Antibody-Based Protein Multiplex Platforms: 
Technical and Operational Challenges

Allison A. Ellington,1 Iftikhar J. Kullo,1 Kent R. Bailey,2 and George G. Klee3*

BACKGROUND: The measurement of multiple protein biomarkers may refine risk stratification in clinical settings. This concept has stimulated development of multiplexed immunoassay platforms that provide multiple, parallel protein measurements on the same specimen.

CONTENT: We provide an overview of antibody-based multiplexed immunoassay platforms and discuss technical and operational challenges. Multiplexed immunoassays use traditional immunoassay principles in which high-affinity capture ligands are immobilized in parallel arrays in either planar format or on microspheres in suspension. Development of multiplexed immunoassays requires rigorous validation of assay configuration and analytical performance to minimize assay imprecision and inaccuracy. Challenges associated with multiplex configuration include selection and immobilization of capture ligands, calibration, interference between antibodies and proteins and assay diluents, and compatibility of assay limits of quantification. We discuss potential solutions to these challenges. Criteria for assessing analytical multiplex assay performance include the range of linearity, analytical specificity, recovery, and comparison to a quality reference method. Quality control materials are not well developed for multiplexed protein immunoassays, and algorithms for interpreting multiplex quality control data are needed.

SUMMARY: Technical and operational challenges have hindered implementation of multiplexed assays in clinical settings. Formal procedures that guide multiplex assay configuration, analytical validation, and quality control are needed before broad application of multiplexed arrays can occur in the in vitro diagnostic market.

© 2009 American Association for Clinical Chemistry
tion times due to shorter diffusion distances, compared to traditional immunoassays. Two-dimensional planar multiplexes consist of high-density microspots of capture ligands (<250 μm diameter; >1000 spots/cm²) immobilized on a rigid surface at spatially discrete locations such that multiple capture ligands are immobilized in 1 well (Fig. 1). Lumiphores are the most common reporter in planar assays because the resulting chemiluminescent signal confers high sensitivity and wide dynamic range (approximately 5 logs) (9). Recently, electrochemiluminescent technology has been used, in which labels such as Ru(bpy)₃²⁺ emit signal only when in close proximity to a stimulated electrode surface (9). Signals are enhanced by microscopy, and captured images are analyzed with platform-specific software packages. Although planar assays are often performed manually, automation of assays can increase assay robustness and sample throughput. The feasibility of automation has been demonstrated using an automated liquid pipettor to add samples and assay reagents (10).

Three-dimensional suspension multiplexes use microspheres of approximately 5.3 μm diameter as the solid support to which different capture ligands are covalently coupled using N-hydroxysuccinimide ester chemistry (Fig. 1) (11). Compared to planar multiplexes, in which protein identification relies on x,y position, microspheres use classifiers such as size or internal fluorophores for assay assignment. Flow cytometric principles are used to generate results: assay-specific microspheres are distinguished by either light scatter (size) or internal fluorescent ratio, and an assay-dependent signal is generated by additional fluorophore labels. Fluorescence-activated cell sorting, the basic technology used in suspension array detection, has been routinely used in diagnostics for >20 years (12, 13). Suspension multiplexes have a dynamic assay range of approximately 3 logs (14). An advantage of multiplexed suspension assays, compared with planar multiplexes, is improved imprecision as a result of multiple (50–100) independent measurements within each microsphere population. Suspension assays are

![Fig. 1. Planar and suspension multiplexed immunoassay formats. In planar assays, capture ligands are immobilized on a rigid 2-dimensional support and probed with sample. Fluorescent or chemiluminescent signal is identified by x,y coordinates. In suspension assays, capture ligands are immobilized on color- or size-coded microspheres. Assays are distinguished by coding attributes, and flow cytometry is used to detect assay-specific fluorescent signal.](image-url)
also readily automated with commercial liquid handling solutions. A disadvantage of suspension assays is that the filter-bottom microplates used for washing microspheres are prone to leaking, clogging, and nonspecific analyte adsorption to filter surfaces (15). Implementation of magnetic beads may overcome this limitation. Magnetic beads composed of Co:Nd:Fe₂O₃ allow manipulation by an external magnetic field in separation and washing assay steps (16).

**Multiplex Immunoassay Configuration**

Developing multiplexed immunoassays requires rigorous validation of assay configuration and analytical performance to minimize measurement imprecision. Optimization of capture ligand selection and immobilization, assay calibration, and assay performance (determination of sample incubation times, reduction of cross-reactivity between assays, and nonspecific protein binding) are important development components for both planar and suspension multiplex formats.

**Capture Ligands**

A key step for development of robust multiplexed protein immunoassays is generation and characterization of capture ligands. These systems depend on the availability of highly specific capture ligands that bind tightly to target proteins. However, it is difficult to predict which materials bind to target proteins with high affinity because of the heterogeneity in size and hydrophobicity of constituent amino acids (17). Monoclonal antibodies produced by classical hybridoma techniques are the most common capture ligands (18). However, monoclonal antibody production and characterization of affinity, specificity, cross-reactivity, and kinetic parameters (association and dissociation rates) are labor- and cost-intensive processes. Despite the direct effects of antibody characteristics on assay performance, current antibody “validation” protocols are often inadequate; some evaluate only molecular weight of the protein band on a Western blot (19). Recently, 2 high-throughput methods for antibody characterization using multiplexed immobilized proteins or peptides have been described. One uses a planar format to provide concurrent analyses of epitope recognition and binding affinity to determine antibody specificity and affinity (20). The other uses a suspension format on the Luminex platform to determine antibody specificity toward as many as 100 antigens in several hours (21).

Antibodies validated for monoplex immunoassays may display cross-reactivity with other proteins in multiplex formats, highlighting a need for application-specific antibody validation (22). Additionally, sandwich immunoassays require 2 validated antibodies for each constituent protein.

Given the limitations of antibodies, alternative capture ligands such as engineered protein scaffolds and nucleic acid scaffolds are being evaluated (23). For example, aptamers, highly specific nucleic acid molecules that distinguish between protein isoforms and conformations, possess target recognition features as antibodies do (24). Recent initiatives in Europe and the US are aimed at improving the availability and quality of affinity reagents, with the ultimate goal of creating standardized collections of well-validated binding molecules for proteomic analyses to improve assay quality (23).

**Antibody Immobilization**

Antibody immobilization is crucial for reproducible measurements. The ideal chemistry for antibody immobilization must have high binding capacity, the ability to retain immunologic activity, high signal-to-noise ratio, and low variability between batches (25, 26). Although antibodies have less potential for denaturation than do other proteins, even minimal denaturation may expose hydrophobic regions, thereby increasing nonspecific protein binding and potentially altering the lower limit of quantification. In planar microarrays, immobilization parameters that influence robust immunoassay performance include spot size and morphology, total antibody binding capacity, background signal, limit of detection, and spotting reproducibility within and across arrays (26). The importance of evaluating these parameters was illustrated by the generation of calibration curves for 23 assays across 17 commercially available slide chemistries; these curves showed that slide surface properties affect immobilized antibody activity and subsequent data quality (26). In multiplexed suspension assays, standard antibody immobilization protocols may mask antigen recognition epitopes composed of large fractions of lysine or arginine residues, resulting in failure of an antibody to recognize its respective antigen (21).

**Manufacturing Process**

Variability in assay manufacturing processes also contributes to assay imprecision. In planar multiplexes, array printing is a major source of variation. Robots print arrays by depositing spots of capture antibodies onto microplate wells in a predefined pattern. Printing tip architecture such as solid pins and piezoelectric (inkjet) distinguishes microarray printing formats (19, 27). Solid pin contact printers draw antibodies
from a reservoir into a pin and deposit the antibodies by capillary action. The quantity of antibody deposited depends on the length of time the pin rests on the microplate. A potential limitation of contact array printers is decreased antibody deposition as the source reservoir is depleted (27). Noncontact piezoelectric printers use an electrical pulse to dispense a discrete amount of capture ligand per spot, theoretically decreasing spot-to-spot variability and variability overall between plates (27). However, this technology is not ideal for depositing viscous materials or complex mixtures and requires lengthy wash cycles (27). High-quality planar arrays depend on stringent printing procedures. For example, we observed printing errors by visually inspecting chemiluminescent images obtained from planar arrays, which highlights a need for more robust printing methods (Fig. 2A) (28).

Microsphere production is a key source of variability in suspension assay signals. In multiplexed suspension immunoassays, 10% to 32% of measurement variability may be attributable to variations in microsphere diameter. Assay imprecision may be improved by modifying the manufacturing process to provide tighter lot QC (Fig. 2B) (29).

Multiplex Calibration

Quantitative tests are preferred over qualitative tests because the former provide greater information about modulation of protein concentrations in response to disease progression and treatment interventions. The accuracy of quantitative immunoassays depends on development and validation of appropriate calibrators and calibration methods. Calibration materials, including high-quality, purified natural and recombinant proteins, provide the basis for relating signal responses to concentrations of proteins. Optimal conditions for high-throughput in vitro production of soluble proteins can be identified using multifactorial analyses (30). Full factorial experiments that include combinations of categorical and continuous factors allow testing of multiple conditions. Multiple regression analyses identify important factors affecting soluble protein expression, and additional statistical techniques optimize these expression conditions (30). Optimal conditions are then used for high-throughput production of soluble proteins that can be used to calibrate planar and suspension assays.

Calibration methodologies also require careful consideration. In general, calibration curves consist of a series of known protein concentrations that are plotted against assay signal, and a mathematical expression is fitted to the curve. Nonlinear regression models with either 4 or 5 parameters are frequently used (31). It is difficult to apply the same calibration curve across multiple microtiter plates; therefore, calibration standards are added to each plate, decreasing sample testing space and increasing assay cost. Calibration curve signals are also vulnerable to random variability between

Fig. 2. Variations in multiplex assay manufacture may introduce measurement error.
(A), A planar immunoassay 6-plex emits chemiluminescent signal from 7 discrete locations. Green boxes enclose calibration curve. Red boxes enclose duplicate samples with imprecision >30%. (B), A pair of microspheres (49, 9.1 μm; 50, 6.3 μm) illustrates the size differences of beads used in suspension assays [Hanley et al. (29)]. Panel B is adapted from Hanley et al. (29) and is being reprinted according to the Creative Commons Attribution license (http://creativecommons.org/licenses/by/2.0); the original article has the following copyright and licensee information: "© 2007 Hanley et al; licensee BioMed Central Ltd."
batches, a factor that can introduce significant measurement variability.

Cross-Reactivity

Issues relevant to multiplexed protein immunoassay development include elimination of assay interference between reagents and configuration of assay sensitivities to provide acceptable dynamic ranges for each of the multiplexed proteins in the targeted specimen matrix (2). Cross-reactivity between detection antibodies and immobilized capture ligands and nonspecific analytes limits the number of proteins that can be used in a given multiplex. Some protein combinations are not possible because nonspecific binding may produce a large background signal, thereby decreasing assay sensitivity. For example, the sensitivity of an 11-plex planar assay was decreased by factors of 1.7–5.0, compared with monoplex ELISAs, because of higher background signals in the multiplexed format (2). Antibody cross-reactivity may limit protein measurements to 30–50 proteins in planar format (32). Theoretically, suspension assays are more vulnerable to cross-reactivity between proteins because cross-linking may occur as beads circulate in fluid, a factor that may limit the ability to multiplex (1).

Antibody-related assay interference can be evaluated in either assay format with 3 experiments that measure the signal produced when (1) single proteins are incubated with complete detection antibody cocktail; (2) complete protein mixture is incubated with single detection antibodies; and (3) antibody cocktails with 1 antibody removed are incubated with complete protein mixtures to detect cross-reactivity between detection antibodies and specific proteins (33).

Assay diluents and assay buffers must interact effectively with all reagents and proteins included in a multiplex under common assay conditions. This presents a challenge because proteins require specific conditioning to maintain conformation; differences in electrical charge, hydrophobicity, and posttranslational modifications may alter requirements. Minor changes in buffer pH and ionic strength may irreversibly modulate structure, impairing assay performance. For example, comparison of commercially available diluents demonstrated that diluent selection significantly suppressed signal intensity and assay sensitivity in a 19-plex suspension assay (34).

Optimal temperatures, incubation times, and concentrations of reagents must be configured during multiplex assay development. Application of multifactorial design to optimize assay parameters enables rapid development of robust assays with interassay and intraassay imprecision of <20% (35). A basic screening design might include multiple levels of variables such as concentrations of capture and detection antibodies and reporter molecules, and incubation times and temperatures. Also, the dynamic ranges of the constituent assays must be compatible so that all proteins can be measured using a single sample dilution. Although samples may be measured at multiple dilutions, doing so would result in higher assay costs. In addition to adjusting sample dilution, the response of the measurement device may be modulated. For example, on the Luminex platform, adjusting the photomultiplier tube setting modulates the dynamic range and sensitivity of assay calibration curves; increasing voltage enhances sensitivity at the expense of lower recovery in the upper end due to saturation (11).

Analytical Validation

Analytical validation of multiplexed immunoassay performance should be conducted before clinical validation. Proposed analytical performance criteria include range of linearity, analytical specificity, recovery, limits of detection and quantification, reasonable imprecision, and comparison to a quality reference method (36).

The CV, estimated using multiple replicates of an identical sample, describes measurement variation. Ideally, intraassay and interassay imprecision will be similar, so no additional variation is introduced into measurements between batches. Immunoassays have intrinsic intraassay variability such that assaying replicates of the same sample yields a distribution of results. These effects can be decreased by improving the assay or running replicated specimens and averaging the values. In smaller studies using planar multiplexes, intraassay imprecision ranged from 4.2% to 67.0% and interassay imprecision ranged from 9.4% to 56.0% (37–41). However, the reliability and consistency of planar multiplexed protein arrays in large-scale studies is unknown, and allowable levels of imprecision for multiplex immunoassays in general have not yet been defined. In particular, the ability to analyze many samples without affecting interassay imprecision has not been adequately documented. Interassay variability can be controlled to some extent at the study design and analytic stages (discussed below). Pilot studies may be useful for determining reliability and validity of assays. We evaluated sources of interplate variability by fitting to each analyte the linear models that included fixed effects for age, sex, race, and the nominal variable “plate” to production data: 24.4% of the analyte total variability was explained by plate-to-plate differences in assay performance (28). These results highlight the
need for development of formal performance guidelines by the US Food and Drug Administration (FDA).4

Observed measurements of signal intensities contain not only measurement noise but also systemic variations of other kinds. Adequate data transformation, standardization, and normalization before statistical analyses are essential for valid results, yet analyses of protein multiplex data are often reported in the absence of documented data quality assurance programs. Various analytic techniques have been described to minimize the effects of intraassay and interassay imprecision. For example, laboratory drift in patient samples can be assessed and used as a QC parameter, and analyses can be recalibrated using a mean ratio value or regression calibration (42). Rescaling measured responses to a constant level may also normalize assay results (43, 44). Alternatively, linear mixed models can assess quality, minimize random noise, and remove systemic measurement errors by estimating variance components. This methodology has been used to quantify the magnitude of systemic biases (analyst, day, and experiment) and extract data that has been adjusted for these effects (45). Of note, these strategies have not been validated, and data adjustment should not replace development of robust multiplex assays.

Operational Control

Appropriate statistical methods that guide assay validation, QC, and analyses are needed for clinical implementation of protein immunoassay multiplexes. Although reproducibility and reliability are critical for diagnostic assays, QC algorithms are not well developed for multiplexed protein immunoassays (36). Single-analyte QC protocols have been implemented in most clinical laboratories (46). Imprecision (random error) for the analytical method is characterized by repeatedly measuring control materials, and these data are used to calculate reference ranges for control materials. Analytical performance of monoplex immunoassays is routinely monitored by analyzing materials with known low, medium, and high protein concentrations. Observed results are compared with the control reference ranges, and observed results falling outside of the reference ranges signify unacceptable assay performance and rejection of data. QC considerations are compounded as the number of individual reactions on an array increases, and generation of control materials with appropriate concentrations of each constituent protein is a major challenge to clinical use of protein multiplexes (36). Control materials must be manufactured in large quantities to provide a sustained source of protein in a matrix similar to the testing specimen, with minimal variability in concentration from aliquot to aliquot or from lot to lot. Defined mixtures of control proteins and protein solubility at high concentrations may be problematic for multiplexed assays (47). These issues are receiving attention from agencies, such as the National Institute of Standards and Technology, that produce and distribute standard reference materials for clinical applications (48).

Even as hurdles in material development are overcome, devising rules for interpreting multiplex QC data and determining the acceptability of the analytical results present major challenges (49). When all results are within the analyte-specific reference ranges for all analytes, QC is not problematic. However, the chances of a QC failure increase with increasing numbers of individual reactions performed on multiplexed assays, and it is unclear how to proceed when measurements for 1 or more constituent proteins fall outside the limits. One option is to reject test results for all constituent analytes and reanalyze controls and samples on additional multiplexed panels. An alternative is to selectively accept data for constituent proteins that passed each level of QC (49). The action is dependent on whether multiplexed data are considered in their entirety or whether each constituent assay is considered separately. An example of the former is the FDA-cleared BioPlex suspension assay for antinuclear antibody testing on the Luminex BioPlex 2200 analyzer that includes 2 control materials that contain normal and abnormal quantities of each of the 13 proteins. Results are reported only for proteins that pass QC specifications; since the Medical Decision Support Software requires data from each test, no output is provided if any 1 analyte fails QC (50). For the latter, retesting strategies are unclear. An option is to develop monoplex or smaller multiplex panels that include subsets of the assays that can be used to retest runs that fail QC, but analytical performance of smaller panels may differ significantly from the original, larger panels, and the effort and cost of validating unique combinations of proteins would be substantial. These QC issues must be resolved before the use of protein multiplex arrays in clinical laboratories becomes routine.

Clinical Utility

Despite the introduction of hundreds of multiplexed protein immunoassays to the research market in recent years, only a limited number have been cleared by the FDA for clinical use, an observation that illustrates the complexity of constructing robust arrays. FDA-cleared planar protein multiplex arrays consist primarily of
the lateral flow immunoassays used for point-of-care evaluation (51). For example, the Triage® Cardio ProfilER® 4-plex measures troponin-I, creatine kinase-MB, myoglobin, and brain natriuretic peptide (BNP) to assist with evaluation of chest pain using a portable lateral flow platform. At present, suspension immunoassay is the prevailing technology for FDA-cleared multiplex protein measurements, especially for testing antibodies in the serum of patients with allergies or autoimmune or infectious diseases in clinical laboratories (15).

Conclusion

Multiplexed immunoassays, either planar or suspension, have the potential to deliver data of diagnostic and prognostic value. Multimarker strategies may improve medical diagnostic and prognostic information; therefore, additional protein microarrays for use in cancer, stroke, diabetes, and cardiovascular diseases are under development. Implementation of these assays in the clinical setting has lagged because of technical and developmental challenges. Although multiplex protein immunoassays are able to measure multiple analytes, analytical and QC problems must be solved before broad application of multiplexed arrays in the in vitro diagnostics market. Clinical applications must follow establishment of globally accepted calibration standards, performance criteria, and QC programs.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures of Potential Conflicts of Interest: Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:

Employment or Leadership: None declared.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
 Honoraria: None declared.
Research Funding: NIH grant U01-HL-81331 from National Heart, Lung and Blood Institute.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

References

42. Tworoger SS, Hankinson SE. Use of biomarkers in epidemiologic studies: minimizing the influence of measurement error in the study design and analysis. Cancer Causes Control 2006;17:889–99.