Journal of Immunological Methods 437 (2016) 53–63

Contents lists available at ScienceDirect

Journal of Immunological Methods

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

Research paper

Evaluation of highly sensitive immunoassay technologies for quantitative measurements of sub-pg/mL levels of cytokines in human serum

David Yeung [1], Shawn Ciotti, Shobha Purushothama [1], Elham Gharakhani [1], Geoffrey Kuesters [1], Brian Schlain, Chase Shen, Douglas Donaldson, Alvydas Mikulskis *

Biogen, 250 Binney Street, Cambridge, MA 02142, USA

ARTICLE INFO

Article history:
Received 30 June 2016
Received in revised form 19 August 2016
Accepted 19 August 2016
Available online 21 August 2016

Keywords:
Ultra-sensitive immunoassay
Ligand-binding assay
Parallelism
Emerging technologies
Cytokine
Biomarker

ABSTRACT

A comprehensive cross-platform and cross-assay evaluation using nine technology platforms and four cytokine immunoassays (IL-6, TNFα, IL-17a, IL-2) was performed by comparing assay precision, sensitivity, parallelism and data correlation between platforms. The precision was acceptable for most evaluated assays. In addition to comparing the analytical assay sensitivity using a spiked recombinant analyte in buffer, forty serum samples from both normal controls and multiple sclerosis patients were used to measure the frequency of endogenous analyte detection (FEAD) as a parameter of each assay's ability to detect the endogenous analyte. The highest FEAD measurements were observed on the Simoa™, Erenna®, Milliplex® and Imperacer® platforms. However, only Simoa and Erenna results showed a high correlation across all evaluated cytokine assays, followed by a more moderate correlation of results across platforms for the V-plex™, high sensitivity ELISA and the Ella™ IL-6 and TNFα assays. In contrast, results from the evaluated cytokine assays on the Milliplex, AMMP™ ViBE® and Imperacer platforms did not correlate to each other nor to other evaluated assays. Acceptable parallelism was observed for the Simoa, Erenna, V-plex and Ella assays but not for the Milliplex, AMMP ViBE and Imperacer assays. In conclusion, the Simoa, Erenna,V-plex and Ella platforms performed well in one or more evaluated cytokine assays. Among those, the Simoa and Erenna assays had the highest sensitivity for detection of cytokines present at sub-pg/mL levels in human serum. In addition, the cross-platform and cross-assay comparisons demonstrated that different immunoassays may yield different results, which underscores the importance of performing such comparative evaluations, especially in the absence of reliable reference standards for the quantitative assessments of biomarkers in immunoassays.

1. Introduction

Biomarkers are important indicators of physiological or pathological processes and play a key decision-making role in drug development. For example, dysregulated cytokines are involved in a broad spectrum of diseases and utilized routinely as biomarkers for various drug development programs (Turner et al., 2014). These cytokines are typically present in low pg/mL to sub-pg/mL levels in human blood and are challenging to measure using traditional ELISAs or even more sensitive ligand binding assays (LBA). It is rather common that biomarker measurements in biological fluids (e.g., serum, plasma, cerebrospinal fluid) fall below the limit of quantitation (BLQ) thereby affecting quantitative analyses and interpretation of biomarker results. Therefore, the use of highly sensitive technologies may help to reduce the rate of BLQ results.

Over the years, the mainstay of enhanced assay sensitivity has been based on immunoassays using electrochemiluminescence (ECL), tyramide amplification, chemiluminescence, time-resolved fluorescence (TRF) and other detection technologies. However, these conventional assay technologies are not always capable of detecting biomarkers at low levels (low pg/mL or lower) in biological samples. This has led to the development of ligand binding assay technologies and assay platforms with improved assay sensitivity. For example, Erenna® (Singulex) is capable of single molecule counting, which typically improves assay sensitivity in comparison to other assay platforms such as MSD® and Gyrolab™ (Fraser et al., 2014). Significant improvements in assay sensitivity have also been reported with Imperacer® technology (Chimera Biotec GmbH), which is based on signal amplification using the immunopolymerase chain reaction (iPCR). Among other recent technologies that offer improved immunoassay sensitivity and performance are Ella™(ProteinSimple), Simoa HD-1 Analyzer™

Abbreviations: FEAD, frequency of endogenous analyte detection; LBA, ligand binding assay.

* Corresponding author.
E-mail address: Alvydas.Mikulskis@biogen.com (A. Mikulskis).
[1] Formerly Biogen, Cambridge, MA, USA.

http://dx.doi.org/10.1016/j.jim.2016.08.003
0022-1759/

(Quanterix), and Acoustic Membrane Micro-Particle (AMMP™) ViBE® (Bioscale).

Several papers devoted to the comparison of emerging sensitive immunoassay technologies to conventional ligand binding assays (LBA) have been published recently (Fraser et al., 2014; Leary et al., 2013; Myzithras et al., 2016; Soderstrom et al., 2011; Wu et al., 2015;). These evaluations were often focused on assessments of either several assays on a single ultrasensitive assay platform (e.g., Erenna or Simoa) or on a single assay evaluation across several assay platforms. In this study a comprehensive comparison of nine emerging or well-established LBA technology platforms and up to 4 cytokine assays on each platform was performed. The performance of each assay and platform was compared based on sensitivity and accuracy of cytokine measurements in human serum using one or more IL-2, IL-17a, IL-6, and TNFα cytokine assays, which are available commercially and were optimized by each technology vendor on the applicable assay platform. IL-6 and TNFα are typically present at low pg/mL levels in human serum and represent the moderate abundance cytokines that can be detected by most commercial assays (Locksley et al., 2001; Wolf et al., 2014). IL-2 (Kovarova et al., 1989; Sharma et al., 2011) and IL-17a (Iwakura et al., 2008), on the other hand, are low abundance cytokines present at sub-pg/mL to low pg/mL levels in human serum and are well-suited for the evaluation of assay sensitivity limits for a given technology. The same set of methodically procured serum samples and controls were evaluated using the commercially available kits optimized by each vendor for the best performance. In addition, evaluations of these samples were performed at vendor sites when possible, in order to ensure that assays were run by experienced operators in a controlled environment. The evaluated assay technology platforms included both plate- and bead-based assay formats using single- or multi-plex modalities, integrated signal measurements, or single molecule counting, and included a wide variety of signal outputs ranging from fluorescence, ECL, and chromogenic detection to a novel acoustic detection system. The lessons learned from these cross-assay and cross-platform evaluations underscore the importance of performing such comparative evaluations for the quantitative assessments of biomarkers in immunoassays, especially, in the absence of reliable biomarker reference standards.

2. Materials and methods

2.1. Cytokine assays and technology platforms

Simoa (Quanterix), Erenna (Singulex), Biochip Array Technology (RANDOX), Ella (ProteinSimple), AMMP ViBE (Bioscale), and Imperacer (Chimera Biotec GmbH) were evaluated by the vendors at their respective laboratories using optimized protocols. Milliplex (Merck Millipore), V-PLEX (MSD), and High Sensitivity ELISA (eBioscience or R&D Systems) assays were evaluated at Biogen (Cambridge, MA) using vendor-specified protocols.

Four cytokine assays were evaluated: IL-2, IL-6, IL-17a, and TNFα using the commercially available kits optimized by each vendor. Cytokines were evaluated either in singleplex or multiplex assay formats. Singleplex assay kits: IL-2 (cat# 03-0051), IL-6 (cat# 03-0089), and IL-17a (cat# 03-0103) from Singulex; IL-2 (cat# 100,195), IL-6 (cat# 100,190), IL-17a (cat# 100,153), and TNFα, (cat#100,191) from Quanterix; IL-6 (cat# 75,062-0001), IL-17a (cat# 75,103-0001) from Bioscale; IL-17a (cat# K151RFD-1) from MSD; TNFα (cat# HSTA00D) from R&D Systems; IL-2 (cat# BMS221HS), IL-6 (cat# BMS213HS), and IL-17a (cat# BMS2017HS) from eBioscience; IL-2 (custom-developed kit for Biogen) from Imperacer; and IL-2, IL-6, IL-17a, and TNFα (cartridge custom-assembled for this study) from ProteinSimple. Multiplex kits for IL-2, IL-6, IL-17a, TNFα (cat# HSCYTMAG-60SK) from Merck Millipore; IL-2, IL-6, TNFα (cat# K15049D) from MSD; and IL-2, IL-6, IL-17a, and TNFα (cat# EV3623) from RANDOX. IL-17a, IL-6, and TNFα on Imperacer, TNFα on Erenna, and IL-2 and TNFα on AMMP ViBE were not evaluated either due to resource constraints or the lack of commercially available kits.

Recombinant cytokine proteins included within the commercial assays were used as calibrators and controls. In addition, the endogenous cytokine controls generated from stimulated PBMCs were included in each assay run.

A brief description of each technology and the evaluated cytokines are described in Table 1.

2.2. Samples

Serum samples were prepared by collecting whole blood samples from healthy donors and multiple sclerosis patients (MS) in Becton Dickinson Vacutainer SST Serum Separation Tubes (Franklin Lakes, NJ). Twenty samples each from healthy donors and MS patients were collected locally or purchased (Sanguine Biosciences, Sherman Oaks, CA). The tubes of whole blood were allowed to clot for 30–60 min at room temperature and then transported at 4 °C within 2–3 h to a centralized lab at Biogen for processing. To standardize the process for sample preparations across sites, a 2–3 h storage period at 4 °C following the clotting incubation was introduced for all blood sample collections and processing. Each whole blood tube was centrifuged at 1200 ×*g* at 4 °C for 10 min after the 2–3 h storage at 4 °C. The serum was aspirated from the collection tubes, aliquoted, and stored at −80 °C. All samples underwent only one freeze-thaw before testing on each technology platform.

2.3. Preparation of endogenous cytokine controls

Whole blood was collected in BD Vacutainer Sodium Heparin Tubes from healthy donors. The blood was centrifuged at 1200 x g for 10 min in a SepMate50 tube from StemCell Technologies (Vancouver, Canada) containing Ficoll-Paque solution (GE Healthcare, Wilmington, MA). The peripheral blood mononuclear cells (PBMC) were collected and washed twice using PBS. Stimulation of cytokine production was performed by incubation of PMBCs (2×10^6 cells/mL) overnight at 37 °C in pooled human serum (Bioreclamation, Westbury, NY) containing 2% phytohemagglutinin (PHA) (GE Healthcare, Wilmington, MA), 0.5 μg/mL ionomycin (Sigma, St Louis, MO) and 50 ng/mL phorbol myristate acetate (PMA) (Fisher Scientific, Waltham, MA). Cell supernatant containing cytokines was collected and aliquoted for storage at −80 °C following centrifugation at 1200 ×*g* for 20 min. The levels of cytokines in the cell supernatant were determined using the V-plex (MSD) assays resulting in 14,646, 346, 560, and 4,361 pg/mL concentrations for the IL-2, IL-6, IL-17a, and TNFα cytokines, respectively. These concentrations were used as a guide for preparations of the endogenous quality control (EQC), high-spiked, and low-spiked samples approximating the midrange, high, and low levels of analyte for each assay, respectively. The supernatant from the stimulated PBMC culture was custom-diluted at Biogen for each assay in a pooled human serum and shipped to the vendors. The resulting levels of spiked cytokines without accounting for preexisting endogenous cytokine levels in the pooled human serum are shown in Table 2.

2.4. Evaluation parameters

2.4.1. Assay precision (inter-assay and intra-assay)

Assay precision was assessed with three independent runs on different days using low- and high-spiked samples (Table 2). Calculations of intra- and inter-assay precision were performed using the published ANOVA method (DeSilva et al., 2003).

2.4.2. Analytical sensitivity

Analytical assay sensitivity on each platform was estimated from the performance of the standard curve, which was generated using a vendor-provided recombinant protein spiked in buffer. The estimated