.

21 Miettinen O, Nurminen M. Comparative Analysis of Two Rates. *Stat. Med.* 4(2) 1985: 213–226.

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;* ***Carlos 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.*