

## Interlaboratory Comparison of Three Multiplexed Bead-Based Immunoassays for Measuring Serum Antibodies to Pneumococcal Polysaccharides<sup>∇</sup>

Melissa J. Whaley,<sup>1</sup> Charles Rose,<sup>1</sup> Joseph Martinez,<sup>1</sup> Gouri Laher,<sup>2</sup> Deborah L. Sammons,<sup>3</sup> Jerry P. Smith,<sup>3</sup> John E. Snawder,<sup>3</sup> Ray Borrow,<sup>2</sup> Raymond E. Biagini,<sup>3</sup> Brian Plikaytis,<sup>1</sup> George M. Carlone,<sup>1</sup> and Sandra Romero-Steiner<sup>1\*</sup>

*Meningitis and Vaccine Preventable Diseases Branch, Division of Bacterial Diseases, National Center of Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia 30333<sup>1</sup>; Health Protection Agency, Vaccine Evaluation Unit, Manchester Medical Microbiology Partnership, Manchester Royal Infirmary, Manchester M13 9WZ, United Kingdom<sup>2</sup>; and Biomonitoring and Health Assessment Branch, Division of Applied Research and Technology, National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Cincinnati, Ohio 45226<sup>3</sup>*

Received 15 January 2010/Returned for modification 11 February 2010/Accepted 15 March 2010

**Serotype-specific IgG, as quantified by a standardized WHO enzyme-linked immunosorbent assay (ELISA), is a serologic end point used to evaluate pneumococcal polysaccharide-based vaccine immunogenicity. Antibodies to each vaccine polysaccharide in licensed multivalent vaccines are quantified separately; this is laborious and consumes serum. We compared three bead-based immunoassays: a commercial assay (xMAP Pneumo14; Luminex) and two in-house assays (of the Health Protection Agency [HPA] and Centers for Disease Control and Prevention [CDC]), using the WHO-recommended standard reference and reference sera ( $n = 11$ ) from vaccinated adults. Multiple comparisons of the IgG concentrations for seven conjugate vaccine serotypes were performed by sample (percent error), serotype (equivalency testing), and laboratory (concordance correlation coefficient [CCC]). When comparing concentrations by sample, bead-based immunoassays generally yielded higher antibody concentrations than the ELISA and had higher variability for serotypes 6B, 18C, and 23F. None of the three assays met the current WHO recommendation of 75% of sera falling within 40% of the assigned antibody concentrations for all seven serotypes. When compared by serotype, the CDC and HPA tests were equivalent for five of seven serotypes, whereas the Luminex assay was equivalent for four of seven serotypes. When overall mean IgG concentrations were compared by laboratory, a higher level of agreement (CCC close to 1) was found among bead-based immunoassays than between the assays and WHO assignments. When compared to WHO assignments, the HPA assay outperformed the other assays ( $r = 0.920$ ; CCC = 0.894; coefficient of accuracy = 0.972). Additional testing with sera from immunogenicity studies should demonstrate the applicability of this methodology for vaccine evaluation.**

*Streptococcus pneumoniae* (pneumococcus, or Pnc) has over 90 serotypes based on its capsular polysaccharide (Ps). Following introduction of the 7-valent polysaccharide-protein conjugate vaccine (PCV-7) in the United States in 2000, the incidence of invasive pneumococcal disease (IPD) due to vaccine serotypes declined (18). IPD due to nonvaccine serotypes has increased in some countries, making expanded-valence vaccines important. At present, serotype-specific IgG as quantified by enzyme-linked immunosorbent assay (ELISA) is the major serologic end point used to evaluate the immunogenicity of Pnc polysaccharide-based vaccines. A consensus Pnc ELISA protocol (3; www.vaccine.uab.edu) was generated after two multilaboratory assay comparisons for IgG antibodies (12). In addition to the protocol, a set of 12 reference sera with serotype-specific assignments is available from the World Health Organization (WHO). A reference standard serum (89SF) is

available from the Food and Drug Administration (FDA; MD) (13). The purpose of these reference materials is to assist in establishing the protocol in laboratories worldwide as well as in the evaluation and implementation of new technologies. Laboratories have developed multiplex technologies to meet the increasing demands of multivalent vaccines (4, 8, 9). These multiplex technologies greatly reduce material waste and amount of serum sample, reagents, and operator time, whereas the Pnc Ps single-plex ELISA requires individual serotype-specific assays to detect and quantify antibody to each Ps constituent in the vaccine.

A multiplex bead-based immunoassay was first described for the measurement of antibodies to Pnc Ps antigens by Pickering et al. in 2002 (11). Those authors introduced the Luminex (Austin, TX) flow cytometric system, which utilized two lasers in the detection of serum IgG antibodies to 14 different Pnc Ps (serotypes 1, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 12F, 14, 18C, 19F, and 23F) within a single reaction well. This methodology was further expanded upon to include additional polysaccharides and by other methods of Ps-bead conjugations. Biagini et al. described the use of sodium periodate to oxidize the Ps covalent link to each of 23 Pnc Ps to amino groups on the beads (1). The

\* Corresponding author. Mailing address: Vaccinology Laboratories, MVPD, DBD, Centers for Disease Control and Prevention, 1600 Clifton Road, Building 18, Room B-105, MS A-36, Atlanta, GA 30333. Phone: (404) 639-2473. Fax: (404) 639-4518. E-mail: Ssteiner@cdc.gov.

<sup>∇</sup> Published ahead of print on 24 March 2010.

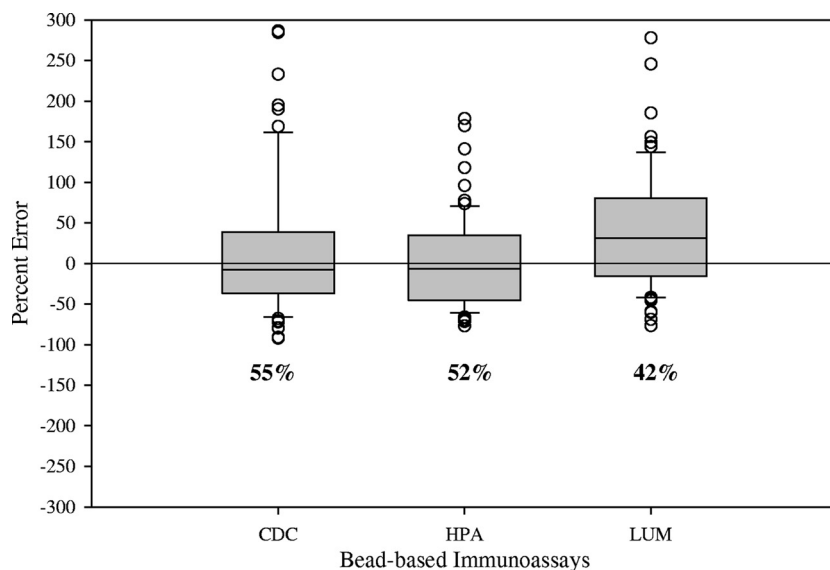


FIG. 1. Comparison of IgG concentrations using the percent error from the mean to calculate agreement between bead-based immunoassays and WHO reference assignments for PCV-7 serotypes. Box plots give the mean concentration and the 25% and 75% quartiles. Concentrations outside of the 95% confidence intervals are represented by open circles. Percentages below the box plots are the levels of agreement.

assay described by Lal et al. as a nonplex assay has been validated successfully at the Health Protection Agency (HPA) in the United Kingdom (4). This HPA assay uses a modification of the poly-L-lysine conjugation technique described by Pickering et al. in 2002. Schlottmann et al. described a modified assay that uses Pnc Ps conjugation via the carboxyl functional groups in the microspheres and 4-(4,6-dimethoxy[1,3,5]triazin-2-yl)-4-methylmorpholinium (DMTMM) (14).

In this study, we compared three different Pnc Ps bead-based immunoassays, one commercial and two in-house assays, which were evaluated in separate laboratories using the WHO reference sera to determine how these methodologies agree with each other and with the WHO reference assignments. To our knowledge this is the first interlaboratory comparison of Pnc Ps bead-based immunoassays, although the basic methodologies for coating beads with Pnc Ps have been compared by Scholtmann et al. (14).

(This research was presented in part as a poster at the 6th International Symposium on Pneumococci and Pneumococcal Diseases, 8 to 12 June 2008, Reykjavik, Iceland [17].)

#### MATERIALS AND METHODS

**Serum samples.** The serum samples ( $n = 11$  or  $12$ ) used were the reference sera available at the National Institute for Biological Standards and Control (NIBSC), HPA, Hertfordshire, United Kingdom, for the evaluation of pneumococcal assays. Due to low quantities of certain sera, 11 of the 12 samples were evaluated in some laboratories. These serum samples have reference assignments that were established by a WHO working group and are readily available ([www.vaccine.uab.edu](http://www.vaccine.uab.edu)). These sera were from young adults (mean age, 39 years) vaccinated with a single dose of Pneumovax II (Merck Sharp and Dohme, Ltd.). Serum samples were lyophilized and stored at  $-20^{\circ}\text{C}$  until resuspension with double-distilled  $\text{H}_2\text{O}$ . Resuspended samples were aliquoted and stored at  $-70^{\circ}\text{C}$ . All laboratories, following an established consensus ELISA protocol (13; [www.vaccine.uab.edu](http://www.vaccine.uab.edu)), used the same standard reference serum (89SF) with known antibody concentrations for 23 pneumococcal capsular polysaccharides. Each laboratory preadsorbed the reference sera with cell wall polysaccharide (CPS) and pneumococcal serotype 22F polysaccharide to adsorb cross-reacting antibodies (2).

**Bead-based immunoassays.** We compared three different bead-based immunoassays for the quantitation of IgG antibodies to pneumococcal capsular polysaccharide in multiplex formats: one was a commercially available product (xMAP Pneumo14, pneumococcal immunity panel; Luminex), and two were in-house immunoassays from HPA and the CDC. The Luminex assay has been developed and optimized for diagnostic use but has not been validated for vaccine evaluation. The in-house HPA assay has been validated at the HPA. The in-house CDC assay has been standardized and only partially validated at the CDC for vaccine evaluation.

**CDC assay.** An in-house bead-based immunoassay for 22 Pnc Ps serotypes (1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12, 14, 15B, 17F, 18C, 19A, 19F, 20, 23F, and 33F) was performed using Ps-coated beads purchased from Flow Applications, Okawville, IL. Bulk uncoated beads were purchased from Luminex, Austin, TX. Each batch of beads had a unique fluorescent signal and was coated with a different polysaccharide using the method described by Biagini et al. (1) with modifications indicated by the manufacturer. Briefly, this method uses sodium periodate to covalently link the polysaccharide to the amino residues on the surface of the beads. The CDC assay was also performed according to the protocol described by Biagini et al. (1) with the following modifications: longer incubation of serum samples with target bead mixture (1 h instead of 30 min) and goat anti-human IgG Fab fragments conjugated to R-phycoerythrin (Jackson ImmunoResearch). For simplicity we will refer to this assay as the CDC assay. This assay was evaluated at CDC, Atlanta, GA.

**HPA assay.** The HPA (Manchester, United Kingdom) in-house assay was performed according to the protocol described by Lal et al. (4) for a total of 12 Pnc Ps serotypes. This assay used beads coated with type-specific polysaccharides for serotypes 1, 3, 4, 5, 6B, 7F, 9V, 14, 18C, 19A, 19F, and 23F and a modification of the method described by Pickering et al. (11). This method uses poly-L-lysine and cyanuric chloride for conjugation of the polysaccharide to the beads.

**Luminex assay.** XMAP Pnc immunity panels were purchased by the Biomonitoring and Health Assessment Branch, Division of Applied Research and Technology, NIOSH, CDC, from Luminex for the evaluation of this technology for a total of 14 Pnc Ps serotypes (1, 3, 4, 6B, 7F, 8, 9N, 9V, 12F, 14, 18C, 19A, 19F, and 23F). The Luminex xMAP assay was followed as per the manufacturer's protocol. For simplicity we will refer to this bead-based immunoassay as the Luminex assay (LUM), even though this panel was evaluated at CDC, Cincinnati, OH.

**Data analysis.** We used three different methods for data analysis. The first method was the calculation of the percent error to determine the agreement between samples for the bead-based immunoassays and the WHO-assigned IgG concentrations (concentration), also known as the WHO assignments, for seven vaccine serotypes (<http://www.vaccine.uab.edu>). The following formula was used:

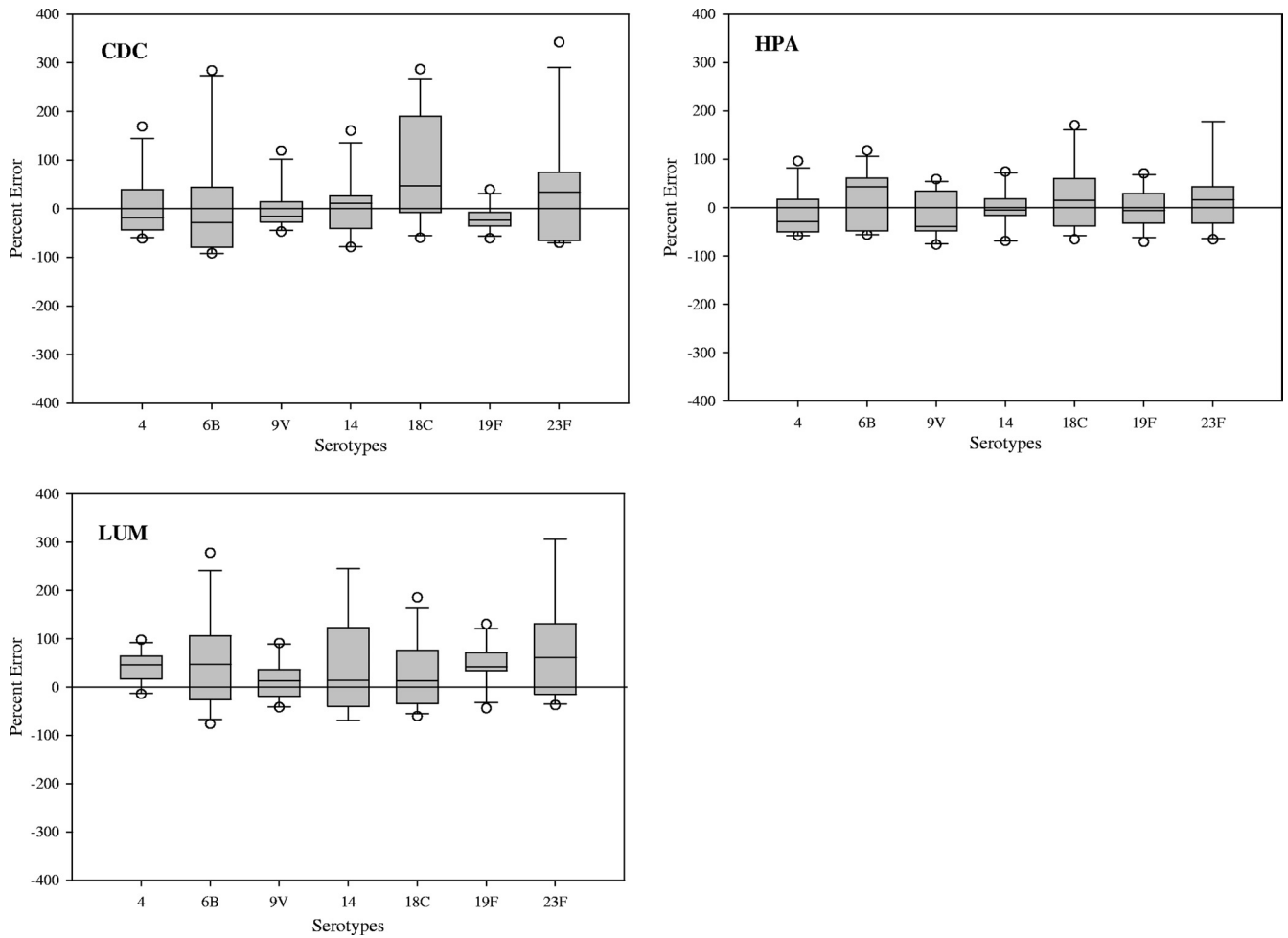


FIG. 2. Box plots illustrating the comparisons of IgG concentrations using the percent error from the mean to calculate agreement between the indicated bead-based immunoassays and WHO reference assignments per serotype.

percent error =  $100 \times \frac{|(\text{assigned concentration}) - (\text{lab-determined concentration})|}{(\text{assigned concentration})}$ .

We followed the WHO recommendation of having the IgG concentrations for  $\geq 75\%$  of the serum samples within a 40% error of the assigned mean for qualifying a new quantitative method (12; <http://www.vaccine.uab.edu>). With this method, we compared the assignments to the newly calculated concentrations for each individual reference serum.

The second method for assessing overall agreement was to compare concentrations between laboratories, with all sera and serotypes combined, which involved the following three measurements: the coefficient of accuracy ( $C_a$ ), precision ( $r$ ), and the concordance correlation coefficient (CCC). The accuracy describes the amount of compliance among laboratories and between the laboratories and WHO assignments for the IgG concentrations (6). Hence,  $C_a$  measures how far the best-fit line deviates from a 45° line. Precision (reproducibility) is the degree to which IgG concentrations determined by laboratories showed the same or similar results and was measured here by using Pearson's correlation coefficient ( $r$ ) (10). The CCC was used to evaluate the agreement between pairs of laboratories, taking into account both precision and accuracy (5).

Our third method used an analysis of variance (ANOVA) random effects model to investigate the sources of variability and test for equivalence between serotypes for each bead-based immunoassay and the WHO assignments and among all bead-based immunoassays. Random effects ANOVA models were fit separately to test for equivalence between each bead-based immunoassay and the WHO assignments and for equivalence among the bead-based immunoassays. Models were estimated using restricted maximum likelihood estimation (REML) in the SAS PROC MIXED model (7). Models consisted of a natural log transformation for the IgG concentrations and the following independent explanatory variables: overall mean, serotype, laboratory, and the interaction

among the serotypes and laboratories. After investigation it was determined that there was substantial variability due to samples and serotype by sample. Hence, the samples and interaction of serotypes by samples were treated as random effects in the models. Our first ANOVA model compared the CDC, LUM, and

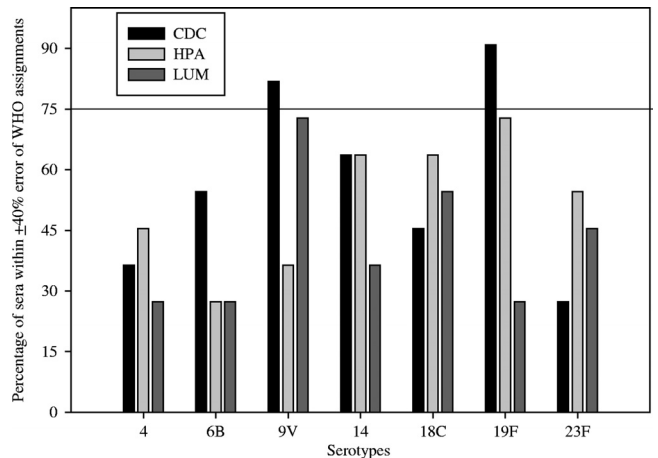


FIG. 3. Percentages of sera that were within a 40% error of the WHO assignments, graphed according to serotype for each assay.

TABLE 1. Measures of agreement among bead-based immunoassays and the WHO assignments

Assay	Statistic	WHO	CDC	HPA	LUM
WHO	Accuracy ( $C_a$ )	1.0	0.945	0.972	0.976
	Precision ( $r$ )	1.0	0.883	0.920	0.867
	CCC ( $r_c$ )	1.0	0.835 (0.767, 0.885) <sup>a</sup>	0.895 (0.847, 0.928)	0.846 (0.769, 0.898)
CDC	Accuracy ( $C_a$ )	1.0	1.0	0.988	0.926
	Precision ( $r$ )	1.0	1.0	0.921	0.891
	CCC ( $r_c$ )	1.0	1.0	0.910 (0.875, 0.936)	0.825 (0.769, 0.868)
HPA	Accuracy ( $C_a$ )	1.0	1.0	1.0	0.896
	Precision ( $r$ )	1.0	1.0	1.0	0.912
	CCC ( $r_c$ )	1.0	1.0	1.0	0.817 (0.759, 0.862)
LUM	Accuracy ( $C_a$ )	1.0	1.0	1.0	1.0
	Precision ( $r$ )	1.0	1.0	1.0	1.0
	CCC ( $r_c$ )	1.0	1.0	1.0	1.0

<sup>a</sup> Values in parentheses are the 95% CIs for the CCC.

HPA assays to the WHO assignments. This model was fit by defining the dependent variable as the natural log of the ratio of the lab value to the WHO assignment for the sample and serotype. Our second ANOVA model compared the CDC, HPA, and LUM assay concentrations to each other by fitting a model to the natural log of the assay concentrations and computing contrasts to compare each laboratory to the other laboratories by serotype. All comparisons were conducted using equivalency testing with *a priori* boundaries of 0.5 and 2.0 for the lower and upper confidence bounds, respectively.

### RESULTS

#### Agreement of sample concentrations to WHO assignments.

When comparing the IgG concentrations generated by each of the bead-based immunoassays to the WHO reference concentrations, the CDC and the HPA assays had higher agreement (55% and 52%, respectively) than the LUM assay (42%), as shown in Fig. 1. When comparing IgG concentrations of the samples by serotype (Fig. 2), all bead-based immunoassays generally resulted in higher concentrations than the WHO assignments. There was higher variability in the calculated concentrations generated by the bead-based immunoassays for

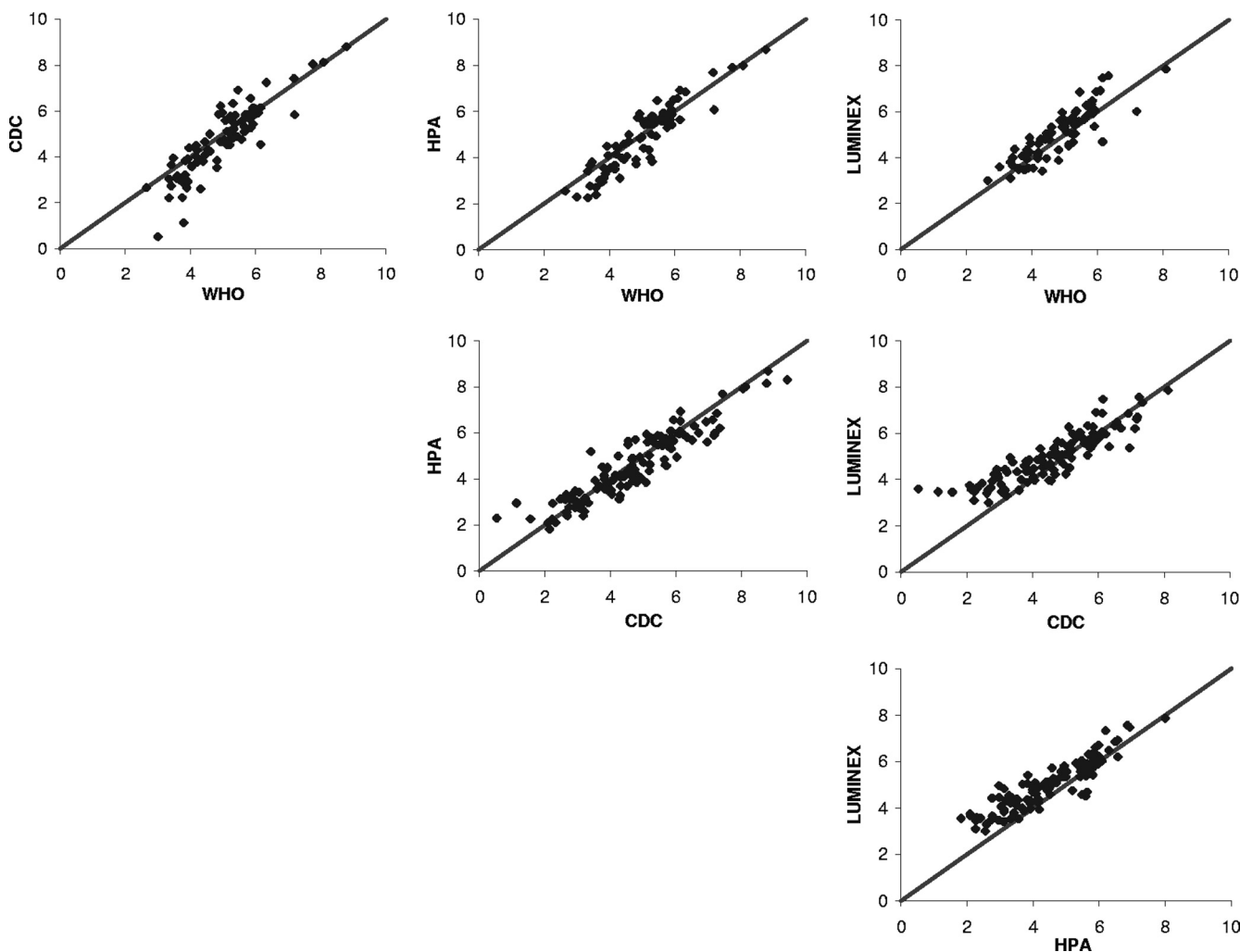


FIG. 4. Matrix of scatter plots for the pairwise comparisons of assay concentrations. IgG concentrations were plotted after log transformation and addition of 3 [ $\ln(\text{IgG concentration}) + 3$ ]. The top three plots are comparisons to WHO reference assignments and therefore include only the PCV-7 serotypes. All other plots are comparisons between bead-based immunoassays, and they include 11 serotypes common to all three assays.

serotypes 6B, 18C, and 23F among the seven serotypes analyzed. Of the bead-based immunoassays, only the CDC assay had two serotypes (9V and 19F) that met the WHO requirement of  $\geq 75\%$  of the sera within 40% percent error of the assignments (Fig. 3). For the other bead-based immunoassays, none of the serotypes was within the WHO recommendation. Since none of the bead-based immunoassays met the requirements for assay qualification based on the percent error analysis, we compared these assays using statistical methods to assess the level of agreement within the bead-based immunoassays and the WHO assignments.

**Overall agreement between laboratories and WHO assignments.** When determining the  $C_a$ , the precision ( $r$ ), and the CCC ( $r_c$ ) among the three different bead-based immunoassays and the WHO assignments (Table 1), the HPA assay had the highest agreement with the WHO assignments ( $C_a = 0.972$ ;  $r = 0.920$ ;  $r_c = 0.895$ ). All bead-based immunoassays had  $r_c$  values above 0.60, indicating a high degree of concordance (B. Plikaytis, unpublished observations) with the WHO assignments. Among the bead-based immunoassays, the HPA and CDC assays gave the highest measures of agreement ( $C_a = 0.988$ ;  $r = 0.921$ ;  $r_c = 0.910$ ). The greatest difference was seen between the LUM and CDC assays. CCC values close to 1 are desirable, as this indicates a high degree of precision and accuracy.

In the scatter plots shown in Fig. 4, we compared the bead-based immunoassays to the WHO assignments and to each other. The LUM assay had the highest accuracy, followed by the HPA and CDC assays. However, the LUM assay showed a bias toward overestimating concentrations. This can be observed in Fig. 4 by the upward shift of the data.

**Equivalency tests based on serotype.** When the bead-based immunoassays were compared to the WHO assignments, we found that both the CDC and HPA assays passed the equivalency test for five out of seven serotypes (71.4%). The LUM assay had four out of seven serotypes pass the equivalency test (57%). When we compared the bead-based immunoassays to each other for 11 serotypes common to all three assays, the same level of agreement was found between the HPA and CDC assays and between CDC and LUM assays. Six of the 11 serotypes passed the equivalency test (54.5%). The LUM and HPA assays passed the equivalency test for five serotypes (45.5%). Table 2 lists the comparisons to WHO assignments or among the bead-based immunoassays outside of the preset boundaries. It is worth noting that in this table, five comparisons were borderline (0.5 and 2.0). Comparisons of HPA concentrations to WHO assignments for serotype 4 and CDC to LUM concentrations for serotypes 4 and 14 resulted in lower confidence intervals (LCIs) between 0.45 and 0.49. The comparison of LUM concentrations to WHO assignments for serotypes 4 and 23F resulted in upper confidence intervals (UCIs) barely over 2.0.

**Variability of serum samples within serotype.** A graphic representation of the variability of the IgG concentrations for each serum sample within a serotype as determined by the bead-based immunoassays is given in Fig. 5. In this figure we noted that some serum samples exhibited a higher degree of variability ( $\geq 40\%$  error compared to WHO assignments) regardless of the bead-based immunoassay used. The presence of highly variable sera could be seen in most serotypes, except for

TABLE 2. Paired comparisons among bead-based immunoassays by serotype using ANOVA

Assays compared	Serotype	LS estimate <sup>a</sup>	LCI	UCI
CDC vs WHO	6B	-0.48	<b>0.42<sup>b</sup></b>	0.91
	18C	0.35	0.97	<b>2.09</b>
HPA vs WHO	4	-0.33	<b>0.49</b>	1.06
	9V	-0.47	<b>0.43</b>	0.92
LUM vs WHO	4	0.33	0.94	<b>2.03</b>
	19F	0.40	1.01	<b>2.21</b>
	23F	0.32	0.94	<b>2.03</b>
CDC vs HPA	1	0.62	1.31	<b>2.62</b>
	3	0.74	1.48	<b>2.97</b>
	6B	-0.50	<b>0.43</b>	0.86
	7F	0.80	1.57	<b>3.15</b>
	9V	0.36	1.02	<b>2.03</b>
CDC vs LUM	4	-0.46	<b>0.45</b>	0.89
	6B	-0.66	<b>0.36</b>	0.73
	14	-0.34	<b>0.49</b>	1.04
	19A	-0.78	<b>0.32</b>	0.67
	19F	-0.65	<b>0.36</b>	0.75
HPA vs LUM	1	-0.75	<b>0.33</b>	0.67
	3	-0.81	<b>0.31</b>	0.63
	4	-0.66	<b>0.37</b>	0.73
	7F	-1.00	<b>0.26</b>	0.52
	9V	-0.52	<b>0.42</b>	0.84
	19A	-0.59	<b>0.38</b>	0.80
	19F	-0.51	<b>0.42</b>	0.86

<sup>a</sup> The LS estimate is the difference between the least squares means of the assays compared. A negative LS estimate represents a reduction in the calculated IgG concentration of the first assay compared to the second assay.

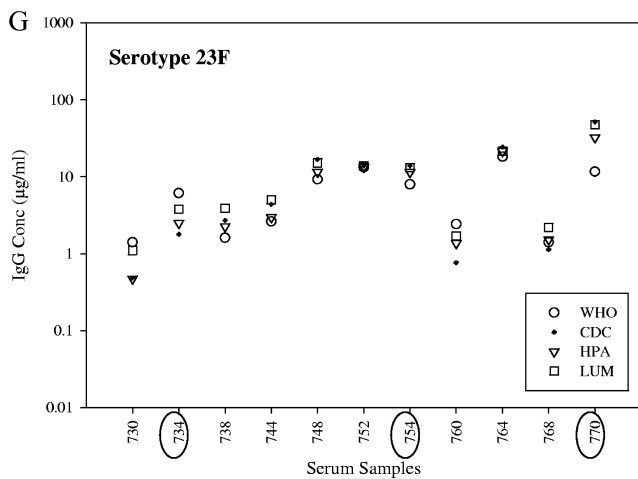
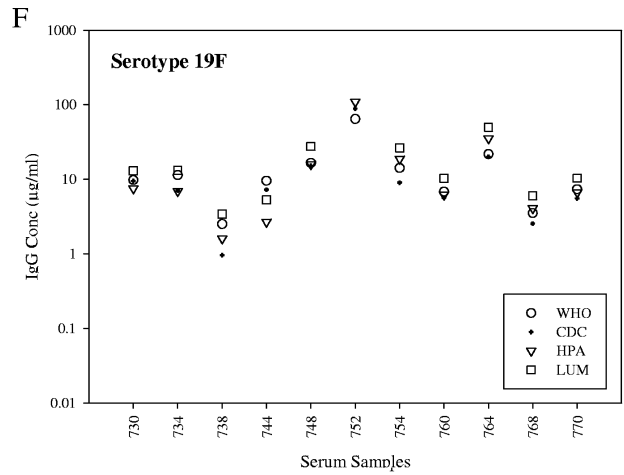
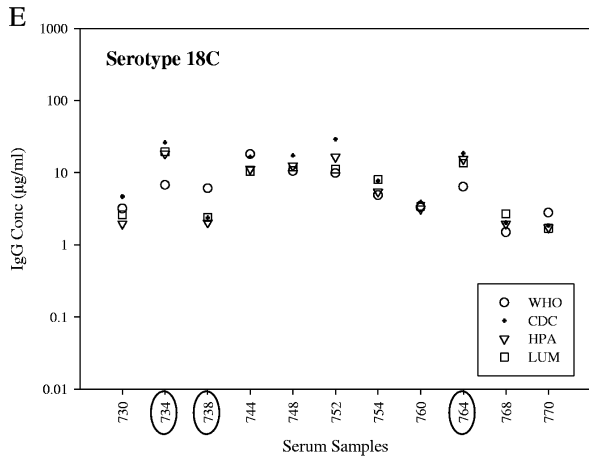
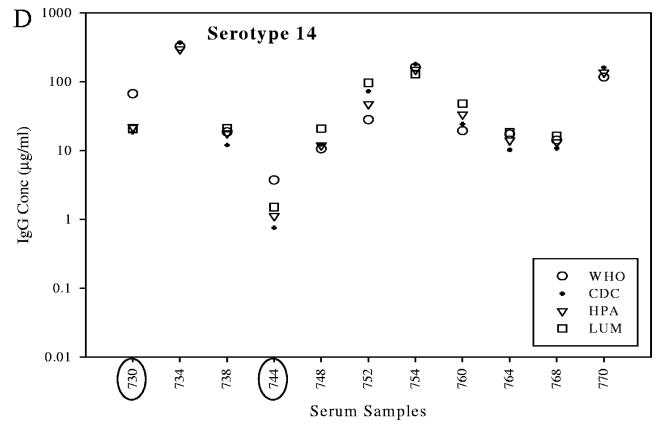
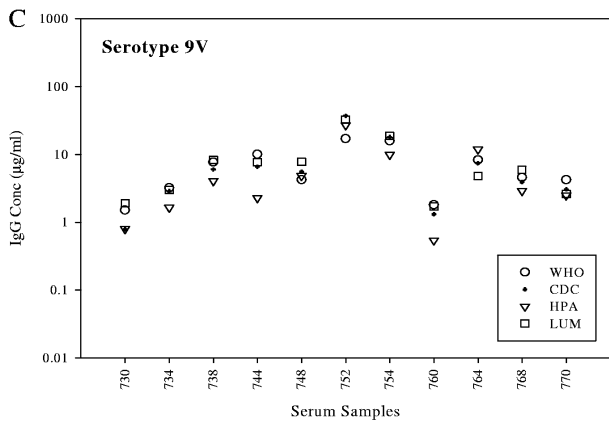
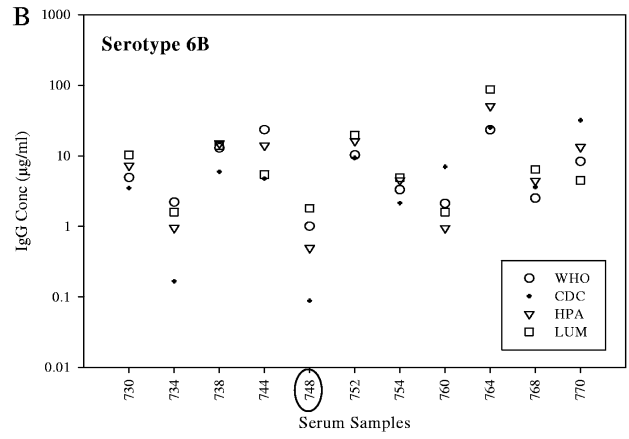
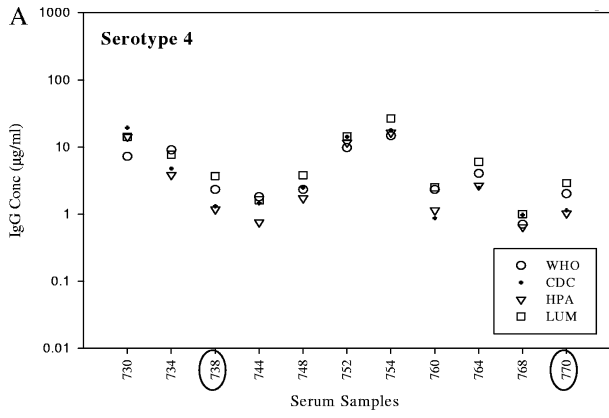
<sup>b</sup> Lower and upper confidence intervals outside of the *a priori* boundaries (0.5 and 2.0) are shown in bold.

serotypes 9V and 19F (Fig. 5). Not all sera identified as variable in the ELISA were identified as having high coefficients of variance in the bead-based immunoassays (data not shown). In the bead-based immunoassays, the estimate for the sample variability was 0.057, for sample by serotype it was 0.21, and the residual was 0.14. Hence, the sample by serotype variability accounted for almost 51% of the variability.

## DISCUSSION

Despite not being within the 40% error limit of the WHO assignments, bead-based immunoassays correlated well with the WHO assignments generated by ELISA using statistical methods. The calculation of antibody concentrations specific to pneumococcal polysaccharides is complicated due to the number of serotypes evaluated. Differences in one serotype may not be seen in another serotype when comparing methods. The presence of cross-reactive antibodies among serotypes in addition to specific antibodies in a multiplex reaction may contribute to these differences. Another factor may be the varied avidities of the antibodies to multiple antigenic targets in the reaction. Due to these factors, it is difficult to establish total agreement between concentrations derived by using ELISA and those derived by using bead-based methodologies.

Several multiplex assays have been developed for the eval-



uation of Pnc Ps antibody concentrations, with various degrees of success in terms of agreement with the consensus ELISA (1, 4, 8, 11, 14). The electrochemiluminescence-based detection assay, which utilizes the ELISA format, is the only multiplex assay currently with agreement within 40% of WHO assignments for all seven serotypes (8). Although there are some differences in bead-based immunoassay protocols, the major differences can be found in the way that the polysaccharides are covalently bound to the beads. The study of Schlottmann et al. compared five different methods for coating beads with polysaccharide and found significant differences in the mean fluorescence intensity (MFI) signals, depending on the coating method used (14). For example, the assay described by Biagini et al., which is used as the basis for the CDC assay protocol, requires great care in the oxidation step with periodate to obtain reproducible concentrations (1, 9). The CDC assay was further standardized by evaluating the type and amount of CPS adsorbents (9, 15), adsorption of the standard with either one or a combination of two CPS preparations and 22F Ps, adsorption of the samples with lysate from rough Pnc strain R36A, calibration of the number of beads per assay, data analysis formats, and intra-assay and interassay variability. None of these factors significantly changed the calculated concentrations for the 12 reference sera (data not shown).

Schlottmann et al. used Pnc Ps conjugation with DMTMM via the carboxyl rather than amino functional groups in the beads (14). The same group showed low reactivity (types 4, 5, 12F, and 19F) when they used poly-L-lysine conjugation and poor reactivity for 9 of 11 serotypes evaluated with periodate oxidation of the polysaccharides (14). However, the multiplex assay by Lal et al. used the poly-L-lysine conjugation described by Pickering and colleagues with good reproducibility, linearity, and range for calculation of IgG concentrations (4). This group successfully validated this multiplex bead-based immunoassay for the measurement of antibodies to 12 different anticapsular Pnc Ps serotypes. In this multilaboratory study, the poly-L-lysine assay had the highest level of agreement with the WHO-assigned antibody concentrations by ELISA. Previously, a nanoplex version of this assay also had high correlations with the ELISA results ( $r$  values between 0.95 and 0.98, depending on the serotype) (4). The nanoplex assay has the capacity of being combined in a 13-plex format along with a tetraplex assay for quantitation of antibodies to meningococcal polysaccharides A, C, Y, and W135 (4). This multiplex assay can reduce the amount of sera needed for the evaluation of both meningococcal and pneumococcal antibodies.

In general, the LUM assay yielded higher concentrations than the other bead-based immunoassays and the WHO assignments. Perhaps the LUM assay may need additional blocking steps to reduce the background signal due to nonspecific binding of antibody, as described by Waterboer et al. (16). In addition, the LUM assay had a limited performance range. Multiple samples were found out of range and, therefore, required repeated testing at higher dilutions (1:200, 1:400, and

1:800) rather than using the recommended 1:100 dilution. Even at higher dilutions some serum samples were still out of the range of the standard curve.

In this study, none of the three assays evaluated met the WHO recommendation of  $\geq 75\%$  of sera within 40% of the assigned values ([www.vaccine.uab.edu](http://www.vaccine.uab.edu)) for all seven serotypes. We chose to use this parameter rather than the recommendation of  $\geq 85\%$  of sera within the same percent error as described by Plikaytis et al. (12). This lower cutoff is even more applicable when we consider that three out of five laboratories submitting ELISA data to WHO did not pass the recommendation of  $\geq 85\%$  of sera within 40% of the assigned concentrations for all seven serotypes ([www.vaccine.uab.edu](http://www.vaccine.uab.edu)). These five laboratories were selected to participate in the calculation of assignments for a pneumococcal ELISA study in which sera were preadsorbed with CPS and 22F. At this time the WHO recommendations, based on these calculated assignments by ELISA, are still provisional and have provided a nonstatistical method for laboratories to compare results. The WHO recommendation for assays to be within 40% of the assigned means is a guideline for ELISAs only. There is no guarantee that it will be applicable for other types of assays. If we follow the WHO recommendations, the bead-based immunoassays cannot be used in vaccine evaluation for all seven serotypes.

Depending on the analytical approach used, the level of agreement varies between bead-based immunoassays and WHO assignments. For example, values for precision above 0.7 and for concordance correlation coefficients above 0.6 have been suggested as measures of good agreement when different methodologies are compared (B. Plikaytis, unpublished observations). If we use these suggested cutoffs for this study, all of the bead-based immunoassays have a good level of agreement with WHO assignments. When pairwise comparisons were performed by ANOVA for testing equivalency among assays on the basis of serotype, once more the HPA and CDC assays had the highest levels of agreement to WHO assignments. In general, paired comparisons were outside of the LCI for the HPA assay and outside of the UCI for the LUM assay. These findings indicate that the LUM assay had a trend toward overestimating IgG concentrations. The differences in agreement for each bead-based immunoassay to the WHO assignments were not biased to a given serotype; no serotype was commonly found outside of the preset boundaries. When bead-based immunoassays were compared with each other, 12 of the 17 comparisons resulting in a value outside of the equivalency boundaries were due to elevated LUM assay values.

A similar ANOVA was performed at the sample level (data not shown). This analysis yielded lower percentages (10% lower) in the equivalency between bead-based immunoassays and WHO assignments than the  $\pm 40\%$  error analysis. Considering the analysis methodologies followed in this study, all three bead-based immunoassays passed the overall equivalency at the laboratory level and had good agreement for most of the serotypes evaluated. The ANOVA confirmed that, overall,

FIG. 5. Graphical presentation of the variability of the mean IgG concentrations for each serum sample by serotype as determined by either bead-based immunoassay or the WHO ELISA. Samples with increased variability ( $>40\%$  error in all bead-based immunoassays) are circled on the  $x$  axis.

bead-based immunoassays correlate well with the WHO assignments but will not likely result in the same antibody concentrations.

Multiplex assays offer incredible time-saving advantages by increasing the throughput of assay determinations by a single operator and by reducing sample volume. Although these bead-based immunoassays have undergone further standardization, future validation of a consensus protocol may require the use of reference sera with reference concentrations established by using this methodology. In addition, parallel testing with appropriate target sera from immunogenicity studies should demonstrate the utility and validity of this methodology. Upon establishment of these reference concentrations, bead-based immunoassays should be applicable for vaccine evaluation.

#### ACKNOWLEDGMENTS

We thank Carl Frasch (FDA, MD) for providing the 89SF reference standard; David Goldblatt (Institute of Child Health, University of London, United Kingdom) and the National Institute for Biological Standards and Control (NIBSC), Health Protection Agency, United Kingdom, for providing the reference sera used in this study; and Daniel Schmidt (CDC) for technical support in the evaluation of the CDC in-house assay parameters.

This project was funded in part by an interagency agreement between NIOSH and NIEHS (Y1-ES-0001, Clinical Immunotoxicity).

#### REFERENCES

1. Biagini, R. E., S. A. Schlottmann, D. L. Sammons, J. P. Smith, J. C. Snawder, C. A. F. Striley, B. A. MacKenzie, and D. N. Weissman. 2003. Method for simultaneous measurement of antibodies to 23 pneumococcal capsular polysaccharides. *Clin. Diagn. Lab. Immunol.* **10**:744–750.
2. Concepcion, N. F., and C. E. Frasch. 2001. Pneumococcal type 22F polysaccharide adsorption improves the specificity of a pneumococcal-polysaccharide enzyme-linked immunosorbent assay. *Clin. Diagn. Lab. Immunol.* **8**:266–272.
3. Jódar, L., J. Butler, G. Carlone, R. Dagan, D. Goldblatt, H. Käyhty, K. Klugman, B. Plikaytis, G. Siber, R. Kohberger, I. Chang, and T. Cherian. 2003. Serological criteria for evaluation and licensure of new pneumococcal conjugate vaccine formulations for use in infants. *Vaccine* **21**:3265–3272.
4. Lal, G., P. Balmer, E. Stanford, S. Martin, R. Warrington, and R. Borrow. 2005. Development and validation of a nonplex assay for the simultaneous quantitation of antibodies to nine *Streptococcus pneumoniae* serotypes. *J. Immunol. Methods* **296**:135–147.
5. Lin, L. I. K. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **45**:255–268.
6. Lin, L., and L. D. Torbeck. 1998. Coefficient of accuracy and concordance correlation coefficient: new statistics for methods comparison. *PDA J. Pharm. Sci. Technol.* **52**:55–59.
7. Littell, R. C., G. A. Milliken, W. W. Stroup, R. D. Wolfinger, and O. Schabenberger. 2006. SAS for mixed models, 2nd ed., p. 498–499. SAS Institute Inc., Cary, NC.
8. Marchese, R. D., D. Puchalski, P. Miller, J. Antonello, O. Hammond, T. Green, L. J. Rubinstein, M. J. Caulfield, and D. Sikkema. 2009. Optimization and validation of a multiplex, electrochemiluminescence-based detection assay for the quantitation of immunoglobulin G serotype-specific antipneumococcal antibodies in human serum. *Clin. Vaccine Immunol.* **16**:387–396.
9. Martinez, J., M. Whaley, S. Romero-Steiner, and G. M. Carlone. 2008. Factors influencing the qualification of anti-pneumococcal IgG antibodies using a fluorescent bead-based system, poster P1-142. *Progr. Abstr. 6th Int. Symp. Pneumococci Pneumococcal Dis.*, 8 to 12 June 2008, Reykjavik, Iceland. Kenes International, Geneva, Switzerland.
10. Pearson, K. 1909. Determination of the coefficient of correlation. *Science* **30**:23–25.
11. Pickering, J., T. Martins, R. Greer, M. Schroder, M. Astill, C. Litwin, S. Hildreth, and H. Hill. 2002. A multiplexed fluorescent microsphere immunoassay for antibodies to pneumococcal capsular polysaccharides. *Am. J. Clin. Pathol.* **117**:589–596.
12. Plikaytis, B. D., D. Goldblatt, C. E. Frasch, C. Blondeau, M. J. Bybel, G. S. Giebink, I. Jonsdottir, H. Käyhty, H. B. Konradsen, D. V. Madore, M. H. Nahm, C. A. Schulman, P. F. Holder, T. Lezhava, C. M. Elie, and G. M. Carlone. 2000. An analytical model applied to a multicenter pneumococcal enzyme-linked immunosorbent assay study. *J. Clin. Microbiol.* **38**:2043–2050.
13. Quataert, S. A., C. S. Kirch, L. J. Wiedl, D. C. Phipps, S. Strohmeier, C. O. Cimino, J. Skuse, and D. V. Madore. 1995. Assignment of weight-based antibody units to a human antipneumococcal standard reference serum, lot 89-S. *Clin. Diagn. Lab. Immunol.* **2**:590–597.
14. Schlottmann, S. A., N. Jain, N. Chirmule, and M. T. Esser. 2006. A novel chemistry for conjugating pneumococcal polysaccharides to Luminex microspheres. *J. Immunol. Methods* **309**:75–85.
15. Slotved, H. C., C. Gluttman, C. D. Pedersen, J. N. Jacobsen, and K. A. Kroghfelt. 2009. Evaluation of the specificity of pneumococcal polysaccharide enzyme-linked immunosorbent assay and effect of serum adsorption based on standard pneumococcal serogroup- or serotype-specific rabbit antisera. *Clin. Vaccine Immunol.* **16**:1279–1284.
16. Waterboer, T., P. Sehr, and M. Pawlita. 2006. Suppression of non-specific binding in serological Luminex assays. *J. Immunol. Methods* **309**:200–204.
17. Whaley, M. J., J. Martinez, R. E. Biagini, R. Borrow, D. L. Sammons, J. P. Smith, J. E. Snawder, G. Laher, S. Romero-Steiner, and G. M. Carlone. 2008. An interlaboratory comparison of heptavalent bead-based assay for determinations of antibodies to pneumococcal polysaccharides, poster P1-141. *Progr. Abstr. 6th Int. Symp. Pneumococci Pneumococcal Dis.*, 8 to 12 June 2008, Reykjavik, Iceland. Kenes International, Geneva, Switzerland.
18. Whitney, C. G., M. M. Farley, J. Hadler, L. H. Harrison, N. M. Bennett, R. Lynfield, A. Reingold, P. R. Cieslak, T. Pilishvili, D. Jackson, R. R. Facklam, J. H. Jorgensen, A. Schuchat, and Active Bacterial Core Surveillance of the Emerging Infections Program. 2003. Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. *N. Engl. J. Med.* **348**:1737–1746.