Journal of Immunological Methods 387 (2013) 1–10

Contents lists available at SciVerse ScienceDirect

Journal of Immunological Methods

journal homepage: www.elsevier.com/locate/jim

Research paper

A condensed performance-validation strategy for multiplex detection kits used in studies of human clinical samples

Chahrazed Belabani [a], Sathyanath Rajasekharan [a], Viviane Poupon [a], Trina Johnson [a], Amit Bar-Or [a,b,*] and on behalf of the CIHR/MSSC NET in Clinical Autoimmunity [1] and The Canadian Pediatric Demyelinating Disease Network [2]

[a] Experimental Therapeutics Program, Montreal Neurological Institute, McGill University, Montreal, QC, Canada
[b] Neuroimmunology Unit, Department of Neurology and Neurosurgery, Montreal Neurological Institute, Montreal, QC, Canada

ARTICLE INFO

Article history:
Received 30 April 2012
Received in revised form 2 August 2012
Accepted 3 August 2012
Available online 11 August 2012

Keywords:
Validation
Multiplex
Accuracy
Precision
Sensitivity
Working range

ABSTRACT

Quantification of soluble phase analytes represents one of the most commonly used techniques applied to a broad range of samples in both basic and clinical immunology laboratories, as well as in context of drug development and diagnostic programs. The recent increase in the application of multiplex immunoassays, such as Luminex, has resulted in a growing array of commercially available multiplex kits. Validated, highly sensitive, and precise methods for such quantification is critical, especially when applied to precious sample collections. While vendors are expected to carry out kit performance validation, discrepancies between technical specifications provided with multiplex kits and their actual performance can be relatively common. Here we present a validation strategy that will aid users to select the optimal kits for their purpose and most validly interpret results from the multiplex assays. To illustrate key considerations when validating and comparing kits, we assess the performance of three conventional multiplex cytokine kits. Our findings confirm the importance of validating the performance of commercial multiplex kits and provide a practical and cost-effective approach that can be readily implemented in both academic and translational laboratory settings.

1. Introduction

Recent advances building on the fundamental principles of Enzyme-Linked Immunosorbant Assay (ELISA) have led to the development of microsphere-based multiplex immunoassays, which are capable of quantifying multiple proteins within the same small fluid volume, resulting in time- and cost-savings (Probst et al., 2003). Several reports have shown a strong correlation between ELISA results and results generated using multiplex technology (Dupont et al., 2005; de Jager and Rijkers, 2006; Ray et al., 2005; Pickering et al., 2002). The use of multiplex immunoassays such as Luminex technology is thus expanding in both basic and applied research leading to a growing array of commercially available multiplex kits.

While commercially available kits undergo a vendor-initiated validation process, several comparative studies have shown considerable variability between different Luminex kits, as well as discrepancies between the kit technical specifications provided by the vendor and the actual assay performance for particular molecules such as cytokines (Khan et al., 2004; Prabhakar et al., 2004; Nechansky et al., 2008; Khan et al., 2009 ; Berthoud et al., 2011). One conclusion of these studies is that performing an assay according to a vendor's recommendation does not

* Corresponding author at: Neuroimmunology Unit and Experimental Therapeutics Program, Montreal Neurological Institute, McGill University, 3801 University Avenue, Montreal, Quebec, Canada H3A 2B4.
E-mail address: amit.bar-or@mcgill.ca (A. Bar-Or).
[1] Members of the CIHR/MSSC NET in Clinical Autoimmunity.
[2] Members of the Canadian Pediatric Demyelinating Disease Study are listed in the Acknowledgments.

0022-1759/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jim.2012.08.002

guarantee reliability of the data, thus underscoring the importance of implementing a user-based validation approach to ensure the validity of data interpretation.

The validation of individual kit performance, including the use of 'validation samples' (Lee et al., 2005), enables the documentation of the specific performance characteristics of the assay including accuracy, precision, selectivity, sensitivity, stability, and reproducibility. The Food and Drug Administration (FDA) and other regulatory agencies require a particularly stringent validation assay for bioanalytical methods relevant to either new drug applications or the modification of established clinical and pre-clinical analyses (CDER et al., 2001). The rigor of such a validation assay, which optimizes sensitivity and selectivity in the drug discovery process, requires the use of up to six additional multiplex kits for each validation process. This approach can be unrealistically time-consuming and cost-prohibitive for many academic labs, which commonly use multiplex kits as tools for exploratory and translational research. Such labs also wish to ensure the reliability of their data, often generated from precious or difficult to obtain samples. Therefore, there is a need for a practical and affordable validation strategy that, even when performed in a non GLP-certified laboratory setting, still ensures that kits are assessed for acceptable accuracy, precision and reproducibility.

Here we outline a condensed validation method that is practical to use with commercially available kits in an academic or translational research setting. Our proposed strategy is based on the "exploratory validation" suggested in the "fit-for-purpose" method validation (Lee, 2009; Lee et al., 2006), and is extrapolated from the FDA guidelines intended for an industry context FDA, 2001. We investigate our suggested validation method using two key steps: (i) A "Kit Performance Verification" step where the standards and the controls are run according to the recommendation of the vendor in order to verify their linearity, and (ii) A "Kit Validation" step, where the working range, sensitivity and the intra- and inter-assay precision and accuracy are determined, as described below.

In addition to the validation of an individual kit, we also consider the likely scenario in which investigators may wish to compare the performance of several commercially available multiplex kits in order to select the preferred kit for their particular study. As a single set of performance parameters may not properly capture relevant parameters for each individual analyte, our validation approach aims in part to discern parameters such as sensitivity, precision, accuracy, and the working range for each analyte within the muti-analyte combination. To guide this decision-making process, we use our proposed validation strategy to compare the performance of three commercially available cytokine multiplex kits. Our findings confirm the importance of validating the performance of commercial multiplex kits and provide a practical and cost-effective approach that can be readily implemented in both academic and translational laboratory settings.

2. Methods

2.1. Multiplex kits and Luminex analysis

We assessed three different kits: a BioRad magnetic bead-based 10-plex kit (Cat M50036S98J, Lot 5019034, BioRad, USA); a conventional Invitrogen polystyrene bead-based 10-plex kit (Cat LCH0001, Lot 806011A, Invitrogen, Camarillo); and an Invitrogen polystyrene bead-based 10-plex ultrasensitive kit (Cat LHC6004, Lot 820239A, Invitrogen, Camarillo). We focused on cytokines that are common between the three kits: IL-1β, IL-2, IL-4, IL-5, IL-10, GM-CSF and IFNγ. Cytokine concentrations were measured using a Luminex LX100 instrument (Luminex, Austin, TX, USA). Data generated from the instrument was evaluated using the Bio-Plex Manager 5.0 software (BioRad). Standards and controls were reported as the average of duplicate measurements. The concentration of each target for each sample was calculated using a 5-parameter logistic fit curve generated from the standards.

2.2. Confirming assay working range and sensitivity

To account for a potentially wider assay range based on lower true sensitivities of the different analytes, the standard curve of each kit was extended beyond the manufacturer's recommendations by the addition of points at the lower level (highest dilution) of the standard curve (Fig. 1). Specifically, the standard curves were extended down to 0.1 pg/ml for conventional kits, and down to 0.02 pg/ml for the ultrasensitive kit. The resulting three lowest standards were considered as the 'Lower Limit of Quantification validation samples' and included the added samples (LLOQ1 and LLOQ2), and the lowest standard in the kit-recommended calibration curve (LLOQ3). At the high concentration levels, ULOQ1 and ULOQ2, corresponding to the two highest standards in the kit-recommended calibration curve, were considered as the two 'Upper Limit of Quantification validation samples', to confirm the upper limit of the working range. For any standard that did not meet the acceptance criteria (see below), one or both of the replicates of that standard were removed. For each analyte, the fluorescence index (FI) of the LLOQ had to be greater than the blank. The final calibration curve had to contain at least six acceptable standards.

2.3. Incorporation of validation samples

In addition to confirming the true working range and sensitivity of the kits we incorporated a series of validation samples that served to determine the intra-assay and inter-assay precision and accuracy of each multiplex kit as well as provided flanking quality control samples (flanking QCs) to assess the reliability of measurements across the plate. These flanking QCs were prepared based on the reference standards at three different concentration levels: low levels (QC-L1 and QC-L2), mid level (QC-M) and high level (QC-H). Inclusion of two lower level QC samples (QC-L1 and QC-L2) was done in case both LLOQ1 and LLOQ2 failed to meet acceptance criteria (see below).

2.4. Evaluation of assay precision and accuracy

Precision was defined as the ability of the measured concentration values to be replicated under the same conditions (repeatability). We considered both intra-assay (within-plate) and inter-assay (plate-to-plate) precision. Precision was calculated as the degree of error in a series of measurements (replicates of the same validation sample) obtained from the identical sample and was calculated as a coefficient of variation (%CV) using the following expression: %CV = Standard