Clinical and Vaccine Immunology

Statistical Approach To Estimate Vaccinia-Specific Neutralizing Antibody Titers Using a High-Throughput Assay

Richard Kennedy, V. Shane Pankratz, Eric Swanson, David Watson, Hana Golding and Gregory A. Poland

Published Ahead of Print 17 June 2009.

Updated information and services can be found at:
http://cvi.asm.org/content/16/8/1105

REFERENCES

These include:
This article cites 52 articles, 18 of which can be accessed free at:
http://cvi.asm.org/content/16/8/1105#ref-list-1

CONTENT ALERTS

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: http://journals.asm.org/site/misc/reprints.xhtml
To subscribe to another ASM Journal go to: http://journals.asm.org/site/subscriptions/
Statistical Approach To Estimate Vaccinia-Specific Neutralizing Antibody Titers Using a High-Throughput Assay

Richard Kennedy,¹ V. Shane Pankratz,² Eric Swanson,¹ David Watson,² Hana Golding,³ and Gregory A. Poland¹,4*

Mayo Vaccine Research Group, Mayo Clinic, Rochester, Minnesota 55905;¹ Division of Biostatistics, Mayo Clinic, Rochester, Minnesota 55905;² Division of Viral Products, CBER, Food and Drug Administration, Washington, DC;³ and Program in Translational Immunovirology and Biodefense, Mayo Clinic, Rochester, Minnesota 55905

Received 10 March 2009/Returned for modification 19 May 2009/Accepted 8 June 2009

Because of the bioterrorism threat posed by agents such as variola virus, considerable time, resources, and effort have been devoted to biodefense preparation. One avenue of this research has been the development of rapid, sensitive, high-throughput assays to validate immune responses to poxviruses. Here we describe the adaptation of a β-galactosidase reporter-based vaccinia virus neutralization assay to large-scale use in a study that included over 1,000 subjects. We also describe the statistical methods involved in analyzing the large quantity of data generated. The assay and its associated methods should prove useful tools in monitoring immune responses to next-generation smallpox vaccines, studying poxvirus immunity, and evaluating therapeutic agents such as vaccinia virus immune globulin.

Smallpox was one of the most feared diseases known, killing hundreds of millions of individuals in the last several centuries alone. Following a massive surveillance and vaccination campaign led by the WHO, the disease was eradicated in 1980 (17). Recent concerns about bioterrorism have renewed scientific interest in this disease, especially in the areas of pathogen biology, host response, vaccines, and therapeutic agents (3, 6, 21). As eradication occurred before the advent of modern immunologic techniques, the fine details regarding poxvirus immunology and vaccine efficacy are only now coming to light. Historically, humoral responses were thought to be primarily responsible for vaccine-induced protection against smallpox disease (17). Although there was some evidence for a role of cellular immunity in disease protection, modern tools to accurately measure these responses were not developed until after the eradication effort (52). Measurements of antibody (Ab) titer play a central role in most vaccine efficacy studies. Classically, plaque reduction neutralization tests (PRNT) were used to determine serum neutralizing Ab titers in smallpox vaccine recipients (15, 34, 39, 42). While effective, these assays suffer from a number of disadvantages, including the use of large volumes of clinical test material, 3 to 4 days of incubation for plaque development, manual counting of plaques, subjective readout interpretations, and high labor costs. Although a number of improvements have been made to increase the efficiency of PRNTs (5, 41), they remain extremely time and labor intensive and do not lend themselves well to large-scale applications, and comparing results directly between laboratories is difficult. Given the renewed interest in variola virus as a potential biological weapon (21) and documented monkeypox outbreaks in the United States and Africa (11, 27, 31), there is a considerable impetus to develop next-generation vaccines against smallpox and myriad other select agents and emerging pathogens. As a result, substantial effort has been expended in the development of new and improved assays to circumvent the limitations of PRNTs (4, 44, 46). Many of these assays utilize high-throughput methodologies to detect levels of viral activity in biological samples (16, 26). Several assays have been developed which rely on recombinant vaccinia viruses expressing reporter genes such as green fluorescent protein or β-galactosidase (β-Gal), with subsequent fluorometric or colorimetric detection of the reporter (12, 28, 35). The β-Gal-based assay was initially described in 2003 and is as sensitive as PRNTs but is more rapid and easily adapted to high throughput and has an objective readout conducive to large-scale statistical analyses (35). Since its initial description, it has been used for a number of studies, including a comparison of Dryvax (Wyeth Laboratories, Inc.) and modified vaccinia virus Ankara immunization in a mouse study (37), the detection of vaccinia virus-specific Ab titers in samples from human immunodeficiency virus-positive adults (29), the evaluation of different vaccination sites for human subjects (53), a nonhuman-primate study showing that Ab responses are sufficient for protection against monkeypox (13, 14), and the evaluation of a subunit recombinant vaccine for use in protection against monkeypox (22). While each of these studies illustrated the utility of the β-Gal-based assay, the number of samples tested in each study was relatively small. We have now optimized and employed this assay in a large population-based study examining immune responses in over 1,000 human subjects. The assay proved to be readily adaptable to large-scale use. However, the analytical methods for this assay suffered from the need to manually inspect the data and identify assay outliers prior to estimating the 50% infectious dose (ID₅₀). We now report the development of a robust statistical analysis plan that requires relatively little user intervention.

* Corresponding author. Mailing address: Mayo Vaccine Research Group, Mayo Clinic, Guggenheim 611C, 200 First Street SW, Rochester, MN 55905. Phone: (507) 284-4968. Fax: (507) 266-4716. E-mail: Poland.gregory@mayo.edu.

1 Corresponding author. Mailing address: Mayo Vaccine Research Group, Mayo Clinic, Guggenheim 611C, 200 First Street SW, Rochester, MN 55905. Phone: (507) 284-4968. Fax: (507) 266-4716. E-mail: Poland.gregory@mayo.edu.

7 Published ahead of print on 17 June 2009.
MATERIALS AND METHODS

Subject recruitment. Eligible subjects were healthy adults between 18 and 40 years of age who had received a single dose of Dryvax within 4 years of recruitment and at least 1 month before the blood draw. Olmsted county residents who were vaccinated as part of the Department of Health and Human Services civilian health care worker program were recruited by the Mayo Clinic (49), while eligible armed forces personnel were recruited by the Naval Health Research Center in San Diego, CA. Institutional review board approval for all study procedures was granted, and written, informed consent was obtained from each subject. Serum samples were obtained, aliquoted into sterile microcentrifuge tubes, and stored at -70°C until use. Prior to the first use, each serum sample was heat inactivated at 56°C for 30 min and then kept at 4°C until the assays for that individual were completed.

Viruses and cell lines. HeLa, HeLa S3, and Vero cells were all obtained from the ATCC and grown in Dulbecco's modified Eagle's medium (Vero) or Eagle's minimal essential medium (HeLa, HeLa S3) supplemented with 10% fetal bovine serum, 2 mmol/liter L-glutamine, 0.1 mmol of nonessential amino acids, 1.0 mmol/liter sodium pyruvate, 1.5 g/liter sodium bicarbonate, and penicillin and streptomycin at 500 IU/ml and 500 µg/ml, respectively. All cell lines were used as originally described (35). A single lot of fetal bovine serum was used for all assays. This serum, while not directly screened for neutralizing activity against vaccinia virus, was included in both the positive and negative controls and any background neutralizing activity would be adjusted for during the normalization procedures for this assay. It is possible that lot-specific differences in background vSC56 activity contribute to the small differences seen between laboratories with respect to optimal viral activity.

Vaccinia virus strain vSC56, expressing β-Gal under the control of a synthetic early/late promoter, was kindly provided by B. Moss (NIAID). The virus was grown in HeLa S3 cells and subjected to titration in Vero cells. The virus stock used in the assay was partially purified by ultracentrifugation (15,800 rpm for 80 min at 4°C in a SW28 rotor) over a 36% sucrose cushion, resuspended in 1 mM Tris-HCl (pH 9.0), aliquoted, and stored in single-use vials at −70°C until use.

Reagents. Vaccinia virus immune globulin (VIG) (MBL/Dynport VIG intra- venous, lot 1) was kindly provided by Christine Anderson (CBER/FDA). VIG was diluted 1:10 in phosphate-buffered saline–0.1% bovine serum albumin, aliquoted, and stored at −70°C until use. CPRG (chlorophenol red-β-d-galacto-pyranoside) substrate was obtained from Calbiochem (San Diego, CA), diluted in phosphate buffer to 4 mg/ml, and stored at −20°C until use. Recombinant β-Gal enzyme was purchased from Roche (Indianapolis, IN). Igepal CA630 and the other chemicals necessary for the various buffers were obtained from Sigma (St. Louis, MO). The cell lysis buffer containing cell medium and Igepal CA630 was made fresh for each assay, and all other buffers were made in bulk and stored appropriately until needed.

Neutralization assay. The assay procedure was performed as initially reported (35) with several modifications as described in this report. Briefly, serum samples were thawed, heat inactivated, briefly spun at 10,000 rpm in a table-top microcentrifuge, and serially diluted. These dilutions (five replicate samples per dilution) were incubated with a known quantity of vSC56 for 1 h at 37°C and added to wells containing 50,000 Vero cells. Positive and negative controls consisted of serially diluted VIG (five replicates per dilution) and virus-only samples (eight replicates per 96-well plate), respectively. One set of VIG dilutions was performed for each assay, while eight virus-only samples were included on each plate in every assay. The plates containing Ab, virus, and cells were incubated overnight and the cells lysed with detergent (Igepal CA630). We did not perform the freeze-thaw step outlined in the original protocol. An aliquot of each lysate was then incubated with a colorimetric substrate (CPRG) in an appropriate buffer system. The enzymatic reaction was quenched using Na2CO3, and the plates were read using a ThermoMax plate reader (Molecular Devices, Sunnyvale, CA). Data were collected and transferred to a Microsoft Excel template for standard curve calibration and initial analysis.

Statistical analysis. Each standard curve was assessed, verified to have an R² ≥ 0.99, and used to convert the values of the optical density readings on that plate to the levels of β-Gal activity. The levels of β-Gal activity in the virus-only wells, as well as the ID₅₀ of the VIG standard, were then estimated and verified to be within appropriate ranges (virus-only β-Gal activity, 1,000 to 6,000; VIG ID₅₀, 5 µg/ml to 30 µg/ml). All well-specific results were retained for analysis. Assays where both the virus-only and VIG averages fell outside of these ranges were discarded, and the subject's samples were tested again.

The basis of the innovation in the estimation of ID₅₀ values was the use of robust regression techniques. A bottleneck in the use of the assay in a high-throughput situation has been the need to inspect the data, identify assay outliers, and make a determination to accept or reject each outlier prior to the analysis. Our intent was to remove the need to identify and eliminate aberrant measures of β-Gal activity required by least-squares approaches to curve fitting. To accomplish this, we employed the M estimation approach introduced by Huber (25). This approach is a generalization of standard least-squares estimations; while the least-squares approach minimizes the sum of squared residuals, M estimators minimize the sum of the functions of the residuals that are increasing less rapidly, thus reducing the effect of reducing the influence of outliers on the estimates of the regression coefficients. We relied on the iteratively reweighted least-squares approach in performing M estimations, with a bisquare weighting function, that is implemented in the ROBUSTREG procedure of the SAS software package (Cary, NC). Raw optical density readings were converted to units of enzyme activity based on standard curves for each plate and then scaled to percentages of control activity. These percentages of control values were then log transformed and all analyses carried out using SAS software. We developed a web-based tool that estimated the ID₅₀ value for each individual whose samples had been assayed on each plate. Together with the assay results, this tool returned a series of model fitting and other assessments that allowed the evaluation of data quality.

We fit a quadratic trend to the data to accommodate dilution readings which fell outside of the linear range of the assay. The resulting regression coefficients from the M estimation procedure were extracted and used with reference to the virus-only levels to estimate the ID₅₀ for each individual. Our web tool returned this estimate along with a collection of assessments of data quality for each assay, including (i) a plot of the observed data, with the fitted line superimposed, as a visual check of the agreement between the model and the data; (ii) the model deviance, a measurement for which large values are indicative of a poor agreement between the observed data and the fitted model; and (iii) indicators of whether the ID₅₀ values were markedly greater than 100% of the percentage of control values. Once all replicated ID₅₀ estimates were obtained in this way, a linear mixed model, with a per-subject random effect, was applied to the replicated log (ID₅₀) values on the logarithmic scale. This enabled additional assessments of the overall performance of the assay and provided an estimate of the intraassay correlation coefficient, which measures assay repeatability across all subjects, as well as measurements of the residuals, which assess the degree to which the ID₅₀ values deviated from their within-individual averages. Experiments with values more than 3.0 standard deviations from the per-subject means estimated from the linear mixed-model fit to all the data were repeated.

ASSAY OPTIMIZATION

Objective. Our study population consisted of 1,076 healthy adults ranging from 18 to 40 years of age. Study subjects were military personnel or civilians involved in the U.S. Smallpox Preparedness and Response Program with documented “takes” (i.e., development of the characteristic pox at the vaccination site) after receipt of a single dose of smallpox vaccine (45, 49). Table 1 lists additional demographic information. Ab responses peak 4 weeks after receipt of the smallpox vaccine (1, 17, 32); consequently, serum samples from each subject were obtained at least 6 weeks after smallpox vaccination, allowing sufficient time for the development of humoral responses. Each sample was tested in at least three separate experiments run on different days. Our protocol had several notable changes from the originally described procedure. These changes, briefly mentioned above in Materials and Methods, are reiterated here along with their effect on assay utility. (i) The addition of a partial purification step performed by ultracentrifugation of the infected HeLa lysate over a 36% sucrose cushion followed by sonication greatly increased the consistency of viral infection and the enzyme activity levels in our assays. (ii) Heat-inactivated serum samples were spun briefly at 10,000 rpm in a table-top microcentrifuge before use to remove Ab complexes and reduce the variability seen in our assays. (iii) The freeze-thaw step before the development phase of the assay was omitted. Given the volume of samples assayed at one time, this step was not necessary. Additionally, the throughput situation has been the need to inspect the data, identify assay outliers, and make a determination to accept or reject each outlier prior to the
Neutralizing Ab titers in study population. Figure 1 shows the distribution of neutralizing Ab responses found in our study population, as measured by at least three separate experiments (each performed after optimization and assay modification). The median values of the individual assay ID_{50} measurements are plotted. These ID_{50} values ranged from 15.7 to 3,612.54. Our range of VIG ID_{50} values was slightly lower than that originally published, with a median of 13.8/5.2 to 25.9). Given that the acceptable ranges for the MOI and virus activity in the assay differed between our laboratory and that of the FDA, these parameters may differ from laboratory to laboratory.

Initial testing of the robust modeling approach. In order to assess the quality of the estimates provided by our high-throughput, robust modeling approach, we identified a sample of 30 subjects whose data were run in a short time interval and assessed the agreement between the ID_{50} estimates calculated using both the original and the modified analytical methods. Using all data from the 30 assays resulted in a correlation of 0.94 between the log-transformed estimates from the two methods. However, differences between these estimates ranged from −0.199 to 0.153, with a mean ± standard deviation of −0.005 ± 0.080 (See Fig. 2). Of these assays, a total of 17 (57%) had outliers; as recommended by the original method, these were removed. Following the re-estimation of the ID_{50} values, the correlation value climbed to 0.97 and the differences between the results obtained for the original and the robust methods ranged from −0.107 to 0.091, with a mean ± standard deviation of −0.007 ± 0.061, on the logarithmic scale. The ID_{50} values from our modified approach were on average within 3% of the ID_{50} values obtained by the original approach without outliers.

The approach to controlling the quality of the individual-specific assay readings described by Manischewitz et al. requires hands-on monitoring of each assay and outlines criteria for the removal of outlying observations (35). In our study, a single operator performed the assay on 38 samples at once. The hands-on time was ~6 h of effort, after which the operator was required to monitor each of approximately 1,000 data points to determine whether there were outliers in any set of replicates and which outliers were to be removed. This step takes additional time and effort. The approach described here employed statistical methods that are robust (i.e., that result in the presence of only a small number of outliers) and eliminate the time-consuming process of defining and then finding and removing outliers. In Fig. 3, one can see the success of this approach in estimating a value of the ID_{50} in the presence of suboptimal data (i.e., data containing outliers).

For the bulk of the subjects (94.1% [1,013/1,076]), all of the replicated ID_{50} values fell within the 1:20 to 1:540 range. Assays where a subject sample had an ID_{50} value greater than 540 were repeated using an extended dilution range (1:60 to 1:1,620). After these assays were performed, using the extended range and fitting the linear mixed model, the intraclass correlation was found to be 0.70. While this value suggests a high degree of reproducibility of the results, 0.95% (31/3,260) of the assays had residuals that were more than 3.0 standard deviations from zero). Each of these 31 subjects was retested an additional three times. All of these second assays confirmed to our quality control constraints, and a repeated residual analysis verified that all samples demonstrated acceptable degrees of variation among the replicate experiments. The value for intraclass correlation for the revised data set was 0.74.

The coefficient of variation (CV) was obtained for each subject in this final data set of log-transformed ID_{50} values. The average CV was 6.9%, with a standard deviation of 4.1%. The largest CV was 26.1%, and the 25th and 75th percentiles of the CV distribution were 3.7% and 9.2%, respectively. The CV for more than 99% of the subjects was less than 20%.

Additional variables which affect data analysis. As we found it necessary to establish a different range of VIG and virus-only values from those originally reported, we examined the effects
of several different parameters on ID\textsubscript{50} results. These parameters included virus activity, VIG ID\textsubscript{50}, date of assay, and interpretation of data by operator. With respect to virus activity, we found that higher levels of virus were associated with lower ID\textsubscript{50} measurements (\(P < 0.0001\)) and that doubling the level of virus activity resulted in a decrease in the average ID\textsubscript{50} measurements by a factor of about 0.88, a 12% decrease. Given that sufficiently high levels of virus are likely to overcome the neutralizing activity of most serum samples, this correlation is not surprising. Likewise, with VIG activity, there was a significant association with ID\textsubscript{50} values (\(P < 0.0001\)), and a doubling of the level of VIG activity was associated with a 7% increase in the average ID\textsubscript{50} values. We did not expect VIG ID\textsubscript{50} values to be directly related to those of serum ID\textsubscript{50}; however, the VIG value is an indirect measure of assay performance, and differences in VIG activity may reflect systematic differences in serum neutralization. Finally, we analyzed VIG ID\textsubscript{50} values for evidence of assay drift. While VIG ID\textsubscript{50} values remained within a narrow range, there was no obvious trend in the observed values (See Fig. 4). Interestingly, while the amount of variability of our VIG ID\textsubscript{50} values was nearly identical to that originally described, the actual range was somewhat lower and may reflect laboratory-specific differences in reagent usage and/or technique. The average month-to-month difference in VIG ID\textsubscript{50} values was 3.8%. We observed no significant association between operator identity and ID\textsubscript{50} estimates (\(P = 0.6184\)). Determinations of average ID\textsubscript{50} values differed among operators by less than 3%.

**DISCUSSION**

With any assay or technique, scaling up of sample sizes can pose multiple problems and necessitate numerous procedural changes and refinements. Fortunately, the original protocol
was highly amenable to being scaled up. Most of the assay materials can be manipulated in a 96-well format with multichannel pipettes. One of the largest bottlenecks in the procedure was the transfer of serum from individual microcentrifuge storage tubes to the 96-well plate for dilution. Initial serum storage in 96-well plates would result in considerable time savings. Without any automation, we were able to readily test up to 20 96-well plates per operator per assay while remaining within the time constraints of the various protocol steps. The use of robotic workstations would dramatically increase this number. Another of the larger bottlenecks was the time and effort required for growing, maintaining, and expanding the HeLa cell cultures (we required \( \approx 4 \times 10^7 \) cells each week). Laboratories utilizing this assay for a large numbers of samples are encouraged to invest in appropriate bulk cell culture systems.

We also found that acceptable levels of virus activity, MOI, and duration of infection may differ between laboratories. Care should be taken to optimize these variables within each laboratory before large-scale testing is performed. In spite of these differences, the inclusion of the VIG standards should still allow results to be compared between laboratories. The development of comprehensive sets of assay standards would allow direct comparisons between laboratories and serve to address this concern.

Our results showed that the range of ID\(_{50}\) values was quite wide (15.7 to 1,314.2), with 100% of our subjects demonstrating detectable levels of humoral immunity higher than those found with a small panel of vaccinia virus-naïve subjects (average ID\(_{50}\) = 1.11). Our study inclusion criteria included a documented “take.” Historically, this take has been considered evidence of successful vaccination and protection from disease (17). Recent studies have confirmed that the vaccine take is associated with seroconversion and positive vaccinia virus-specific T- and B-cell responses (2, 8, 36), although in the absence of endemic disease, we can only estimate correlates of protection. Historically, titers of 1:20 or 1:32 were assumed to rep-
ID50 values well above 540, it was necessary to use a higher dilution range of the assay. For those individuals with documented vaccine “take,” a small number of reports cite 50% neutralization titers of between 1:10 and 1:100 (10, 20), 64 and 256 (19), and 64 and 128 (50), indicating that the measurement error of the assay is quite high, indicating that the measurement error of the assay is quite small relative to the person-to-person differences.

We acknowledge Megan Ryan and Kevin Russell and their Naval Health Research Center team as well as the nurses and study coordinators from the Mayo Vaccine Research Group for their tremendous support from the Mayo Vaccine Research Group for their tremendous efforts.
efforts in recruiting the large number of subjects needed for the study. We thank Cheri A. Hart for her editorial assistance.

This work was supported by the NIAID though NIH Population Genetics Analysis Program grant AI40065.

Disclosures and conflicts of interest: G.A.P. has provided consultant advice to Acambis. The other authors do not have any conflicts of interest.

REFERENCES


