Impact of Analytical Variability on Clinical Interpretation of Multiplex Pneumococcal Serology Assays

Xiaochun Zhang, Kelly Simmerman, Belinda Yen-Lieberman, Thomas M. Daly
Robert J. Tomsich Pathology and Laboratory Medicine Institute, Cleveland Clinic, Cleveland, Ohio, USA

The response to pneumococcal vaccination can be used to assess a patient’s humoral immune response to polysaccharide antigens. Multiplex assays measuring serotype-specific levels of pneumococcal antibodies are often used for this purpose, and clinical algorithms have been published to assist in the definition of an adequate immune response. We evaluated whether interlaboratory variability in multiplex pneumococcal serology assays would affect the clinical classification of the immune response. Specimens from 57 patients were analyzed at three reference laboratories with different multiplex assays to measure pneumococcal serology. Analytical correlation and clinical agreement in the classification of a patient’s vaccination status by the three methods were compared. Although substantial variation in the quantitative antibody levels measured by different laboratories was seen, the qualitative classification of individual serologic results showed a high degree of agreement between labs and the ultimate classification of a patient as “protected” or “nonprotected” was the same for most patients. The majority of discordant classifications were driven by a systematic bias in results from one of the assays rather than by random error. These data suggest that the use of integrated assessments based on multiple serotypes can compensate for much of the analytical variability seen between laboratories. Knowledge of the analytical performance characteristics of a particular assay is most important when evaluating patients with results near clinical cut points.

Pneumococcal serology can be used to measure the vaccination response to polysaccharide antigens in patients who are being evaluated for defects in humoral immunity. Pneumococcal vaccines such as Pneumovax 23 and Prevnar 13 contain a mixture of antigens from multiple pneumococcal serotypes to provide coverage against a broad range of commonly encountered strains (1). The earliest measurements of serotype-specific antibodies were performed with individual enzyme-linked immunosorbent assays (ELISAs) for different serotypes (2). However, the increasing availability of multiplex immunoassay techniques has led to the development of assays that can provide quantitative, serotype-specific measurements of multiple antipneumococcal antibodies in a single reaction (3, 4). Although this level of detail was initially used primarily for research into the nature of the immune response to pneumococcal vaccines (5, 6), multiplex pneumococcal serology assays are increasingly being measured in clinical practice to assess a patient’s response to pneumococcal vaccination.

Clinical interpretation of pneumococcal serology assay results can be complex, and several guidelines have been proposed to define an adequate vaccination response when using serotype-specific measurements (7–10). Although the details of these proposals vary, most approaches incorporate two primary factors, the level of the antibody response achieved against a given pneumococcal serotype and the number of serotypes for which this level is reached. A combination of these factors is then used to determine if the patient has achieved an adequate response to vaccination (10). A key parameter in such approaches is the definition of the clinically relevant cutoff that indicates an adequate antibody response. One strategy is to define a “protective” level of antibody that has been associated with reduced rates of a given clinical outcome on the basis of epidemiologic data. In adult patients, a level of 1.3 μg/ml has been proposed as a relevant cutoff on the basis of correlation with a reduced likelihood of primary infection and pneumococcal colonization in early studies (10, 11). A lower cut point of 0.35 μg/ml is often cited for pediatric patients on the basis of studies correlating this level with the prevention of invasive pneumococcal disease in infants (12, 13).

The earliest multiplex assay for measuring pneumococcal serology was developed on a bead-based flow cytometric assay platform (14), and most currently available clinical assays are based on this approach. At present, there is only one FDA-cleared in vitro diagnostic (IVD) kit for multiplex pneumococcal serology assay available in the United States. This assay measures antibodies against 14 pneumococcal serotypes (xMAP Pneumo14; Luminex Corporation, Austin, TX). In addition to the IVD assay, several commercial reference laboratories have developed laboratory-developed tests (LDTs) that are also based on the Luminex platform and use reagents and assay conditions developed and validated by the individual laboratories under Clinical Laboratory Improvement Amendments regulations. A recent interlaboratory comparison evaluated the performance of three multiplex pneumococcal assays with a small series of reference standards provided by the WHO (15). The three assays showed 42 to 55% agreement with the WHO-assigned values, with various levels of correlation between serotypes. However, in practical terms, it was not clear whether the level of variation seen was large enough to have an impact on the clinical use of these results when evaluating vaccine responses. To address this issue, we performed multiplex pneumococcal serology testing of a large series of clinical samples to determine whether interlaboratory variability could impact the
clinical classification of patients when using published clinical al-
gorithms.

MATERIALS AND METHODS

Sample acquisition. Specimens were obtained from residual portions of clinical test material sent to the laboratory for multiplex pneumococcal serology assay. Initial sample selection was performed without knowledge of patient clinical history or vaccination status. Samples were acquired from a total of 57 patients (generating 741 individual serology assay results per laboratory) for use in the final validation set. The majority of these specimens were single, unpaired specimens from adult patients, reflecting the population that our laboratory serves. All samples were collected according to local Institutional Review Board policies governing the use of residual material for assay validation.

Pneumococcal serology testing. Samples were tested with three different multiplex assays for the measurement of serotype-specific pneumococcal IgG antibodies. Samples were tested on site at the Cleveland Clinic with the xMAP Pneumo14 pneumococcal immunity panel (Luminex Corporation, Austin, TX). This IVD assay detects antibodies against 14 pneumococcal serotypes, i.e., 1, 3, 4, 6B, 7F, 8, 9N, 9V, 12F, 14, 18C, 19A, 19F, and 23F. Standard values for the assay are calibrated to FDA-89 reference serum, and assay buffers incorporate cell wall polysaccharide and polysaccharide type 22 as blocking agents. The performance char-
acteristics of this assay have been reported previously (14, 16, 17). In addi-
tion, aliquots of each sample were sent in a blinded fashion to two com-
nercial reference laboratories that perform multiplex pneumococcal serologic testing. Both reference laboratories use LDT assays based on the Luminex platform, but the specific analytical details of these assays have not been published. Both LDT assays A and B measure antibodies against the following 14 serotypes, 1, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 12F, 14, 18C, 19A, 19F, and 23F, which differ from the IVD assay by one serotype, substituting serotype 19A for serotype 5.

Statistical analysis. Method comparisons and correlation statistics for quantitative results were calculated with EP Evaluator software (release 9). Comparison of population medians between assays was performed by one-way analysis of variance (ANOVA) on ranks with SigmaPlot 11.2 software. For serotypes where significant differences were identified by ANOVA, pairwise multiple-comparison testing (Tukey’s test) was performed to identify the group that differed from the others.

RESULTS

Interlaboratory variation in quantitative results. We initially compared the absolute levels of pneumococcal antibodies measured for each serotype in order to determine agreement in quan-
titative results between laboratories. A total of 2,223 serotype results were generated from the 57 samples in the test set. The 13 serotypes that were common to all of the assays were evaluated independently, and correlation statistics were calculated for all of the possible pairwise combinations of the assays. Substantial varia-
tion between laboratories was seen, with only 61% of the compar-
isons achieving r values of >0.8 (Table 1). However, the degree of variation differed widely between serotypes. While some sero-
types showed good correlation across all of the laboratories, others showed very little agreement between assays (Fig. 1). In addition to this sporadic serotype-specific variability, a systematic negative bias was seen for LDT B relative to the other two assays across multiple serotypes (Table 1). This bias was most evident when comparing the median values for each serotype in the sample set across the three assays, where the values generated by LDT B were lower for most of the serotypes (Fig. 2).

Interlaboratory variation in qualitative serotype classifica-
tion. In order to determine whether the variations in pneumococcal antibody measurements would impact clinical decision making, we examined how each assay classified serotypes as “protective” or “nonprotective” versus a clinically defined cut point. Because our sample set was composed primarily of single time point values from adult patients, we chose to use the classification scheme suggested by Paris et al., which defines an adequate re-
response as an antibody level greater than 1.3 μg/ml (10, 18). This type of qualitative classification improved the overall agreement between laboratories, with rates of concordance near 80% for all pairwise comparisons between laboratories (Table 2). When compar-
ing IVD and LDT A, discrepancies between the two assays were fairly evenly distributed in terms of which assay fell above the “protective” cutoff. In contrast, the results of comparisons involving LDT B were more skewed, with 93% of the discrepancies involving samples where LDT B gave a result below the cutoff. This is consistent with the systematic negative bias of the LDT B results seen in the quantitative comparison.

Interlaboratory variation in the clinical classification of pa-
tients as “protected” versus “nonprotected.” Because the inter-
pretation of a patient’s vaccination status incorporates informa-
tion from multiple serotypes, we evaluated whether variations in serotype classifications between assays would impact the final cat-

### Table 1 Interassay correlation statisticsa

| Serotype | IVD vs LDT A | IVD vs LDT B | LDT A vs LDT B |
|----------|--------------|--------------|----------------|
|          | r | Slope | Intercept | Bias | r | Slope | Intercept | Bias | r | Slope | Intercept | Bias |
| 1        | 0.73 | 2.12 | -1.72 | 1.55 | 0.57 | 0.49 | 1.36 | -0.12 | 0.46 | 0.19 | 1.95 | -1.67 |
| 3        | 0.71 | 0.41 | 0.78 | -1.43 | 0.56 | 0.14 | 0.55 | -2.49 | 0.70 | 0.36 | 0.23 | 0.12 |
| 4        | 0.93 | 1.31 | -0.01 | 0.35 | 0.98 | 0.73 | -0.06 | -0.38 | 0.94 | 0.55 | -0.05 | -0.72 |
| 6B       | 0.89 | 0.62 | 0.65 | -1.20 | 0.90 | 0.44 | 0.74 | -2.44 | 0.85 | 0.95 | -0.24 | -0.42 |
| 7F       | 0.83 | 2.19 | -2.29 | 1.99 | 0.85 | 0.64 | -0.60 | -1.90 | 0.79 | 0.36 | 0.17 | -3.99 |
| 8        | 0.88 | 1.12 | -0.39 | -0.01 | 0.65 | 0.46 | 0.37 | -1.80 | 0.76 | 0.75 | -0.52 | -1.42 |
| 9N       | 0.89 | 1.26 | 0.74 | 0.12 | 0.93 | 0.74 | -0.54 | -1.38 | 0.91 | 0.59 | -0.11 | -1.49 |
| 9V       | 0.71 | 1.52 | -0.88 | 0.84 | 0.93 | 0.63 | 0.23 | -1.45 | 0.95 | 0.37 | 0.12 | -2.44 |
| 12F      | 0.92 | 0.29 | 0.88 | 0.15 | 0.98 | 0.25 | 0.22 | -2.52 | 0.73 | 0.51 | -0.12 | -1.35 |
| 14       | 0.74 | 0.95 | -0.71 | -1.04 | 0.84 | 0.53 | 0.26 | -3.43 | 0.80 | 0.54 | 0.26 | -2.39 |
| 18C      | 0.79 | 3.30 | -3.25 | 6.62 | 0.90 | 0.90 | -1.10 | -1.55 | 0.60 | 0.21 | 0.43 | -8.17 |
| 19F      | 0.39 | 4.97 | -16.98 | 3.44 | 0.53 | 0.77 | 0.82 | -2.02 | 0.91 | 0.35 | 0.10 | -5.46 |
| 23       | 0.90 | 1.12 | 0.03 | 0.34 | 0.85 | 0.68 | 0.15 | -0.67 | 0.92 | 0.63 | 0.08 | -1.01 |

a All possible pairwise comparisons of the three assays are shown for each serotype. Regression statistics were calculated by Deming analysis, while bias is presented as the average absolute bias for each comparison.
egorization of a patient’s immune response. On the basis of the previously referenced criteria, we classified a patient as “protected” if 70% of the serotypes (9/13) achieved a level of >1.3 μg/ml. By these criteria, 82% of the patients (47/57) were identically classified by all three assays, with 96% agreement (55/57) between IVD and LDT A. The majority of the discordant results were cases where patients were classified as nonprotected by LDT B and protected by the other two assays (Fig. 3).

DISCUSSION
Multiplex pneumococcal serology assays are among the most analytically complex assays used in clinical practice, and the use of

FIG 1 Quantitative correlation of representative serotypes. Examples of serotypes showing good (serotype 4, $r > 0.9$) and poor (serotype 19F, $r < 0.6$) correlations between assays are shown.

FIG 2 Comparison of pneumococcal antibody levels. For each assay, the median value of the 57 samples tested is shown for each serotype. The results for serotypes marked with an asterisk were statistically significantly different ($P < 0.05$) from the median results of the other assays.
differing methods introduces the potential for analytical variability between laboratories. In this study, we evaluated whether this variability is significant enough to impact the final clinical interpretation when assessing vaccination responses. Substantial variation was seen in the absolute values of pneumococcal antibodies reported by the three different assays. Both serotype- and assay-specific variations were observed. The underlying causes of the differences are not clear but may relate to differences in the specific reagent formulations and calibration materials used by the different laboratories to validate the assays. A more consistent approach to the standardization of pneumococcal antibody assays could potentially help to reduce these differences. Detailed recommendations in this regard have previously been published by the WHO for the standardization of ELISAs (www.vaccine.uab.edu), and reference standards are available for this purpose.

Despite the variability seen in the reported quantitative results of the various assays, the overall classifications of patients’ pneumococcal immune status were remarkably similar between assays. This appears to be due primarily to two factors. First, although pneumococcal serology assays provide quantitative results, many clinical algorithms use these results in a qualitative fashion by converting results into “protective” or “nonprotective” on the basis of a clinically defined cut point. The data shown here suggest that interpreting results in this manner helps to mitigate much of the analytical variability seen between assays. In addition, the evaluation of multiple serotypes limits the impact that any one serotype has on the overall classification of the patient status. Because multiple serotypes are considered in the final classification, isolated differences in individual serotypes are usually offset by the larger agreement across the remaining serotypes, leading to an overall concordance in the classification.

LDT B provided an interesting illustration of these benefits. Despite the systematic negative bias seen in this assay, results from LDT B still agreed with the consensus classification for 80% of the patient samples. In addition, most of the discordant samples were from patients given an overall classification of “nonprotective” by LDT B but achieving protective levels for eight serotypes (narrowly missing the nine required for 70%). Clinicians should be aware of the specific characteristics of the assay they are using, as an awareness that an assay tends to give lower results might lead them to give further consideration to patients with “borderline” results.

A potential limitation of this study is that we did not examine the impact of analytical variability on algorithms that use the difference between pre- and postvaccination levels to define a successful immune response (19), which many authors have suggested is the most sensitive way to evaluate an immune response to pneumococcal variation (9). Unfortunately, such samples are uncommon in the patient population we serve, as less than 5% of the

### TABLE 2 Interassay agreement in qualitative classification of serotypes

| Serotype | IVD vs LDT A | IVD vs LDT B | LDT A vs LDT B |
|----------|--------------|--------------|----------------|
|          | Agree | IVD | LDT A | % Agreement | IVD | LDT B | % Agreement | IVD | LDT B | % Agreement | LDT A | LDT B | % Agreement |
| 1        | 49    | 3   | 5 | 86 | 45 | 4 | 8 | 79 | 51 | 2 | 4 | 89 |
| 3        | 48    | 6   | 3 | 84 | 42 | 15 | 0 | 74 | 45 | 11 | 1 | 79 |
| 4        | 45    | 3   | 9 | 79 | 51 | 6 | 0 | 89 | 47 | 10 | 0 | 82 |
| 6B       | 49    | 6   | 2 | 86 | 50 | 7 | 0 | 88 | 51 | 4 | 2 | 89 |
| 7F       | 48    | 7   | 2 | 84 | 35 | 22 | 0 | 61 | 41 | 16 | 0 | 72 |
| 8        | 49    | 6   | 2 | 86 | 46 | 11 | 0 | 81 | 48 | 9 | 0 | 84 |
| 9N       | 55    | 1   | 1 | 96 | 53 | 4 | 0 | 93 | 53 | 4 | 0 | 93 |
| 9V       | 49    | 2   | 6 | 86 | 49 | 8 | 0 | 86 | 41 | 15 | 1 | 72 |
| 12F      | 36    | 11  | 10 | 63 | 42 | 15 | 0 | 74 | 42 | 13 | 2 | 74 |
| 14       | 50    | 6   | 1 | 88 | 42 | 15 | 0 | 74 | 49 | 8 | 0 | 86 |
| 18C      | 49    | 1   | 7 | 86 | 48 | 9 | 0 | 84 | 41 | 16 | 0 | 72 |
| 19F      | 44    | 7   | 6 | 77 | 34 | 22 | 1 | 60 | 35 | 22 | 0 | 61 |
| 23       | 54    | 2   | 1 | 95 | 50 | 6 | 1 | 88 | 50 | 6 | 1 | 88 |
| Total    | 625   | 61  | 55 | 84 | 587 | 144 | 10 | 79 | 594 | 136 | 11 | 80 |

---

*Serotype results were classified as protective or nonprotective for each sample by using a cut point of 1.3 μg/ml. Results that were classified identically by both assays are listed under "Agree," while samples classified as protective by one assay and nonprotective by the other are listed under the assay giving the protective result (IVD, LDT A, or LDT B).*  

*Each value is the number of samples unless otherwise specified.*
REFERENCES

1. Centers for Disease Control and Prevention (CDC). 1997. Prevention of pneumococcal disease: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Morb. Mortal. Wkly. Rep. 46:1–24.

2. Wernette CM, Frasch CE, Madore D, Carlone G, Goldblatt D, Plikaytis B, Benjamin W, Quataert SA, Hildreth S, Sikkema DJ, Kayhny H, Jonsdottir I, Nahm MH. 2003. Enzyme-linked immunosorbent assay for quantitation of human antibodies to pneumococcal polysaccharides. Clin. Diagn. Lab. Immunol. 10:514–519.

3. Lin J, Kaltoft MS, Brandao AP, Echaniz-Aviles G, Brandileone MC, Hollingshead SK, Benjamin WH, Nahm MH. 2006. Validation of a multiplex pneumococcal serotyping assay with clinical samples. J. Clin. Microbiol. 44:383–388.

4. Marchese RD, Puchalski D, Miller P, Antonello J, Hammond O, Green T, Rubinstein LJ, Caulfield MJ, Sikkema D. 2009. Optimization and validation of a multiplex, electrochemiluminescence-based detection assay for the quantification of immunoglobulin G serotype-specific antipneumococcal antibodies in human serum. Clin. Vaccine Immunol. 16:387–396.

5. Elberse KE, Therniaeva I, Berbers GA, Schouls LM. 2010. Optimization and application of a multiplex bead-based assay to quantify serotype-specific IgG against Streptococcus pneumoniae polysaccharides: response to the booster vaccine after immunization with the pneumococcal 7-valent conjugate vaccine. Clin. Vaccine Immunol. 17:674–682.

6. Uddin S, Borrow R, Haeney MB, Moran A, Warrington R, Balmer P, Arkwright PD. 2006. Total and serotype-specific pneumococcal antibody titres in children with normal and abnormal humoral immunity. Vaccine 24:5637–5644.

7. Balmer P, Cant AJ, Borrow R. 2007. Anti-pneumococcal antibody titre measurement: what useful information does it yield? J. Clin. Pathol. 60:345–350.

8. Balmer P, North J, Baxter D, Stanford E, Melegaro A, Kaczmarcki EB, Miller E, Borrow R. 2003. Measurement and interpretation of pneumococcal IgG levels for clinical management. Clin. Exp. Immunol. 133:364–369.

9. Kamchaisatan W, Wanwatsuptikul W, Seasman JW, Tangsimmankong N. 2006. Validation of current joint American Academy of Allergy, Asthma & Immunology and American College of Allergy, Asthma and Immunology guidelines for antibody response to the 23-valent pneumococcal vaccine using a population of HIV-infected children. J. Allergy Clin. Immunol. 118:1336–1341.

10. Paris K, Sorensen RU. 2007. Assessment and clinical interpretation of polysaccharide antibody responses. Ann. Allergy Asthma Immunol. 99:462–464.

11. Landesman SH, Schifman G. 1981. Assessment of the antibody response to pneumococcal vaccine in high-risk populations. Rev. Infect. Dis. 3(Suppl):S184–S197.

12. Balmer P, Borrow R, Findlow J, Warrington R, Frankland S, Wight P, George R, Andrews N, Miller E. 2007. Age-stratified prevalences of pneumococcal-serotype-specific immunoglobulin G in England and their relationship to the serotype-specific incidence of invasive pneumococcal disease prior to the introduction of the pneumococcal 7-valent conjugate vaccine. Clin. Vaccine Immunol. 14:1442–1450.

13. Siber GR, Chang I, Baker S, Fernsten P, O’Brien KL, Santosham M, Klugman KP, Madhi SA, Paradiso P, Kohberger R. 2007. Estimating the protective concentration of anti-pneumococcal capsular polysaccharide antibodies. Vaccine 25:3816–3826.

14. Pickering JW, Martins TB, Greer RW, Schroder MG, Astill ME, Litwin CM, Hildreth SW, Hill HR. 2002. A multiplexed fluorescent microsphere immunoassay for antibodies to pneumococcal capsular polysaccharides. Am. J. Clin. Pathol. 117:589–596.

15. Whaley MJ, Rose C, Martinez J, Laher G, Sammons DL, Smith JP, Snawder JE, Borrow R, Biagini RE, Plikaytis B, Carlone GM, Romero-Steiner S. 2010. Interlaboratory comparison of three multiplexed bead-based immunoassays for measuring serum antibodies to pneumococcal polysaccharides. Clin. Vaccine Immunol. 17:862–869.

16. Borgers H, Moens I, Picard C, Jeurissen A, Raes M, Sauer K, Proesmans M, De Boeck K, Casanova JL, Meyts I, Bossuyt X. 2010. Laboratory diagnosis of specific antibody deficiency to pneumococcal polysaccharide antigens by multiplexed bead assay. Clin. Immunol. 134:198–205.

17. Jeurissen A, Moens I, Raes M, Wuyts G, Wellebrods L, Sauer K, Proesmans M, Ceuppens JL, De Boeck K, Bossuyt X. 2007. Laboratory diagnosis of specific antibody deficiency to pneumococcal capsular polysaccharide antigens. Clin. Chem. 53:505–510.

18. Bonilla FA, Bernstein IL, Khan DA, Ballas ZK, Chinen J, Frank MM, Kobrinsky NJ, Levinson AI, Mazer B, Nelson RJ, Oranje JS, Routes JM, Shearer WT, Sorensen RU, American Academy of Allergy, Asthma, and Immunology, American College of Allergy, Asthma, and Immunology, Joint Council of Allergy, Asthma, and Immunology. 2005. Practice parameter for the diagnosis and management of primary immunodeficiency. Ann. Allergy Asthma Immunol. 94:51–63.

19. Hare ND, Smith BJ, Ballas ZK. 2009. Antibody response to pneumococcal vaccination as a function of preimmunization titer. J. Allergy Clin. Immunol. 123:195–200.