Validation and use of a multiplex assay for the measurement of cytokine concentrations in non-human primate serum

A THESIS
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Sarah Gresch

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Dr. Rebecca Davies, Dr. Michael Murtaugh

November 2014

Acknowledgements

I have been lucky to have committee members who honor the scientific field with a commitment to high quality work. My thesis and, more importantly, my work as a scientist, is better because of their guidance. I would like to thank Dr. Rebecca Davies, my advisor, supervisor, and mentor, for leading by example. She is always willing and excited to share her knowledge and positive attitude. She has been an advocate for my professional progress for years. Thank you for having faith in my abilities, especially when I didn't. I would like to thank Dr. Melanie Graham for her infectious enthusiasm and immensely helpful discussions. Your commitment to good research and continuous improvement sets a high bar that I will keep striving for. I would also like to show my appreciation to Dr. Michael Murtaugh for his excellent practical and scientific advice. I am consistently amazed by your depth and breadth of knowledge; I am lucky to have had the opportunity to learn from you.

My thanks to Dr. James Collins for providing a work environment that encourages professional and scientific advancement. Thank you to Mrs. Sharon Hardy who introduced me to the concept of laboratory validation. I am very lucky that I had a teacher like her with patience and a wealth of knowledge and experience. To Ms. Suzanne Stone for her immense contributions to the lab and constant willingness to help, thank you. To the staff in the Preclinical Research Center, thank you for being so easy to work with and for answering my questions. Thank you to the members of the Murtaugh Lab who have been so kind and inclusive, that means a lot to me.

I would like to thank the students, staff and faculty of the College of Veterinary Medicine, particularly Lisa Hubinger, for their support and encouragement throughout this process.

Thank you to my family and friends. Your words of encouragement are truly appreciated and you all made this process so much more fun. To Steve, one of my most vocal supporters, thank you for being there for me and making sure I know how proud you are of me. To mom, your blind, unwavering, enthusiastic belief in me is both frightening and comforting; thank you for being my very first supporter.

Finally, to my husband, Dave, for being so wonderful and for always being on my side. Thank you for asking how every day went, for listening to every presentation, for proofreading, for taking the time to understand what I was talking about, for rubbing my feet and for working so hard for our future. You are my rock and my best friend and I could not have done this without you.

Abstract

Because of the complexity of the cytokine network, an individual cytokine measurement may be difficult to interpret or may provide insufficient information to inform conclusions. Multiplexing technology, which allows for the simultaneous measurement of multiple analytes, has advanced the study of cytokine secretion patterns. However, if cytokines are to be measured and used to make conclusions on research studies, we must understand the variability around the measurement and the potential for errors associated with testing. Method validation procedures generate data that reflect assay performance and identify the inherent measurement uncertainty, allowing data to be accurately interpreted. This study included the design and performance of a multiplex assay validation assessment for the simultaneous measurement of 23 cytokines in non-human primate (NHP) serum. The validation included analysis of serum samples from 36 healthy cynomolgus macaques in order to determine if the method could be used to detect cytokine concentrations in healthy NHP. Although sensitivity and precision data were generally acceptable, recovery and linearity data were highly variable. Of the 23 cytokines tested, 11 met acceptance criteria, 5 were marginal and 7 were rejected. The ability of the method to detect cytokine concentrations was demonstrated and cytokine concentration ranges were presented for 15 cytokines as measured in serum from healthy cynomolgus macaques. If the validation studies had not been performed, false conclusions could be made (overestimated or underestimated cytokine values) contributing to errors within the Total Testing Process. It would be appropriate to continue to evaluate the impact of pre-analytical error (e.g. storage time and temperature,

blood collection steps) on the measurement of cytokines in serum from cynomolgus macaques. Because this work demonstrates significant variability among measurements within and between animals, other important next steps include the description of typical variation levels within and between animals, followed by the generation of appropriate (population- or subject-based) reference intervals for healthy animals. It is clear after performing this validation study, that clinical and research investment in multi-analyte methods is not warranted without demonstration of method validity for each analyte of interest.

Table of Contents

List of Tables

List of Figures

Introduction

Cytokines

Cytokines are a complex network of regulatory proteins that contribute to cell growth, differentiation and signaling and they play important roles in embryonic development [1], normal physiological function, immune response and tissue repair. As of 2012 there were about 250 proteins categorized as cytokines and it is likely that more will be discovered [2]. Cytokines are secreted from white blood cells and various other cells (e.g. thymic, endothelial) throughout the body at concentrations ranging from ng/mL to pg/mL. These minute levels of cytokines often act in an autocrine or paracrine fashion, exerting their effects locally, rather than systemically. Cytokine actions may be inhibitory or stimulatory. They can be redundant, i.e. two distinct cytokines may act on the same cell to produce the same effect, or they may be pleiotropic, i.e. one cytokine may act on multiple cells to produce different effects [3]. The action a cytokine exerts depends on the environment where it is released, as some cytokines display synergistic or antagonistic effects depending on the presence of other cytokines [4]. This complexity makes it difficult to predict the effects of individual or associated cytokines.

The current cytokine field developed as a merging of four independent areas of research associated with lymphokines, interferons, colony-stimulating factors and growth factors. Lymphokines are cytokines secreted by lymphocytes; they were first discovered in the 1960s when it was observed that lymphocyte-secreted products could affect the growth

and function of other leukocytes [5]. Interferons were first recognized for their ability to act in an antiviral manner in the 1950s, and later it was determined that they were also able to affect cell growth and differentiation [5]. Colony-stimulating factors are proteins that stimulate differentiation of hematopoietic stem cells but can also affect fully differentiated cells. Finally, as the name suggests, growth factors are necessary for the normal growth and development of non-hematopoietic cells [5] and they can aid in chemotaxis, cell proliferation and synthesis of other cytokines. In 1965 it was noted that an antiviral agent was secreted by a lymphocyte which would have made the antiviral agent both an interferon and a lymphokine. As similar connections were observed, the distinctions between these four areas began to break down. At the time, the term lymphokine was commonly applied to many regulatory proteins, even when they were not secreted by lymphocytes, so the term "cytokine" was proposed in 1974 to encompass the expanding field [6]. Since 1960, this field has grown every year until 2011; a search of "cytokines," as the medical subject heading in PubMed shows one result for 1960 followed by a steady increase to almost 29,000 results for 2012. This same search resulted in about 27,000 hits for 2013, suggesting that research in the cytokine field may have reached a plateau and is waning.

Cytokines have been referred to as the 'hormones of the immune system' because cytokines, like hormones, function as biological messengers. However, there are differences. The majority of cytokines are not constitutively produced, but rather are stimulated in the event of a host challenge (presence of tumor cells, bacterial or viral

infection, tissue injury) [5]. In addition, cytokines typically function in an autocrine or paracrine manner, exerting their effects locally, often making them undetectable in circulation. Many hormones function in an endocrine manner and most can be detected systemically. Despite these differences, it is clear that cytokines and hormones are becoming more associated by mechanism and function, and some molecules classically considered hormones, like growth hormone, prolactin and leptin, are considered by some to be cytokines, in spite of their names or historical categorization [7]. Examples of cytokine/endocrine connectivity include: the cytokine interleukin- (IL) 1beta induces adrenocorticotropic hormone (ACTH) release [8], ACTH has been shown to induce the expression of IL-18 from adrenal cortex cells [9] and the pituitary has been shown to be important to immune function [10]. The hormone resistin is secreted mainly by peripheral blood mononuclear cells [11] and is involved in inflammation [12], endowing it with cytokine-like properties.

In this thesis, a commercial analytical method for simultaneously determining the concentration of 23 individual cytokines in NHP serum was evaluated to determine if it met pre-established method validation acceptance criteria. These 23 cytokines are of interest because they are involved in inflammatory and immune response pathways and, when taken together, can signify changes in immune status. A brief overview of each of the 23 cytokines analyzed in this thesis is provided in Appendix A

Methods for Measuring Cytokines

Cytokines are often measured in samples from healthy and diseased laboratory animals to evaluate their potential as biomarkers for dysfunction. Like hormones, cytokines are frequently measured in serum or plasma using enzyme-linked immunosorbent assays (ELISA) [13–17] . This method provides acceptable sensitivity, rapid turn-around time, safety and convenience (radio-isotopes are not required) and affordability. The ELISA method has been used for cytokine analysis in samples from cows [18], rats [19], dogs [20], humans and cynomolgus macaques [21].

ELISA components differ by manufacturer, and results may be affected by antibody characteristics and specific manufacturer reagents [22]. Technical competency is also important to obtaining accurate results. Results determined using ELISA may differ when compared to results generated using radioimmunoassay (RIA) or chemiluminescence. Therefore, it can be challenging to compare results generated by different people using different methods at different facilities on different sample populations. Recognizing and understanding this variability is critical for interpreting (and comparing) data appropriately.

One way to establish the accuracy, reliability and comparability of data generated under varying conditions is to evaluate method validation characteristics determined to identify measurement uncertainty. When a method validation data summary is provided with the research results it is possible to understand the inherent error around the measurement

and a comparison of methods becomes much more straightforward. Method validation procedures for ELISA include the demonstration of analytical sensitivity, linearity, precision (intra- and inter-assay) and accuracy (the ability to recover spiked analyte). For example, in one report, an ELISA method for use with samples from human and cynomolgus macaques, was demonstrated to be accurate for the measurement of cytokines based on the demonstration of dilutional linearity, precision of <20% coefficient of variation (CV) and accuracy in the 80-120% range [116].

Other techniques used less frequently in the past to measure cytokine concentration in serum or plasma have included radioimmunoassay [23], chemiluminescence [24], gene expression assay [25] and two multiplexing techniques, sandwich ELISA multiplex (e.g. SearchLight) [26] and bead based multiplex (e.g. Luminex) [27]. Multiplexing technologies feature the simultaneous measurement of concentrations of multiple analytes within a single sample. This technology is of interest for use in monitoring cytokines because individual cytokine results are frequently considered in relation to other cytokines as a profile of associated values. Due to the complexity of the cytokine network, an individual cytokine measurement may be difficult to interpret or may provide insufficient information to inform conclusions. For these reasons, cytokine panels are one of the most popular commercial multiplexed panels available and they are being used by researchers in both human and animal medicine [28–31].

Multiplexing has clear advantages over traditional ELISA methods. It provides the opportunity to save time, sample volume and reagents because many analytes can be measured at once, commonly using only 25 µL of sample. Cytokine multiplex methods provide improved sensitivity, generating results in the pg/ml range, rather than the ng/ml range achieved using the ELISA method; this is especially important when measuring systemic cytokines as the circulating levels are usually very low even if relatively high levels are present in a specific microenvironment [32].

Bead based multiplexing technology is based on the principles of flow cytometry, whereby particles are suspended in a fluid and flow past a laser used to discriminate among differences [33]. This technology uses microspheres, known as "beads," which are filled with a mixture of red and infrared fluorophores (a chemiluminescent compound) to create a spectral signature unique to each analyte. The outer surface of each bead is coated with antibodies specific for one analyte. After antibody capture, a reporter (streptavidin phycoerythrin) is added to the sample wells which will allow for luminescence of all the bound analyte. The beads flow through the instrument single-file and are first detected by a red laser (635 nm) which excites the internal dyes and classifies the bead by individual analyte. The beads continue past a green laser (532 nm) which excites the reporter attached to the analyte and the mean fluorescence intensity (MFI) for each bead is measured and averaged with all other beads for the same analyte. The beads initially developed for this technology were polystyrene-based and a specific filter-bottom 96-well plate had to be used in conjunction with a vacuum system for liquid

removal. However, magnetic-bead technologies which consist of beads with a superparamagnetic core have been recently developed. The new magnetic bead kits can be used with a hand-held magnet which temporarily attaches to the bottom of a solid 96-well plate, keeping the beads adhered to the bottom of the wells while the plate is washed. Magnetic beads provide for more effective washing/separation steps which result in less bead loss and improved replicate measurements as indicted by lower CV between replicates. With the use of magnetic beads in multiplex kits now becoming the norm, many companies are eliminating polystyrene bead based kits. Beads developed by the Luminex Corporation are the most commonly used for multiplexing kits by various manufacturers and require a special Luminex instrument for detection; other non-Luminex (cytometric bead array) kits can be used with a traditional flow cytometer.

There are, of course, challenges in manufacturing and effectively using multiplex immunoassays. Reagent optimization will affect the quality of a multiplex assay and the production of the specific beads that present the individual antibodies can ultimately affect sample measurements. For example, the purity of the beads (free of biological contamination), and strength of the bond between the antibody and the bead surface can influence the dynamics of the assay. It has also been reported that 10% to 32% of measurement variability could be due to variations in bead diameter during manufacturing [34]. In addition, the pH and ionic strength of buffers can affect protein structure, thus changing the ability of the protein analyte to bind to the appropriate antibody. Optimizing a kit with so many diverse analytes presents challenges as it can be

difficult to select appropriate buffers that will not alter any of the associated analytes present in the multiplex panel [35]. Likewise the optimal incubation times, temperatures and reagent concentrations must be appropriate for all analytes in the panel. Multiplexing is also more prone to cross-reactivity since multiple antibodies to several, and possibly similar, analytes are present within a single test well [36]. For example, IL-2 and IL-15 are structurally similar cytokines [37] and since all reactions occur in the same well during a multiplexed assay, the antibody for IL-2 may also bind to IL-15 molecules, causing the concentration results for IL-2 to be inaccurately elevated. Additional variability in results may occur as a result of the rapid degradation of cytokines in blood and the need for highly sensitive assays in order to measure cytokines that are present (and perhaps biologically relevant) at very low concentrations (picomole range).

Many cytokines are not expected to be present in elevated concentrations in healthy animals (because they are inflammatory and/or involved in the immune response pathways), therefore, representative data reflecting normal patterns of secretion are not available. As a result, some of the test validation procedures are harder to plan and implement (for example if endogenous samples with variable concentrations are not available to combine, it is difficult to create serum pools that can be used as quality control samples that cover the critical range of the standard curve.)

An understanding of the uncertainty that is due to the method as well as the inherent variability within a healthy animal is required before changes in cytokine concentration can be considered significant or informative during dysfunction or disease.

Comparison of Methods: ELISA and Multiplex

Because previously published cytokine data were generated using ELISA, multiplex technology has been compared to ELISA methods in order to assess whether the methods provide similar results with similar conclusions. In 2003, a comparison between a 15-cytokine multiplex assay developed in-house and several individual ELISAs for each of the same 15 cytokines was reported [36]. The multiplex assay was constructed with reagents from 6 different manufacturers and the 13 ELISAs came from 5 manufacturers. The cytokine measurements were found to be strongly correlated by method (correlation coefficients ranging from 0.75 to 0.99). In contrast, Liu et al. reported poor correlation between multiplex and ELISA methods when comparing measurement of human IL-6, IL-8 and TNFα (correlation coefficients ranging from 0.107 to 0.318) [38]. In 2004, Khan et al. compared four different manufacturers multiplex kits (5-plex) to a single ELISA designed for use with serum from humans. They found the concentrations obtained from the ELISA were only similar to the results from the multiplex kit from the same manufacturer, presumably because the same antibody was used in both kits [39]. In 2006, a National Institute of Health (NIH) study group evaluated the performance of ELISAs in comparison to multiplexed assays by performing a comprehensive literature review. Not surprisingly, they reported that the highest correlations were found when

comparing kits from the same manufacturer, as these share a common, proprietary antibody pair [40]. This group also reported that even in the cases of good correlation, the absolute concentrations observed form the two methods generally differ, while the trends remain the same. A study published in 2004 by the Eli Lilly Research Laboratories supports this finding [41]. That group also reported acceptable correlation (concordance correlation coefficient = 85%) between an in-house human five-plex cytokine multiplex assay and ELISAs for one of the cytokines (IL-6). They noted, however, that the multiplex results were, on average, 2.36-fold higher than the ELISA results. Other investigators have compared the two different methods with mixed results: Richens et al. reported good correlation (R^2 values 0.98-0.99) between multiplex kits and an ELISA for the measurement of IL-8 in human blood/serum [42]. Dossus et al. reported highly variable correlations (correlation coefficients from 0.37 – 0.92) when comparing measured concentrations from a multiplex assay to results from ELISA kits for the measurement of IL-6, TNFα, IL-1β, IL-1ra, CRP and sCD40L in paired serum and plasma samples from human patients [43].

This variability in results generated using different methodologies is not unusual and several factors may contribute to it. For instance, proprietary antibodies, reagents and standard matrices from different manufacturers, as well as different sample handling conditions, may influence results and contribute to the range of correlation reported. In addition, Breen et al. have reported that even the same method may generate different results when performed in different locations [44]. In that study, four different labs used

the same three human cytokine multiplex assays to measure paired sets of 18 serum and plasma samples. Although the observed concentrations showed similar trends between labs, the absolute results obtained, even when using the same kits, were highly significantly different ($P < 0.001$) for all kits. The inconsistent correlation between the two methods discussed above should be considered when selecting multiplex technologies for cytokine concentration measurement because it is clear that results cannot always be related to previous data collected using a different analytical method. Investigators need to consider sources of variability and error and establish strategies to account for the influence of these factors. One way to account for this variability and potential error is to consider 'the total testing process, TTP'.

The Total Testing Process

George Lundberg first introduced the concept of the TTP in 1981 when he outlined a series of processes, starting with the clinical question under consideration by the clinician, progressing to specific test selection, sample collection, transport to the laboratory, analysis, reporting back to the clinician, and final interpretation and decision making by the clinician [45]. Lundberg identified the phases associated with these processes as pre-analytical, analytical and post-analytical. Not surprisingly, there are opportunities for error to occur at each phase. Some errors can occur in all phases (e.g. transcription error) while others are unique to the phase of work being performed. The potential for error to occur during the pre-analytical, analytical and post-analytical phases

is present in any laboratory or research investigation that involves sample collection, analysis and interpretation of data resulting from analytical methods.

Pre-analytical errors are those that occur prior to performing the analytical test in the laboratory. They may occur as a result of inappropriate test selection, sample collection, sample handling, sample delivery or receiving within the laboratory. For example, some lithium heparin blood collection tubes were found to contain endotoxin which caused IL-1β and IL-6 to be stimulated [46]. Other examples of pre-analytical errors that may affect cytokine measurement include the time elapsed between blood collection and centrifugation [47], specimen storage time and temperature, and the choice of blood collection tube (plasma vs. serum) [22].

Analytical errors are those that occur at the time the analytical procedure is being performed. Some analytical errors can be reduced by establishing and adhering to standard operating procedures; such errors may include using improper reagents, cross-contamination during sampling, using incorrect procedures and equipment malfunction. A critical procedure for mitigating analytical error is to plan and perform a method validation strategy to identify the error routinely present within the method. Additional examples of analytical errors include the use of an antibody that is not specific for the analyte being tested or the presence of interfering molecules in the sample matrix (such as autoantibodies or binding proteins).

Lastly, post-analytical errors can also affect data reliability. Examples of post-analytical errors include data upload error, incorrect data analysis, data transcription error, use of inappropriate reference intervals, and inappropriate interpretation of results.

The potential for, and frequency of, errors throughout the TTP have been monitored in clinical and diagnostic laboratories for both human and veterinary medicine. In both human and veterinary diagnostic laboratories, the majority of errors occur in the pre-analytical phase (46-77% of total errors), followed in frequency by the post-analytical phase (7-46%). The lowest error rate occurs during the analytical phase of the TTP (7-21%) [48], [49]. While this is encouraging to analysts, it is important to realize that the analytical errors, while less frequent, will likely have the largest negative impact on patient care.

The frequency of errors in the analytical phase is low primarily because technologies and equipment have been standardized, quality control procedures have been incorporated into analytical procedures, and equipment and personnel management systems exist to improve consistency and reliability within laboratories. Many service (but few research) laboratories incorporate quality assurance (QA) programs which monitor laboratory error as a way to ensure continuous improvement. These programs integrate document control processes (documentation of and adherence to specific procedures), and reagent, equipment and personnel management (reagent inventories, equipment maintenance and calibration records, and training and competency review and documentation). In

addition, the potential for analytical errors is reduced by a adhering to a defined Quality Control (QC) strategy (tracking and trending of QC data) to verify method performance over time as a requirement for method validation to establish the reliability and accuracy ('fit for purpose') of laboratory methods.

Unfortunately, an understanding of the TTP and the associated potential for errors using analytical methods is frequently overlooked in research. For instance, most published studies that provide cytokine concentration data and interpretation do not include data describing the validation parameters of the method. For example, a group from the Harvard Medical School recently used an ELISA to measure human IL-2 and IL-15 in samples from cynomolgus macaques, but they do not mention how they (or others) evaluated the assay before use [50]. Similarly, investigators from Japan recently analyzed IL-2 and IFNγ in samples from cynomolgus macaques using both an ELISA and a Millipore multiplex kit. Data were not presented to demonstrate the reliability of either method [51]. In 2009, investigators submitted serum from cynomolgus macaques to a commercial testing service for the measurement of 5 cytokines using the Linco multiplex technique. [52]. It is likely that Linco had performed validation studies; however, validation data are not summarized or referenced for the reader. It cannot be assumed that a commercial method has been fully validated unless the data are presented for review within their kit insert or as part of a publication. Without such information, it is difficult to establish the reliability of the data as presented.

The intent of validation studies is to ensure a test is “fit for purpose” [53], meaning that experimental validation data must prove the assay performance characteristics are acceptable for its intended purpose. The intended purpose of an assay can vary based on individual study goals, for instance analytical data may be used to establish diagnostic cut-off values for disease surveillance, generate reference intervals, or establish normal or abnormal patterns of secretion. Important decisions are made based on interpretation of small changes in laboratory data (for example, research inclusion and exclusion criteria, and clinical and diagnostic cut offs). Therefore it is imperative to understand the potential for error around a given measurement and to describe how the accuracy and reliability of those measurements have been determined.

In veterinary medicine and research, commercially available assay kits designed for use for human diagnostic testing are frequently used to analyze samples from other species. In these cases, it is especially important to evaluate the method to make sure that the analytic characteristics of the assay are not affected by species differences such as matrix variability. The matrix of the blood contains different proteins and these proteins will vary by species; this variability creates the potential for interference (and subsequent error) with assay methodology.

An increasing appreciation for the importance of validation data may be forthcoming. The US National Institutes of Health (NIH) are considering strategies that include additional requirements for experimental validations within grant applications to address

the recent and consistent findings that many published results cannot be replicated [54]. Research proposals that include clearly described validation strategies may be seen as more competitive than those that do not address method reliability and consistency.

Multiplexed Human Cytokine Validations

Validation reports supporting the reliability of methods for the measurement of cytokines in NHP serum are rare. However, previous studies on the validation of human cytokine multiplex assays have been published and show variable results. The Eli Lilly study mentioned previously reported acceptable results for their in-house, human, five-plex cytokine assay (IL-1β, TNFα, IL-6, IL-8 and IL-10) with inter-assay precision <25% and recoveries of 83% - 125%, for all cytokines [41]. A study published in 2008 compared the accuracy of three commercially available human cytokine panels and found that 21%, 67% and 78% of samples tested for accuracy (from BioRad, R&D Systems and Linco respectively) met their acceptance criteria (70% - 130% of expected values) [55]. Prabhakar et al reported that five of six cytokines tested for accuracy (using a Linco multiplex kit) fell within their acceptance range of 80% - 120%, with only TNFα failing to meet this criterion [56]. Another study from 2008 assessed the reproducibility of a Millipore human cytokine assay and found inter-assay precision (n=3) to range from 18% to 44% for ten cytokines [3], while Prabhakar et al reported inter-assay precision of <25% for all six cytokines tested and intra-assay precision ranging from 9-21% (n=5) [56]. In 2009 Chowdhury, Williams and Johnson published a study including the

validation of a 7-plex human cytokine multiplex kit from Biosource. In that study intra-assay precision was below 20% (n=10) for all seven cytokines and intra-assay precision was found to be highly variable with a range from 3% to 80%. In the same study, accuracy ranges (percent of expected to observed concentrations in spiked samples) were found between 38% and 84% [57]. Although this study resulted in highly variable results, the authors concluded it was still a valid method after considering several factors (limit of detection, precision, accuracy) and the tradeoffs between assay efficiency and accuracy. The examples above show the degree of variability possible among validation results, independent of manufacturer (both in-house and commercially available are represented), location, or institution (academic, government and industry are represented). In addition to the range of results reported in these publications, an array of various acceptance criteria and validation methods were also presented.

Validation Strategy

A validation protocol is planned to evaluate the routine error around a measurement [58]. Two types of error can be considered when planning assay validation procedures: 1. measurement error, or the variability around a measurement; 2. method error, or the limitations of the method. Sensitivity and precision determinations will identify measurement error and the assessment of linearity and accuracy will identify method error. A specificity study, to understand the ability to differentiate analytes in the presence of other (possibly similar) components expected in the sample, also helps define

the method error. Understanding these errors, and the limits they infer, can prevent incorrect data interpretation and the associated negative effects on research conclusions, clinical trials and patient therapies.

One of the first steps in defining a validation protocol is to identify the acceptance criteria the method must meet in order to be approved as being 'fit for purpose'. These criteria should be established before beginning the method evaluation to minimize bias and to avoid the tendency to fit the criteria to the generated data [53]. The acceptance criteria may be based on previous data using other well characterized or 'gold standard' methods.

In 2013, the United States Food and Drug Administration (FDA) released an updated comprehensive draft guidance document describing steps and recommendations for acceptance criteria for the validation of ligand-binding assays [59]. According to this document, the following seven parameters may be included in a full validation of an assay: sensitivity, precision, accuracy, recovery, selectivity (specificity), reproducibility, and stability of analyte (robustness). The FDA defines sensitivity as "the lowest analyte concentration that can be measured with acceptable accuracy and precision." The variability around these very low measurements is expected to be greater than at other points along the standard curve and it therefore gives slightly more relaxed acceptance criteria. For example, this guidance suggests that all standards should be within 20% of their expected concentrations except at the lower limit of quantitation (LLOQ), where 25% is acceptable. The second parameter, precision, is the "closeness of individual

measures of an analyte when the procedure is applied repeatedly to multiple aliquots of a single homogeneous volume of biological matrix." The recommendation is to perform at least five determinations per concentration and a minimum of three concentrations in the expected range of sample concentrations, with acceptance levels of 20% CV or less, except at the LLOQ when 25% CV is acceptable. Precision assessment should be performed both within and between assays. As part of the between-assay precision evaluation, incurred sample reanalysis should be performed. To do so, a selected subset of samples (i.e. in biological matrix and not QC pools created in buffer) should be analyzed on separate runs on different days to support precision and accuracy findings. The next parameter, accuracy, is the degree of closeness of the observed value to a "known true value under prescribed conditions." Obtaining a sample of known value (such as a traceable standard) can be difficult when working in non-human species and the FDA guidance suggests confirmation with a validated reference method such as liquid chromatography-mass spectrometry (LC-MS), if possible. It is recommended that accuracy be evaluated the same way as precision but with the use of a known-true (traceable to a standard) sample. The next validation parameter to be tested is recovery and it is the "the measured concentration relative to the known amount added to the matrix." Although the guidance document only recommends that recovery be performed in ligand-binding assays that use sample extraction, when using a known-true sample, it may also serve as an accuracy assessment experiment. The fifth parameter is selectivity/specificity which is described as "the ability of the bioanalytical method to measure and differentiate the analytes in the presence of components that may be

expected to be present. These could include metabolites, impurities, degradants, or matrix components." Selectivity/ specificity addresses two issues of possible interference: 1) from substances structurally similar to the analyte of interest, and 2) from matrix effects. To understand interference, the structurally similar substances should be added in (or 'spiked') at known concentrations, and measured to assess agreement to the known-true value; when possible, this should be confirmed with a validated reference method (like LC-MS). Matrix effects may be studied three ways: 1) by linearity (parallelism) whereby a dilution series from a sample is compared to a dilution series of standards, 2) by comparing the standard curve in buffer to the standard curve in biological fluids and 3) by determining the 'non-specific binding' of the assay. Non-specific binding refers to the amount of signal (radioactivity, fluorescence, luminescence, etc.; used to determine analyte concentration in the sample) that is inherent to the test and is not specific to antibody binding. For example, trace amounts of radioactivity will adsorb to the inner walls of tubes that do not contain sample or antibody, and could affect the result of a radioimmunoassay, if not accounted for. Linearity assessments are also important if a sample with high concentrations must be diluted to be measured accurately. In those instances, not only is the analyte of interest being diluted but so are all of the possible interfering proteins in the matrix, which could result in increased binding of analyte and misleadingly high results [60]. The next parameter, reproducibility, is the "precision between two laboratories" but it also "represents precision of the method under the same operating conditions over a short period of time." Reproducibility can be evaluated by comparing validation results between two or more laboratories performing

the same protocol, or by a single lab repeatedly measuring the same quality control (QC) or study sample. In-house QC samples are prepared by pooling previously-tested samples of similar concentrations and establishing lab-specific acceptance ranges (commonly the mean ± two standard deviations) by testing these QC in each assay to verify that the method is working as expected. The final parameter, robustness, is an assessment of "the chemical stability of an analyte in a given matrix under specific conditions for given time intervals," the sample results should be within 15% of expected concentrations. Recommended analyses for robustness may include assessing freeze/thaw cycles, bench-top handling conditions, long-term storage, stock solution in different states or different buffers, and processed sample stability (time to analysis), if applicable.

The FDA guidance document is meant to provide best practice recommendations and all of the described steps are not always required to consider an assay validated. For instance, the authors suggest that for exploratory methods, less validation (perhaps only a single assay to determine the intra-assay precision and accuracy) may be sufficient [59]. However, this document recommends that a full validation be implemented when a new bioanalytical method is being established.

The multiplex method for measuring the concentration of multiple cytokines in a single sample creates an obvious complexity for planning validation protocols. The FDA guidance document states that "samples involving multiple analytes should not be

rejected based on the data from one analyte failing the acceptance criteria." This approach will become important as the cytokine multiplex method is assessed in this thesis.

Thesis Introduction

The goal of this thesis was to evaluate the performance of Millipore's Milliplex MAP Non-Human Primate (NHP) magnetic cytokine multiplex assay kit (catalog number PCYTMG-40K-PX23) to determine if it is valid or 'fit for purpose' for use as a method to accurately and reliably measure 23 different cytokines in serum from healthy cynomolgus macaques.

Methods verified for use in this species are important. When no suitable alternative methods or species are available, the NHP can add translational value because of its ability to closely mimic the biological complexity of humans. When this research was initiated, data in the literature that described the analytical validation performance of multiplex methods used for cytokine measurements in blood samples from cynomolgus macaques were not available. Previous studies evaluating cytokines in NHP blood using multiplex assays have been published [27], [31], [61], [62], however, these studies do not provide method validation results. A recently published paper by He, Narayanan and Fort compares the performance of three commercially available multiplex kits in cynomolgus macaque serum [63]. However the three kits tested were all polystyrene-

bead based; one of these kits has already been discontinued (R&D Systems), one is in the process of being replaced with a magnetic kit counterpart (Millipore) and the manufacturer of the third kit (Life Technologies) plans to transition all multiplex kits to magnetic format in the future. In addition, the study only reported accuracy measured in the standards (by comparing manufacturer reported values to observed values) but did not assess the accuracy measured in the sample matrix (NHP serum). Nor did they report intra- or inter-assay precision or linearity. The scarcity of validation data supporting the use of this method in the NHP is troubling since cytokine measurements are frequently performed and used to inform research strategy, outcomes and direction.

Hypothesis

The hypothesis of this thesis is that the commercially available Millipore Milliplex NHP magnetic cytokine multiplex assay kit is 'fit for purpose' as demonstrated by meeting pre-established validation criteria for the 23 cytokines being measured in NHP serum.

The specific aims of this study are:

1. To assess sensitivity, precision, recovery and linearity of a method used to measure concentrations of 23 cytokines in one sample aliquot to determine if the method meets pre-established validation acceptance criteria. This aim will determine if the method is *valid, accurate and reliable* for the measurement of these cytokine concentrations in cynomolgus macaque serum.

2. To determine if the assay is suitable for the detection of cytokine concentrations in serum from a study population of healthy cynomolgus macaques.

It is not an aim of this study to optimize or influence the production parameters of this commercial method. Instead, we are evaluating its use in our laboratory to define the error and limitations of the test and to determine if it is suitable for use in our study population.

Materials and Methods

Cytokine Assay Kits

Millipore MILLIPLEX® MAP magnetic bead, premixed 23-plex kits for Non-Human Primate Cytokines (cat. no. PCYTMG-40K-PX23) were used according to manufacturer instructions. This is a magnetic bead based assay kit with premixed beads for 23 cytokines: G-CSG, GM-CSF, IFNγ, IL-1β, IL-1ra, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12/23(p40), IL-13, IL-15, IL-17A, MCP-1, MIP-1α, MIP-1β, sCD40L, TGFα, TNFα, VEGF and IL-18. A standard curve with concentrations ranging from 20,000 pg/mL to 4.9 pg/mL for IL-4; 50,000 pg/mL to 12.2 pg/mL for IL-10 and IL-18; and 10,000 pg/mL to 2.4 pg/mL for the other 20 cytokines was made by serial dilution. All standards and controls were run in triplicate in each assay. The standards and controls for nine of the cytokines (IFNγ, IL-1β, IL-2, IL-4, IL-5, IL-12, IL-13, TNFα and IL-18) were recombinant NHP proteins, the standards and controls for the other 14 cytokines were human proteins, as shown in Appendix B. The detecting antibodies for 21 of the cytokines were generated against human proteins; detecting antibodies for IL-2 were generated against monkey (unspecified) protein while VEGF was generated against mouse protein (this data is summarized in Appendix B). The manufacturer reports “no or negligible” cross-reactivity between any of the analytes of the assay panel. A human serum matrix is provided with each assay kit to normalize the assay; this matrix is added to wells containing standards and controls which is meant to mimic the NHP matrix of the samples.

Validation Plan

Assay sensitivity can be defined two ways: 1) the lower limit of detection (LLOD) which is considered the smallest concentration that can be distinguished from zero. 2) The lower limit of quantification (LLOQ) which is the smallest concentration that can be accurately and precisely tested. The analysis software used for these studies (MILLIPLEX Analyst, version 3.4) determines the LLOD for each individual analyte in every assay performed. LLOD calculations are extrapolated based on the standard curve and the background readings for each analyte in each assay. Therefore the software may determine different LLODs for the same analyte over different assays. Extrapolation was done by extending the measured curve concentrations past the limits of the actual points measured. Therefore, to investigate the LLOD for each analyte using actual, measured data points, the lowest standard was diluted to 50% with assay buffer and tested in two wells in one assay then the lowest standard was diluted to 25% with assay buffer and tested in 6 wells in another assay. The percent of observed concentration was accepted (and deemed accurate) if it was between 75% - 125% of the expected concentration [64].

Assay precision was defined as the ability to repeat observed concentrations under the same conditions both within a single assay (intra-assay precision) and between assays (inter-assay precision). Intra-assay precision was evaluated at low concentration levels and high concentration levels, relative to the standard curve. To study precision at the

lower end of the standard curve, a serum pool was made by combining two samples from the same animal, collected four months apart, and spiking with a known amount of kit-provided standard near the middle of the curve range. To study precision at higher relative concentrations, a second serum pool was made by combining two samples from the same animal (but different than the animal used to make the first pool), taken two months apart, and spiking with a known amount of kit-provided standard at the highest concentration. Both of these pools were made such that the majority (>90%) of the final pool matrix consisted of endogenous NHP serum and the remaining portion consisted of assay kit standard. All samples underwent a freeze-thaw cycle and were then kept at ≤ -20^{o}C until the date of assay. The two pools used to study intra-assay precision were tested on a single assay plate in 12 wells each. Inter-assay (between assay) precision was determined using four method controls, tested in triplicate, on four separate plates. Two of the four controls used to study inter-assay precision were kit-provided controls, using the same lot number throughout the study. The other two controls were created by a mixture of a NHP serum pool (made from the serum of two healthy cynomolgus macaques) spiked with a consistent volume of one of the kit-provided controls; see Appendix B, Table 12. The CV (ratio of standard deviation/mean) was calculated from the measured concentration of each replicate. The pre-established acceptance criteria selected for a multiplexed method is CV <25% for inter-assay precision and <20% for intra-assay precision based on common published parameters [56], [57], [64].

As described above, accuracy and recovery are similar, with the major difference being the use of a “known-true” source for comparison. Although the standards provided with the assay kit are considered a “known-true” source, a confirmatory analysis (for example, HPLC) was not performed. Therefore, the analyses performed for this validation strategy are considered recovery and not accuracy. Recovery is then defined as the degree of closeness to the known true value. Recovery was measured as the percent of observed sample concentration relative to the expected concentration in a sample spiked with a known analyte concentration. Samples used to establish the recovery of the method were created by measuring cytokine concentrations in two NHP serum pools (constructed as described above for precision) with and without spiking of known amounts of two kit standards, resulting is 6 distinct samples: sample 1 neat; sample 1 + standard 5; sample 1 + standard 7; sample 2 neat; sample 2 + standard 5; sample 2 + standard 7; see Table 1. Each of these 6 samples was measured in triplicate and this entire procedure was repeated in 6 additional individual animals (for a total of 8 animals) over 3 assays. Each spiked sample was constructed such that the majority (75%) of each sample consisted of endogenous NHP serum and the remainder of kit standard. Recovery was calculated as a percent of observed to expected concentrations. See Table 2 for a summary of spiking concentrations and an example of calculations for a single cytokine (IL-15). For the multiplexed method, acceptance criteria for recovery were set to be 75 – 125% of expected values based on common parameters used by researchers in the field for this method [56], [57], [64].

Table 1: An example of how samples were constructed from 2 of the 8 animals used in recovery studies so that resulting solutions contained 75% endogenous material.

Sample ID	Endogenous Component	Formula	Testing
Sample 1 neat	Animal # 1	No additions; only serum from Animal # 1	Tested in 3 wells in a single assay
Sample 1 + standard 5	Animal # 1	60 µL serum from Animal # 1 + 20 µL standard 5	Tested in 3 wells in a single assay
Sample 1 + standard 7	Animal # 1	105 µL serum from Animal # 1 + 35 µL standard 7	Tested in 3 wells in a single assay
Sample 2 neat	Animal # 2	No additions; only serum from Animal # 2	Tested in 3 wells in a single assay
Sample 2 + standard 5	Animal # 2	60 µL serum from Animal # 2 + 20 µL standard 5	Tested in 3 wells in a single assay
Sample 2 + standard 7	Animal # 2	105 µL serum from Animal # 2 + 35 µL standard 7	Tested in 3 wells in a single assay

Table 2: Summary of spiking concentrations and an example of recovery calculations for a single cytokine (IL-15; 6 of 8 samples were within acceptable recovery ranges, 75% - 125 %, in this case).

IL-15					
		Added Analyte (pg/mL)	Recovery	Average per Animal	Result
Assay #1	Animal #1	2500	54%	53%	Reject
		625	53%		
	Animal #2	2500	113%	112%	Accept
		625	111%		
Assay #2	Animal #3	2500	88%	84%	Accept
		625	80%		
	Animal #4	2500	106%	102%	Accept
		625	99%		
	Sample #5 (pool)	2500	75%	66%	Reject
		625	57%		
Assay #3	Animal #6	2500	89%	86%	Accept
		625	82%		
	Animal #7	2500	83%	82%	Accept
		625	80%		
	Animal #8	156.25	82%	80%	Accept
		39.08	82%		
		9.75	77%		

An analysis of linearity is performed to determine whether serum matrix effects interfere with measured results and to confirm the working range (range of the standard curve) of the assay. Like recovery, linearity was also measured as a percent of observed sample concentration relative to the expected concentration in a sample diluted a known amount. The eight individual animal pools used in recovery studies were also used in linearity studies. Each of the eight pools that had been spiked with the highest kit standard were then diluted to 50% and 20% with kit-provided assay buffer and each was measured in triplicate (for example, "Sample 1 with standard 7" from Table 1 was diluted to 50% and 20% with assay buffer). The linearity of three of these eight samples was also tested at a 50% dilution from neat samples (i.e. a dilution was made from a serum aliquot that had not been previously spiked with the highest kit standard). Each sample tested for linearity studies was tested in triplicate. See Table 3 for a summary of the linearity studies setup and an example of calculations for a single cytokine (IL-15). For the multiplexed method the acceptance criteria for linearity is 75 – 125% of expected values, based on common acceptance values used in the field [56], [57], [64].

Table 3: Summary of linearity set-up and an example of linearity calculations for a single cytokine (IL-15; 7 of 8 samples were within acceptable ranges, 75% - 125 %, in this case).

IL-15					
		Dilution	**Linearity**	**Average per Animal**	**Result**
Assay #1	Animal #1	Spiked with standard 7 then diluted to 50%	142%	167.7	Reject
		Spiked with standard 7 then diluted to 20%	194%		
	Animal #2	Spiked with standard 7 then diluted to 50%	106%	115%	Accept
		Spiked with standard 7 then diluted to 20%	123%		
Assay #2	Animal #3	Neat, diluted to 50%	91%	117%	Accept
		Spiked with standard 7 then diluted to 50%	121%		
		Spiked with standard 7 then diluted to 20%	141%		
	Animal #4	Neat, diluted to 50%	74%	83%	Accept
		Spiked with standard 7 then diluted to 50%	88%		
		Spiked with standard 7 then diluted to 20%	88%		
	Sample #5 (pool)	Neat, diluted to 50%	undetectable	107%	Accept
		Spiked with standard 7 then diluted to 50%	103%		
		Spiked with standard 7 then diluted to 20%	111%		
Assay #3	Animal #6	Spiked with standard 7 then diluted to 50%	116%	123%	Accept
		Spiked with standard 7 then diluted to 20%	131%		
	Animal #7	Spiked with standard 7 then diluted to 50%	108%	116%	Accept
		Spiked with standard 7 then diluted to 20%	124%		
	Animal #8	Spiked with standard 7 then diluted to 50%	106%	110%	Accept
		Spiked with standard 7 then diluted to 20%	114%		

Specificity data as reported by Millipore in the assay kit insert were assumed to be accurate. Studies to confirm specificity of the antibodies for the 23 cytokines would

require complex protocols using multiple analytes that may not be traceable to known standards. Therefore, specificity was not evaluated in this study.

A comparison of the validation protocol used in this study with the recommended FDA guidelines is summarized in Table 4. The layout of each validation plate is presented in Appendix B, Figure 7.

Table 4: Comparison of cytokine multiplex validation protocol and FDA Industry Guidelines for method validation.

Parameter	**Definition and recommendations from the FDA Industry Guidance**	**How it was performed in this study**
Sensitivity	"The lowest analyte concentration that can be measured with acceptable accuracy and precision." (No specific recommendations are made on how to evaluate this.)	To test precision at low concentrations: the lowest standard was diluted to 50% with assay buffer and tested in two wells in one assay then the lowest standard was diluted to 25% with assay buffer and tested in 6 wells in another assay. In addition, the observed lowest standard concentration was compared to the expected concentrations.
Precision	"The closeness of individual measures of an analyte when the procedure is applied repeatedly to multiple aliquots of a single homogeneous volume of biological matrix." The recommendation is to perform at least five determinations per concentration and a minimum of three concentrations in the expected range of sample concentrations. Incurred sample reanalysis (using actual samples, not QCs) should be performed as part of inter-assay precisions. No specific instructions are provided on how many replicates are needed to evaluate inter-assay precision.	Two NHP serum pools were constructed (one at high concentrations, one at low concentrations) and each was tested in 12 wells on a single plate. For inter-assay evaluation, 2 kit provided QC and 2 pools of NHP serum spiked with kit QC were run in triplicate on 4 separate plates. The NHP serum spiked with QC is meant to satisfy the incurred sample reanalysis as it incorporates actual biological matrix but was spiked with QC to gives concentrations within the working range of the assay.
Accuracy	"The degree of closeness of the determined value to the nominal or known true value under prescribed conditions." Perform at least five determinations per concentration and a minimum of three concentrations in the expected range of sample concentrations using known true samples. Alternatively, a confirmatory method (like LC-MS) can be used to verify the samples evaluated.	Although the kit standards were considered "known true," traditional accuracy was not performed. Rather recovery studies were done. No other confirmatory methods were performed.
Recovery	"The measured concentration relative to the known amount added to the matrix." Generally only used for assays which require sample extraction. Specific number of replicates is not provided.	Although this study does not involve sample extraction, recovery was performed in lieu of traditional accuracy. Since the kit standards were considered our "known true" sample, they were spiked into endogenous NHP serum to mimic biological matrix conditions. The two pools described for precision as well the serum from six other individual NHPs were spiked with two known concentrations of kit standard and each was measured in triplicate.
Selectivity/	"The ability of the bioanalytical method to	Specificity for structurally similar

Specificity	measure and differentiate the analytes in the presence of components that may be expected to be present." To evaluate interference from the structurally similar substances, these substances should be obtained and spiked in at known amounts; when possible, this should be confirmed with a validated reference method (like LC-MS). Interference from matrix effects should be evaluated three ways: 1. by linearity (a dilution series from a sample is compared to a dilution series of standards) 2. By comparing the standard curve in buffer to the standard curve in biological fluids. 3. By determining the non-specific binding (NSB) of the assay.	analytes was performed by Millipore and not confirmed in our study due to the interaction complexity of 23 different analytes and the lack of traceable standards for all 23 analytes from cynomolgus macaques. Linearity was performed by diluting each of the 8 samples used in recovery studies to 50% and 20% with kit-provided assay buffer. The linearity of each of these 8 samples was also tested neat (diluted to 50% prior to any spiking with kit standard). Each sample was assayed in triplicate. The standard curve was not tested in buffer alone, kit-provided NHP matrix was always added to normalize the assay. The NSB of the assay was determined by testing buffer alone (in triplicate) in each assay.
Reproducibility	"The precision between two laboratories" and "the precision of the method under the same operating conditions over a short period of time." It can be evaluated by comparing validation results between two or more laboratories performing the same protocol, or by a single lab repeatedly measuring the same QC or study sample.	The validation results from Millipore were compared to our own. However, Millipore validated using cytokines from phytohaemagglutinin (PHA) or LPS stimulated peripheral blood mononuclear cells (PBMCs) from 2 cynomolgus macaques, except for IL-10, which was tested as LPS challenged serum; therefore validations are not directly comparable. No other lab validation data were found to compare to. Precision studies also provided reproducibility data.
Stability/ Robustness	"The chemical stability of an analyte in a given matrix under specific conditions for given time intervals." Can be studied by assessing freeze/thaw cycles, bench-top handling conditions, long-term storage, stock solution in different states or different buffers, and processed sample stability (time to analysis), if applicable.	Long-term storage will be evaluated in future studies. The NHP serum pool used for inter-assay QC was aliquoted and will remain stored at -20^{o}C for future analysis and comparison.

Measurement of Cytokine Concentrations Using Luminex Multiplex Technology

A Luminex100® system was used to perform the multiplexed assays. This system uses 6.5μM beads, which have a superparamagnetic center with a polystyrene coating. The plastic coating has a distinct spectral signature and the surface is coated with antibodies against distinct analytes [65]. The bead's spectral signature is achieved using a specific ratio of two dyes, red and infrared, containing fluorochromes, which fill the internal portion of each bead. As sample is mixed in plate wells with beads, the analyte of interest is captured by antibodies on the bead surface. A secondary antibody is added and finally a reporter (streptavidin phycoerythrin), that produces orange fluorescence, is added. As these beads flow past a red laser the internal dyes are excited and the bead is categorized based on the analyte it captures. The bead continues to flow past a second, green, laser which excites the reporter and the amount of fluorescence is measured [66]. The level of fluorescence correlates to the analyte concentration in the test sample. Luminex xPONENT Software version 3.1 was used for Luminex operation.

Due to a manufacturer change in kit format (transition from polystyrene to magnetic beads) all measurements of cytokines in serum from NHP associated with Aim 2 were generated using a kit that differed by bead-type (and therefore separation technique) when compared to the measurements generated for the validation studies. The NHP cytokine data were generated using the polystyrene-bead based kit (catalog number MPXPRCYTO-40K-PX23), which was discontinued. An additional difference

associated with the method transition is a change in the ability to measure IL-10. Millipore reported that IL-10 was not detectable in lipopolysaccharide (LPS)-stimulated peripheral blood mononuclear (PBMC) cells from cynomolgus macaques in the polystyrene-bead based kit. When the IL-10 antibody pair was changed upon transition to the magnetic bead format, IL-10 was then considered "strongly reacting," when tested in cynomolgus macaque serum using this assay kit. The manufacturer determined "reactivity" (specificity) for IL-10 using LPS-challenged serum from two cynomolgus macaques. Reactivity for the other 22 cytokines was determined by testing the secreted products of LPS- or phytohemagglutinin-challenged PBMCs from two cynomolgus macaques (e.g. LPS-stimulated PBMCs from cynomolgus macaques generated IL-6, which then strongly reacted with the antibodies used in this kit). A Technical Applications Scientist at Millipore indicated (personal communication) that no other antibody pairs were changed in the transition to magnetic bead technology [67]. However, the standard curve range did change due to adjustments incorporated in the transition. Prior to the integration of the magnetic bead technology, all 23 analytes had the same standard curve range of 0.64 – 10,000 pg/mL. After the transition to the magnetic kit, the standard curve range for IL-4 was 4.9 – 20,000 pg/mL, for IL-10 and IL-18 was 12.2 – 50,000 pg/mL and for the other 20 analytes it was 2.4 – 10,000 pg/mL. As a result of these changes, IL-10 data are not available for the NHPs (as they were all tested using the polystyrene-bead based kit). However, IL-10 validation data is presented (as all validation work was performed using the magnetic-bead based kit).

Serum samples from NHPs that were determined to contain undetectable concentrations of specific cytokines are recorded as containing <0.64 pg/mL as this was the lowest standard for the polystyrene-bead kits used for analyses. Undetectable concentration levels determined as part of the validation of the magnetic bead kit are reported in light of the individual cytokine sensitivity level determined.

Animals

All animal procedures were approved by the University of Minnesota Institutional Animal Care and Use Committee, were conducted in compliance with the Animal Welfare Act, and adhere to principles stated in the Guide for Care and Use of Laboratory Animals. Blood samples were collected from 36 healthy cynomolgus macaques (*Macaca fascicularis*) from a subcutaneous vascular access port [68] to allow for cooperative central venous access in the homecage. Semi-annual veterinary physical examinations were performed and included weight, body condition scoring, heart rate, temperature, palpation of lymph nodes and abdomen and evaluation of the oral cavity, dermis, ears and nose. Complete blood count and a chemistry panel were also performed. Animals were enrolled in a separate islet transplantation study so the exclusion criteria were pertinent to that study and included body weight under 1.8 kg, severe persistent diarrhea or vomiting, active systemic infection, lymphopenia, neutropenia, thrombocytopenia, elevated hemoglobin or liver enzymes or plasmodium infection. Blood (0.6 mL) was collected in Becton Dickinson serum separator microtainer tubes. Tubes were kept cool and allowed

to clot for 30 minutes then centrifuged at 3,500 RPM for 20 minutes. Samples were aliquoted into plastic microfuge tubes and frozen at $\leq$ -20^{o}C.

Of the 36 NHPs, two were female and 34 were male, all aged between 3.4 and 6.4 years (median 5.3 years), and weighed between 3.6 and 7.9 kg (median 5.1 kg). All animals were purpose-bred and purchased from institutionally approved commercial vendors. They were housed in pairs or small groups of the same sex. They had free access to water and were fed High-Protein Monkey Chow biscuits (Purina Mills, St Louis, MO, USA) based on body weight. Their diet was enriched with fresh fruits, vegetables, grains, beans, nuts, and a multivitamin preparation. The animals participated in an environmental enrichment program that included social play, opportunities for foraging, puzzle solving, music, and regularly scheduled access to exercise and swimming areas.

Data Analysis

After Luminex data acquisition, fluorescence data were analyzed with Millipore MILLIPLEX Analyst version 3.4. The Millipore MILLIPLEX Analyst settings were selected according to manufacturer recommendations and were described as: standard curve fit model using a five parameter log fit with threshold slope: 0.10; threshold R^2: 0.50; threshold curve fitting: 2.00; threshold replicate data: 2.00; reporting range low bead count: 35 and high CV: 20.00. Aside from the curve fit model, these settings simply alert the user to potentially poor data from the standard curve (poor slope, poor

coefficient of determination (R^2), for example) or within the sample data (replicates with low bead counts or high CV, for example), but do not affect the analysis.

The statistical analysis of the data for validation purposes was performed with the calculated concentrations of each sample rather than the MFI (mean fluorescence intensity) source data. Validation analysis with MFI numbers may show a falsely better assay performance. Calculations based on concentration values will show higher variation than analysis of MFI numbers as the differences in standard curves of each analyte must be considered. However, analysis using concentration gives a more accurate representation of assay performance since actual results are reported as a concentration (pg/mL) rather than as an MFI number.

The Tukey method (any value 1.5 times outside the inner-quartile range) was used to identify potential outliers. Although outliers were detected in the cytokine concentration distributions, they were not excluded. The data are presented as the full distribution of results because without specific criteria or justification for exclusion, all data may represent true values within a healthy animal. The cytokine data collected from the 36 monkeys were not used to generate a reference interval (RI) for the specific cytokines because best practice recommendations for the establishment of a *de novo* RI suggest that at least 120 reference individuals that have met specific inclusion criteria and categorization (age, sex, etc.) be used to generate a population-based RI [69]. The

objective of Aim 2 was simply to determine if the assay is suitable for the detection of baseline cytokine concentrations in serum from healthy cynomolgus macaques.

Cytokine concentrations were measured in 36 healthy cynomolgus macaques, sometimes multiple samples per animal, for a total of 54 samples tested. An average of each cytokine per animal was determined when more than one sample per animal was tested. If censored data (undetectable concentrations, listed as <0.64 pg/mL) were included in mean calculations, the data point was changed to half the limit of detection (i.e. <0.64 was changed to 0.32). If the resulting mean was below 0.64 pg/mL the mean was reported as "<0.64." Using half the limit of detection is a simple method used in place of more complex statistical methods for analyzing censored data [70] and is suggested as an acceptable method in this case as the differences of calculated averages that could have resulted from using 0, 0.32 or 0.64 pg/mL in analysis are miniscule.

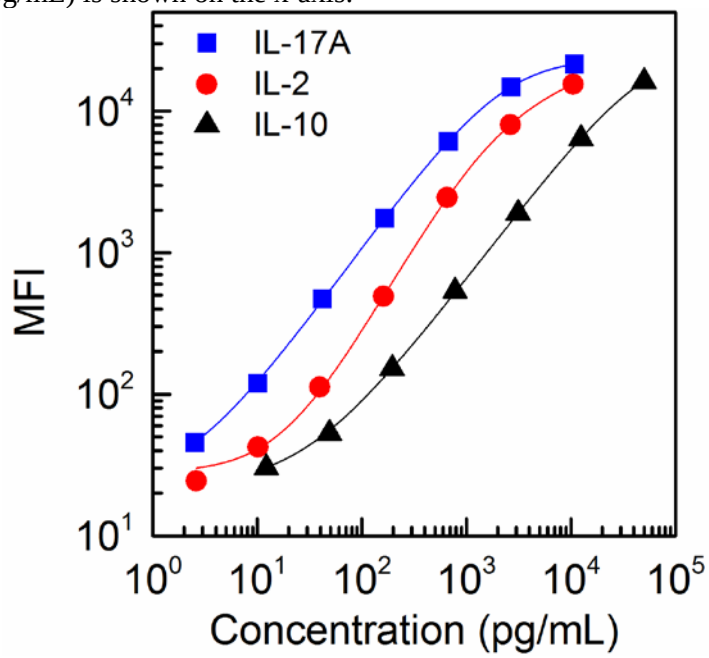
NHP cytokine distribution data will be presented as box-and-whisker plots. The top of the box represents the top 75^{th} percentile, the middle line is the median, the bottom of the box represents the bottom 25^{th} percentile and the whiskers extend to the minimum and maximum (excluding outliers); outliers are shown beyond the reach of the maximum whisker.

Results

Aim 1: Method Validation Data

Representative standard curves from validation data that were determined to be acceptable (IL-2), marginal (IL-10) or rejected (IL-17A) are shown in Figure 1. The recovery (percent of expected to observed concentration) for each point on each standard curve in each assay was determined. Occasionally there were standard curve points that fell outside of the acceptable 75-125% range (13 of 644 total standard curve points, 2%, over four assays) but trends were not identified (there were never more than three points of any curve outside of the expected range and trends of standard curve variability across multiple assays were not present.) All standard curves obtained in this study were acceptable and did not contribute to the rejection of validation acceptance criteria for any cytokine. Representative standard curves are shown in Figure 1 to illustrate the degree of uniformity between standard curves for cytokines from each validation outcome category (accepted, marginal, rejected).

Figure 1: Representative standard curves from an accepted (IL-2), marginal (IL-10), and rejected (IL-17A) cytokines. Mean fluorescence intensity (MFI) is shown on the y-axis while concentration (pg/mL) is shown on the x-axis.



Sensitivity

The results of sensitivity analyses are shown in Table 5. To briefly review terms: the LLOD is the lowest concentration that can be detected; the LLOQ is the lowest concentration that can be tested accurately (± 25% of the nominal value) and precisely (CV < 25%). Both the LLOD and LLOQ was determined for each cytokine, though only the LLOD was used to determine if validation acceptance criteria were met (LLOD ≤ the lowest standard). The lowest standard was diluted to 50% and 25% to determine first the LLOD (if a concentration could be detected) and second the LLOQ (if the observed value was both precise and accurate). If the LLOD did not match the LLOQ, the lowest standard of each cytokine in each assay was reviewed determine precision and recovery

values, and if conditions were met, the lowest standard was set as the LLOQ. With the exception of VEGF, the lowest standard for each cytokine consistently showed acceptable recovery and precision. In two of the four validation assays performed, the standard curve for VEGF showed acceptable results at all eight points (from 0 – 10,000 pg/mL). In one of the other two assays the VEGF standard curve did not show acceptable recovery values (as a percentage of observed concentration/expected concentration) until standard point 5 (156.3 pg/mL) and in the other curve until standard point 6 (625 pg/mL). Since the ability to measure the lowest VEGF standard (2.4 pg/mL) is inconsistent, the LLOQ was set at the vendor's reported sensitivity level of 13.6 pg/mL. Although the LLOQ for VEGF was not below the lowest standard, the LLOD for VEGF was below the lowest standard. Therefore, the results of the sensitivity studies show this method is sensitive at or below (i.e. more sensitive) the lowest standard for all 23 cytokines and this method meets the pre-established acceptance criteria for sensitivity.

Table 5: Results of sensitivity studies. The lower limit of quantification (LLOQ), lower limit of detection (LLOD), lowest standard and Millipore's reported sensitivity for each analyte are shown. In order to meet assay acceptance criteria, it is expected that the LLOD will be at, or below, the lowest standard for that analyte.

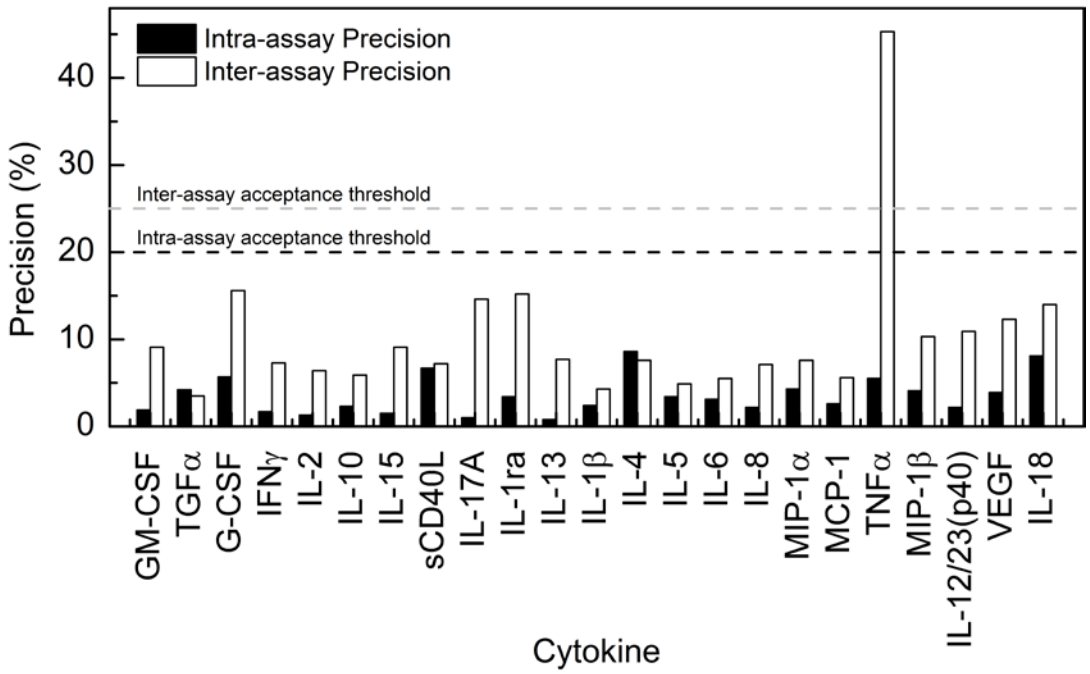
Analyte	Lowest Standard (pg/mL)	LLOD (pg/mL)	LLOQ (pg/mL)	Vendor Reported Sensitivity (pg/mL)	Outcome
GM-CSF	2.4	0.6	2.4	1.8	Pass
TGFα	2.4	1.2	1.2	1.1	Pass
G-CSF	2.4	0.6	2.4	2.1	Pass
IFNγ	2.4	0.6	2.4	1.6	Pass
IL-2	2.4	0.6	2.4	2.1	Pass
IL-10	12.2	3.05	12.2	6.4	Pass
IL-15	2.4	2.4	2.4	0.5	Pass
sCD40L	2.4	2.4	2.4	2.1	Pass
IL-17A	2.4	0.6	2.4	1.3	Pass
IL-1ra	2.4	2.4	2.4	2.4	Pass
IL-13	2.4	0.6	2.4	5.8	Pass
IL-1β	2.4	1.2	1.2	1.2	Pass
IL-4	4.9	4.9	4.9	3.1	Pass
IL-5	2.4	1.2	1.2	1.5	Pass
IL-6	2.4	2.4	2.4	1.6	Pass
IL-8	2.4	1.2	2.4	1.1	Pass
MIP-1α	2.4	0.6	2.4	4.9	Pass
MCP-1	2.4	2.4	2.4	3.1	Pass
TNFα	2.4	0.6	2.4	1.6	Pass
MIP-1β	2.4	0.6	0.6	1.6	Pass
IL-12/23(p40)	2.4	2.4	2.4	1.5	Pass
VEGF	2.4	2.4	13.6	13.6	Pass
IL-18	12.2	6.1	12.2	6.1	Pass

Precision

Intra-assay precision was determined by performing 12 replicate measurements at two targeted concentration levels (low and high with relation to the standard curve) for a total of 24 observations within a single assay. The averages of these 24 observations are reported as the overall intra-assay precision. All cytokines met acceptance criteria of CV

< 20% for intra-assay (within plate) precision. Inter-assay (between plate) CVs were generally higher than within-plate CV, as expected, since assays performed on different days introduce an additional variable. Inter-assay CV was acceptable (<25%) for 22 of 23 cytokines; TNFα showed an unacceptable inter-assay CV of 45.3%. Inter-assay CV was calculated from the average of 4 different control samples (two commercial controls and 2 in-house NHP controls), each tested in triplicate in four assays performed on four separate days over six months. The vendor reported intra-assay precision is <5% and the inter-assay precision is <15% for all analytes. The current study agreed with the manufacturer reported CV of <5% in 18 of 23 (78%) analytes for intra-assay precision and in 20 of 23 (87%) analytes for inter-assay precision. The results of the precision studies are presented in Figure 2; the calculated CV for each cytokine is presented in Appendix C, Table 13.

Figure 2: Precision data. For intra-assay precision n=24 in one assay. For inter-assay precision n=4 (4 assays in which each control was tested in triplicate).



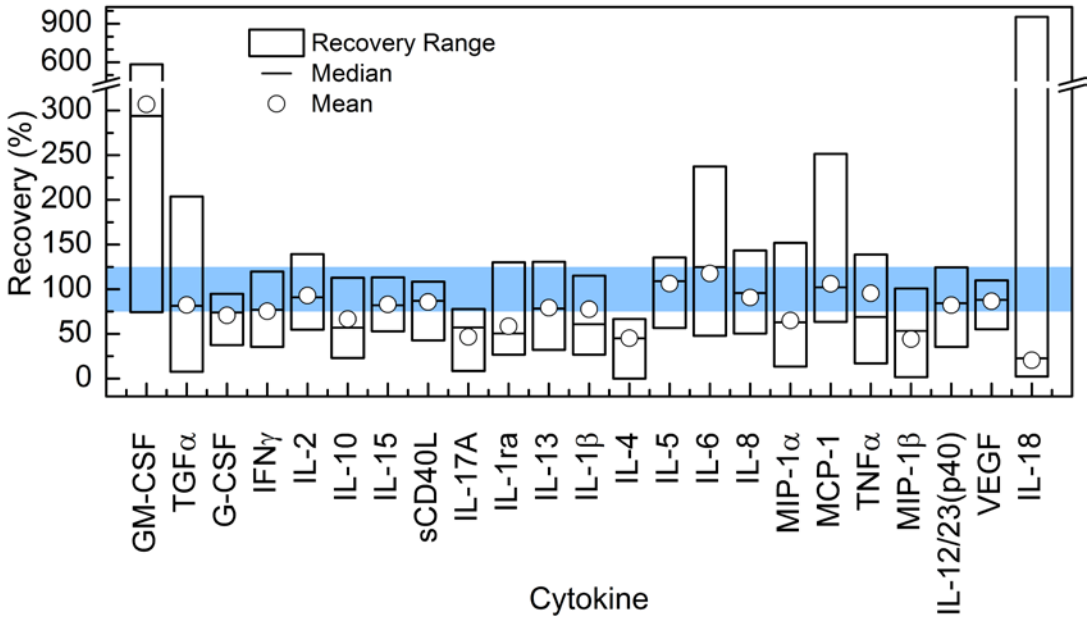
Recovery

Recovery was analyzed in seven individual animal samples and one pool made from the serum of two healthy cynomolgus macaques. Of these eight samples (representing nine animals) recovery performance varied considerably even when testing for the same analyte. For example, in one assay, TGFα showed an average recovery of 114% in animal #1 while animal #2 resulted in an average TGFα recovery of 70%. In another assay, animal #3, #4 and #5 showed average TGFα recovery of 93%, 204% and 88% respectively. This animal-to-animal cytokine recovery variability was also seen in other analytes (IFNγ, IL-10, MIP-1α, TNFα). Some cytokines were consistently recovered

across individual animals and assays. For example, IL-5, IL-8 and MCP-1 all had acceptable recovery levels across all determinations. However, GM-CSF was recovered at percentages from 200% - 300% across all animals and across all assays suggesting that these results are overestimated due to some form of interference or artifact. IL-17A, IL-4, MIP-1β and IL-18 showed consistently low recoveries across animals and assays which could reflect an underestimate of true values.

To meet assay acceptance criteria, recovery is expected to be 75% to 125% of expected values. The manufacturer reported recovery (reported as "accuracy") results ranging from 87-101% for 22 out of 23 cytokines, the 23rd cytokine, VEGF, was reported to have an accuracy of 70%. The results reported here agree with manufacturer reported recovery for one cytokine and disagree (data reported here do not meet recovery expectations) with manufacturer reported recovery for 18 of 23 cytokines. Recovery for 4 cytokines in the current study was better (closer to 100% of the expected values) than that reported by the manufacturer. Results of recovery studies are presented in Figure 3 and calculated recovery values, along with manufacturer reported accuracy values are presented in Appendix C, Table 14.

Figure 3: Recovery results. Seven individual animal samples and one NHP serum pool were spiked with known amounts of standard at various levels to yield 17 separate spiked samples, each measured in duplicate, over 3 assays. Each recovery point plotted here is the average of duplicates from one of the 17 spiked samples. The shaded band from 75% - 125% indicates the acceptable range.

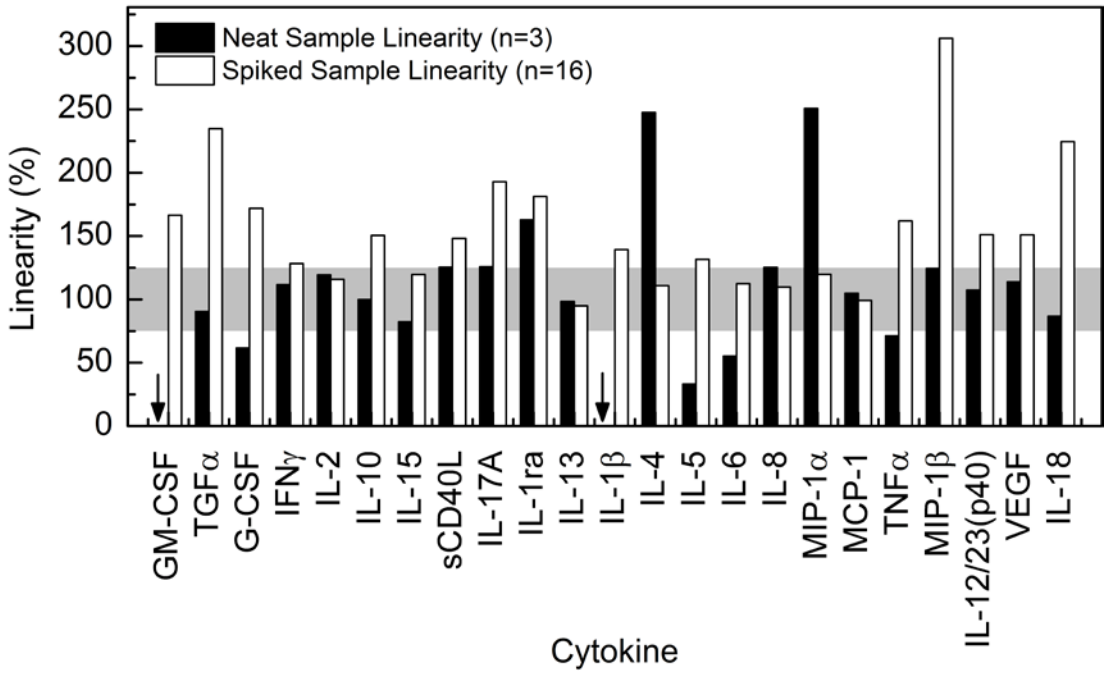


Linearity

Linearity is studied by diluting a sample of known concentration to specific dilution levels and comparing the observed concentration to the expected concentration after analysis. Acceptance criteria require that observed concentrations are determined to be 75-125% of expected concentrations. Endogenous cytokines in healthy animals are generally very low or undetectable for most of the 23 cytokines tested here [5], [71], making dilution studies unreliable in the samples from these healthy animals. Therefore, some NHP samples were spiked with the highest standard in order to create a dilution

series. When possible, the average linearity of neat (un-spiked) samples is shown in addition to average linearity from the sample spiked with standard. Since endogenous cytokine concentration was expected to be low, a dilution series was not constructed from the un-spiked samples; rather a single 50% dilution of these samples was performed. Spiked samples were diluted to 50% and 20% in seven individual animals and one serum pool, each of the resulting dilutions were analyzed in duplicate across three assays, over three weeks. All three samples that were tested neat had undetectable levels of GM-CSF and IL-1β, therefore no linearity data could be gathered for those analytes until additional GM-CSF or IL-1β was spiked in. The results of the linearity studies are presented in Figure 4 and calculated linearity results are presented in Appendix C, Table 15. The manufacturer did not generate linearity data from this method.

Figure 4: Linearity results from neat samples and samples spiked with the highest assay standard. Arrows indicate cytokines that were undetectable when tested neat. The shaded horizontal band indicates the acceptable linearity range (75% - 125%). Arrows indicate undetectable levels where dilutions could not be performed.



Acceptance Criteria

Each of the 23 cytokines were categorized as acceptable, marginal or rejected based upon whether the measurement characteristics met validation acceptance criteria for sensitivity, recovery, precision and linearity. All of the cytokines met the acceptance criteria for sensitivity and intra-assay precision, and all but one cytokine, TNFα, met the acceptance criteria for inter-assay precision. Categorization results for recovery and linearity are summarized in Table 6. The tests for recovery and linearity were performed in seven individual animals (because individual animal variation is expected) and one pool made

from the serum of two cynomolgus macaques. More than half of these eight individual samples tested for recovery had to meet acceptance criteria (75-125%) to be classified as "accepted" for categorization purposes. Additionally, either the linearity of all samples diluted neat (without addition of kit standard) had to fall in the acceptable range, *or* the linearity of more than half of all samples (regardless of dilution neat or after spiking) had to fall in the acceptable range to be classified as "accepted" for categorization purposes. If both conditions were met for recovery and linearity then the cytokine was categorized as "accepted," if neither condition was met then it was classified as "rejected" and if it met only one of the two conditions it was classified as "marginal".

TNFα failed to meet acceptance criteria for inter-assay precision, recovery and linearity. Therefore this method is not suitable for the measurement of TNFα and it is categorized as 'rejected'.

Table 6: Validation categorization and outcomes for the 23 cytokines included in the NHP cytokine panel.

Cytokine	Acceptable Recovery*	Acceptable Linearity**	Validation Outcome
GM-CSF	no	no	Reject
TGFα	yes	yes	Accept
G-CSF	yes	no	Marginal
IFNγ	yes	yes	Accept
IL-2	yes	yes	Accept
IL-10	no	yes	Marginal
IL-15	yes	yes	Accept
sCD40L	yes	no	Marginal
IL-17A	no	no	Reject
IL-1ra	no	no	Reject
IL-13	yes	yes	Accept
IL-1β	no	no	Reject
IL-4	no	yes	Marginal
IL-5	yes	yes	Accept
IL-6	yes	yes	Accept
IL-8	yes	yes	Accept
MIP-1α	no	yes	Marginal
MCP-1	yes	yes	Accept
TNFα	no	no	Reject
MIP-1β	no	no	Reject
IL-12/23(p40)	yes	yes	Accept
VEGF	yes	yes	Accept
IL-18	no	no	Reject

*>50% of the data from the eight samples demonstrated average acceptable recovery (75-125%)
**>50% of the data from the eight animals tested demonstrated acceptable average linearity AND/OR overall acceptable linearity (75-125%) of endogenous (neat) samples

This research demonstrates that when this multiplex method is used, 11 out of the 23 cytokines (48%) can be reliably measured based upon the method being demonstrated to meet validation criteria. Five cytokines (22%) were determined to be of marginal reliability because all validation criteria were not consistently met. Method validity for 7 of the 23 cytokines (30%) could not be assured and so the use of this method for those analytes was rejected. These results are summarized in Table 7.

Table 7: Validation outcome for the 23 cytokines in the multiplex assay kit.

Accept	Marginal	Reject
TGFα	G-CSF	GM-CSF
IFNγ	IL-10	IL-17A
IL-2	sCD40L	IL-1ra
IL-15	IL-4	IL-1β
IL-13	MIP-1α	TNFα
IL-5		MIP-1β
IL-6		IL-18
IL-8		
MCP-1		
IL-12/23(p40)		
VEGF		

The 11 cytokines categorized as accepted have met acceptance criteria for sensitivity, precision, recovery and linearity. All five of the marginal cytokines (G-CSF, IL-10, sCD40L, IL-4 and MIP-1α) had acceptable intra- and inter-assay precision but varied in their recovery and linearity performance. G-CSF showed acceptable recovery (75-125%) in four of the seven animals tested (samples tested for the eighth animal had low bead counts and no results were available), while two of the eight animals showed recovery ranges near 70%. Linearity was acceptable for G-CSF in two of the three samples tested neat, but not for any of the other samples tested. Next, IL-10 showed acceptable recovery in two of eight animals, while the remaining six animals resulted in recoveries ranging from 23% - 70%. Linearity of IL-10 was very similar to recovery data in that only three of eight animals showed acceptable linearity results. Recovery for sCD40L was generally acceptable with four of six animals showing acceptable results (results were not obtained for one animal due to low bead count and for another due to very high results that were above the readable range). Linearity was, however, poor for sCD40L as only

one of seven animals showed acceptable results. In contrast to sCD40L, IL-4 showed poor recovery but acceptable linearity: none of the eight animals tested for recovery showed acceptable results, but half of those same animals showed acceptable linearity results. Finally, recovery for MIP-1α was acceptable in two of eight animals while recovery for two other animals was just outside the acceptable range. Linearity of MIP-1α was acceptable in four of eight animals. This study demonstrates that the multiplex method is not fit for purpose for measuring these 7 cytokines. Concentrations of these 7 cytokines measured in the NHP as part of thesis Aim 2 are provided for completeness (but not interpretation) in Appendix D.

Aim 2: Measurement of Cytokines in Non-Human Primate Serum

Non-human primate data for the 15 cytokines that were classified as accepted or marginal for method acceptability are presented in Table 8 (although IL-10 validation data were considered marginal, it is not included here as NHP data are not available for this cytokine due to the assay format transition described previously). The median, mean, standard deviation and percentage of samples within the working range (standard curve range) of the assay is shown for each cytokine to illustrate concentration distribution.

Table 8: Cytokine concentration distribution for 36 individual NHP (averages were calculated for animals with multiple samples).

Analyte	Range (pg/mL)	Median (pg/mL)	Mean (pg/mL)	Standard deviation (pg/mL)	% of samples within working range of assay
G-CSF	<0.64 - 18.4	3.7	4.6	4.5	68%
IFNγ	<0.64 - 98	2.0	9.4	19.5	78%
IL-4	<0.64 - 2.8	<0.64	<0.64	N/A	11%
IL-2	1.3 - 78	8.4	12.5	13.5	100%
IL-5	<0.64 - 15	<0.64	1.2	2.7	28%
IL-6	<0.64 - 46	<0.64	2.6	4.3	44%
IL-8	271 - >10,000	2,033	2,962	2,923	97%
IL-12/23 (p40)	24 - 584	96.3	130	116	100%
IL-13	<0.64 - 13	2.2	3.0	2.8	83%
IL-15	<0.64 - 18.4	3.4	4.6	4.0	92%
MCP-1	102 - 632	344.2	360	129	100%
MIP-1α	<0.64 - 256	7.9	17.6	42.4	86%
sCD40L	188 - >10,000	4,956	5,511	3,086	89%
TGFα	2.4 - 47	8.4	11.5	9.6	100%
VEGF	<0.64 - 1,798	10.3	93	305	81%

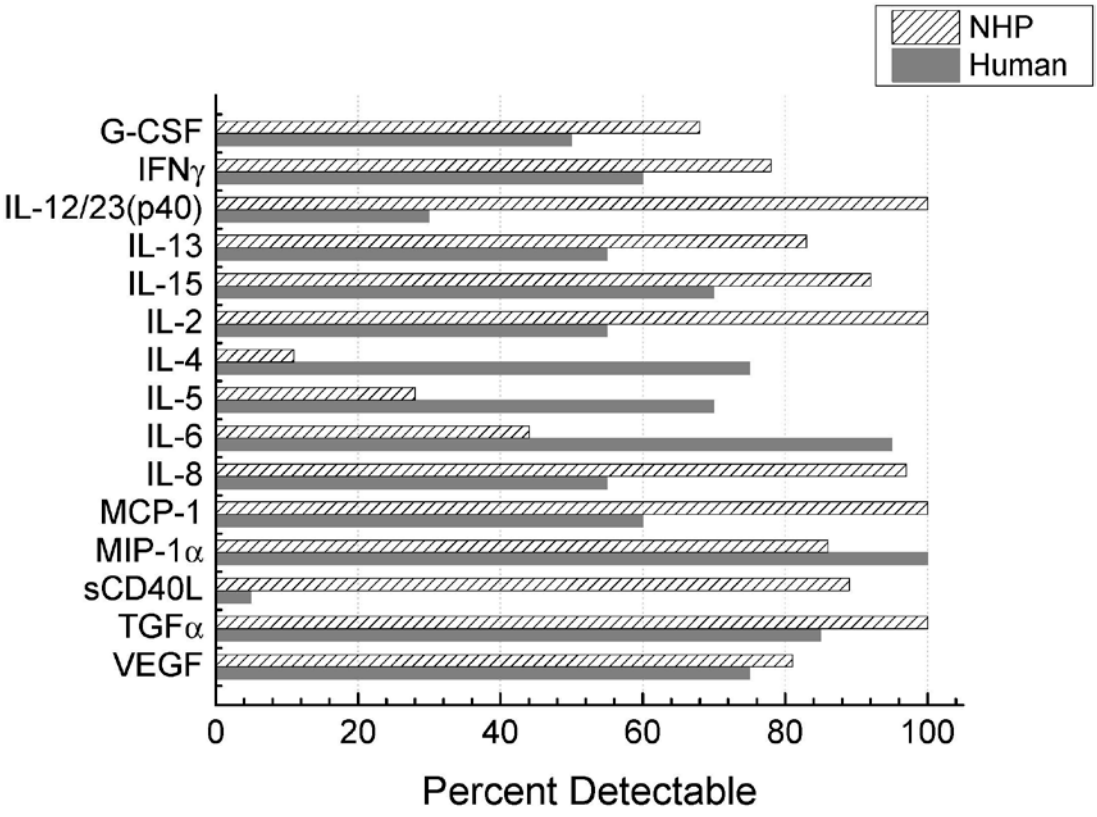
For comparison purposes, Table 9 provides human cytokine concentration ranges obtained using the manufacturer's human multiplex kit counterpart to the NHP kit. The human data presented in Table 9 were provided by the manufacturer with the disclaimer that because of sample size and randomization limitations, data represent a guideline and not an established reference range [72]. Figure 5 displays the percent of samples tested that are within the detectable ranges of the assays. Comparisons between NHP and human are of interest because there are similarities between the human and non-human primate, and the observed cytokine ranges may be similar. In addition, when necessary, NHPs are used in research trials to evaluate safety and efficacy prior to translation for use in the human.

Table 9: Cytokine concentration range determined in human and NHP sera. The concentration range (and percent of samples within the working range of the assay) are presented.

Cytokine	Concentration Range, Human (pg/mL; n=20)*	Concentration Range, NHP (pg/mL; n=36)
G-CSF	<3.2 – 756 (50% detectable)	<0.64-18.4 (68% detectable)
IFNγ	<3.2 – 22.2 (60% detectable)	<0.64-98 (78% detectable)
IL-4	<3.2 – 10.7 (75% detectable)	<0.64-2.8 (11% detectable)
IL-2	<3.2 – 102 (55% detectable)	1.3-78
IL-5	<3.2 – 403 (70% detectable)	<0.64-15 (28% detectable)
IL-6	<3.2 – 10 (95% detectable)	<0.64-22 (44% detectable)
IL-8	<3.2 – 156 (55% detectable)	271- >10000 (97% detectable)
IL-12/23(p40)	<3.2 – 27.9 (30% detectable)	24-584
IL-13	<3.2 – 362 (55% detectable)	<0.64-13 (83% detectable)
IL-15	<3.2 – 14.9 (70% detectable)	<0.64-18.4 (92% detectable)
MCP-1	<3.2 – 81.8 (60% detectable)	102-632
MIP-1α	63.5 – 439	<0.64-256 (86% detectable)
sCD40L	<3.2 – 6.4 (5% detectable)	188- >10000 (89% detectable)
TGFα	<3.2 – 118 (85% detectable)	2.4-47
VEGF	<3.2 – 118 (75% detectable)	<0.64-1798 (81% detectable)

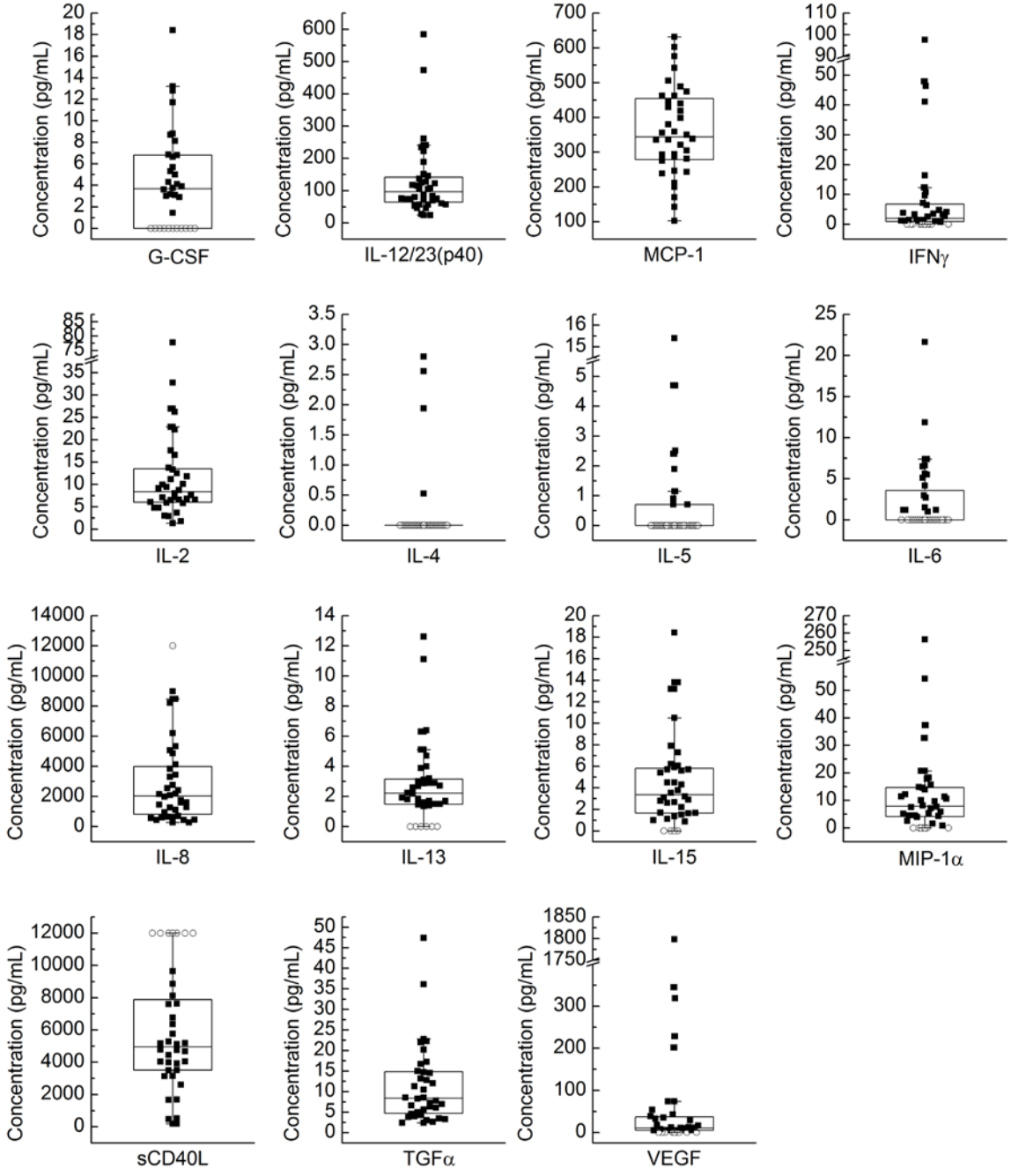
*All human cytokine concentration data provided by Millipore as obtained from their Milliplex MAP Human Cytokine/Chemokine 42-plex panel (catalog number MPXHCYTO-60K) [73].

Figure 5: The percent of NHP and human samples that are within detectable ranges of the respective assays.



The box and whisker distribution plots for cytokines measured in each animal (only accepted and marginal data displayed) are provided in Figure 6. Points representing samples outside the limit of detection are differentiated by shape (open circles). Outliers are shown beyond the reach of the maximum whisker and were calculated as being ≥ 1.5 times the inner-quartile range (IQR) above the 75th percentile.

Figure 6: Distribution of cytokine concentrations obtained from healthy cynomolgus macaques Top of the box = 75th percentile; middle line = median; bottom of the box = 25th percentile; whiskers extend to the minimum and maximum (excluding outliers); outliers are shown beyond the reach of the maximum whisker.



Discussion

The aims of this thesis were 1) to determine if a multiplex method for the measurement of multiple cytokines in NHP serum was valid, accurate and reliable and 2) to use the method to determine if the assay is suitable for the detection of cytokine concentrations in a study population of healthy cynomolgus macaques.

Aim 1: Method Validation Data

Validation is necessary to demonstrate 1) that a method is fit for its intended purpose, and 2) to demonstrate that the method will work as described by the manufacturer under the laboratory and sampling conditions required to meet research objectives. For instance, method validation is especially critical when a test method manufactured for use in one species is applied to other species because the blood or serum matrix will differ among species and this difference may affect method performance. In the current study, a validation of the multiplex method for cytokine analyses in the NHP was performed by evaluating the sensitivity, precision, recovery, and linearity of the procedure when used with serum from cynomolgus macaques. Acceptance criteria for method validity were pre-established and based on commonly used parameters for the multiplexed method [56], [57], [64].

Multiplexing technology has significant animal welfare (reduced animal handling and sample volume) and research (increased efficiency) advantages when compared to traditional methods like ELISA, due to the ability to simultaneously measure multiple analytes (cytokines) from one small volume (25 µL or less) sample. While multiplex assays cost more than ELISA, they become cost effective when four or more analytes can be accurately measured in one sample as opposed to performing four individual cytokine ELISAs) [74]. Understandably, multiplexed assay costs increase as the number of analytes being measured increases; therefore, it is critical to confirm that the method is appropriate for each analyte. Because cytokines exist as part of a complex and inter-connected network, simultaneous detection of these biomarkers presents an opportunity to measure a panel of cytokines, rather than a single cytokine. While these advantages are compelling, the procedures for validating a multiple analyte method will be more complex because error and variability are not likely to be constant or consistent within a panel grouping in spite of effort by the manufacturer to optimize the procedure for all analytes.

In order to identify and understand the error and variability associated with a method, it is necessary to determine if the assay is fit-for-purpose by performing a method validation assessment prior to using the method to meet research or clinical objectives in the species of interest. As summarized in the Introduction, previous studies provide validation data for multiplex methods used with human samples [41], [55–57], [74], [75], however, information related to method suitability for use with NHP samples is limited.

Data from the current study demonstrate that the precision of the multiplex method is acceptable and comparable to that reported by the manufacturer for most of the 23 cytokines in the measurement panel. However, for 18 of the 23 cytokines, the recovery and linearity of the method were highly variable, generally unacceptable, and not favorably comparable with the spike-recovery results reported by the manufacturer. For example, none of the 23 cytokines met the recovery acceptance criteria (75 to 125 %) in all eight (or even seven) of the animal samples tested. Six of the eight animals had acceptable recovery for seven cytokines but, in contrast, there were also six cytokines where none, or only one, of the animals demonstrated acceptable recovery results. Similarly, the linearity results were also highly variable: across the 23 cytokines, anywhere from zero to all eight of the animals had acceptable linearity results. Prabhakar et al. found similar inconsistencies using a human multiplex kit (Linco) for use with human serum samples; they demonstrated acceptable precision (<20% intra-assay and 25% inter-assay), but highly variable recoveries (28% - 127%) for the six cytokines tested [56]. When using a human Biorad 17-plex cytokine kit, Siawaya et al showed recovery results with similar variability to those presented in this thesis; they observed recovery ranges from 20% - 325% [55]. The poor recoveries demonstrated in assays from different manufacturers suggest that accurate measurements can be difficult to obtain with the multiplexed method. Therefore, the tradeoff of convenience and efficiency may not be worth the loss in data reliability for some of the analytes tested.

Because cytokines typically have short half-lives [31] and are potent and can exert their effects at picomolar concentrations [5], it is important to confirm the sensitivity of each method for each analyte. In the validation study reported here, sensitivity measures met the acceptance criteria (at or below the concentration level of the lowest standard) for all 23 cytokines evaluated.

Some cytokines, such as IL-17A, MIP-1β and IL-18, demonstrated consistently low recovery results and were subsequently rejected based on this failure to meet required acceptance criteria. Poor recovery may be caused by cross-reactivity, which can result from matrix-effects. Matrix-effects may arise because of the protein rich environment of biological fluids which includes proteases, binding proteins, or other proteins structurally similar to the cytokine being measured. It has been shown that alpha-2 macroglobulin, a common plasma protein (i.e. part of the blood matrix) can lower the measurements of IL-2, IL-6, IL-4, IL-1β and TNFα in ELISA assays by up to 26% [77]. The role of this protein in the study reported here is unknown, but it is feasible that it may affect other cytokines in ELISA, and perhaps, in multiplex cytokine measurements as well. Similarly, autoantibodies to IL-6 and IL-1α have proven to inhibit their respective cytokines in vitro, resulting in falsely low ELISA measurements [78], [79]. Soluble receptors for cytokines, such as TNFα, can also be present in the blood matrix and have been implicated in erroneously high cytokine measurements, adding additional potential for interference [80]. In the pilot study for Enbrel (etanercept), a soluble TNFα receptor used to neutralize biological TNFα activity, patients with multiple myeloma had a

significantly increased plasma TNFα concentration during treatment than prior to treatment (median 8.0 pg/mL pre-treatment, 244 pg/mL during Enbrel treatment) [81]. Eason et al. reported that etanercept acts as a carrier, extending the half-life of TNFα which may account for the higher concentration observed in treated patients [82]. This group also reported that although high TNFα concentrations were detected, there was an undetectable level of TNFa bioactivity, which leads to the conclusion that both free and bound TNFα could be detected using the methods employed in that study (ELISA from R&D Systems). Due to the proprietary nature of manufacturer antibody binding sites, it is unknown if Millipore's NHP cytokine assay would detect both free and bound TNFα. If etanercept is administered to NHPs (as may be done during transplant studies) it may result in higher serum TNFα concentrations than in animals that did not receive etanercept. Diurnal and seasonal fluctuations in certain matrix proteins (blood clotting factors and various hormones) have been documented in humans [82–84] indicating a potential source of the intra- and inter-animal variability observed in the samples tested. In fact, cortisol, which has a diurnal fluctuation, has been observed to inversely affect the inflammatory cytokines IFNγ, TNFα, IL-1 and IL-12. Normal, lower levels of cortisol in humans from midnight to 5:00 am may contribute to the reported increase of inflammatory symptoms during these times [86]. This relationship also suggests that the cortisol secreted in response to acute or chronic stress may influence concentrations of cytokines at any given time.

The multiplex kit being tested in this thesis is described as being a NHP based method. However, many of the reagents are human based. A human protein source is used for the standards and controls for 14 of the 23 cytokines, detection antibodies were raised against human proteins for 21 of the cytokines and the normalizing matrix is made from human serum. To ensure that the assay works with NHP samples, the manufacturer determined reactivity for all 23 analytes in seven species of NHP (chimpanzee, baboon, rhesus, cynomolgus, pigtail, African green and marmoset) by collecting serum and PBMCs from two animals of each listed species. Serum was challenged with LPS, and PBMCs were challenged with LPS or phytohemagglutinin, then the samples were testing using the assay kit and categorized as "+++," (strongly reacting) "++," "+/-" or "-," (no detectable response). Although this technique semi-objectively indicates reactivity, it does not describe efficiency in antibody binding of the different cytokines. In a study that compared amino acid homology for 14 cytokines (ten in common with the cytokines studied in this thesis) between human, rhesus macaque, pigtailed macaque and sooty mangabey, cytokine amino acid homology ranged from 93% to 99% between humans and NHPs, and ranged from 97% to 100% homology between NHP species [87]. Although minor, if these species-specific differences occur at a critical antibody binding site, they could contribute to the variability identified in this study.

Five cytokines were determined to be 'marginally' acceptable for meeting method validation criteria. In general, the main problem with these analytes was the lack of consistency among observed results within animals and between assays. Possible

explanations again include the interfering effects of matrix components both within animals and over time (analyte stability). Protein instability challenges the optimization of multiplex systems because preservatives may not equally and consistently protect all analytes.

Some individual cytokine measurements were consistently imprecise. For example, TNFα measurement achieved acceptable intra-assay precision performance, but failed when evaluated for inter-assay precision, with a high CV of 45%. Of the four controls used to determine inter-assay precision for TNFα, three were at the low end of the detectable range; with average concentrations of 2.5 pg/mL, 12.9 pg/mL and 18.6 pg/mL (detectable range for TNFα is 2.4-10,000 pg/mL). Therefore it is expected that precision will be reduced at the lower limit of the detectable range, and as recommended by the FDA, higher acceptable CVs are appropriate at the lower level of quantitation (LLOQ). However, in this case, even the fourth control for TNFα, which had an average concentration of 373 pg/mL, replicated poorly (39% CV) between plates. Given that the detectable range of this cytokine extends to 10,000 pg/mL, 373 pg/mL is a relatively low concentration; however the majority of the other 22 cytokines had average control concentrations as low as, or lower than, 373 pg/mL but still performed within acceptable limits. TNFα is known to be particularly labile so it is possible that even the use of freshly reconstituted controls will not assure its stability and acceptable inter-assay precision.

In a study by Breen et al., in which the differences in results between labs, manufacturers and lot numbers of cytokine multiplexed kits were reviewed, it was shown that reproducibility between labs can be poor and that relative concentrations and changes in animals over time were of more value than the absolute concentration observed [44]. With different manufacturers, this likely has to do with variable antibody pairs, which remain proprietary secrets. In the current study, all data represent work performed using the same kit lot number, laboratory site and technician. Therefore, the testing environment should be optimal for generating consistent test results. These complexities suggest that, not unexpectedly, consistency in laboratory methods is important.

To summarize the first aim of this thesis, we have described the sensitivity, precision, recovery and linearity for each cytokine in the multiplex panel. We have categorized the method validity for the measurement of each of the 23 cytokines as being approved, marginal (G-CSF, IL-10, cSD40L, IL-4 and MIP-1α) or rejected (GM-CSF, IL-17A, IL-1ra, IL-1β, TNFα, MIP-1β and IL-18). These results are critical when selecting cytokine analyses for research or clinical use and for the appropriate interpretation of cytokine results when this method is used. For instance, for cytokines like IL-2, IL-5 and MCP-1 (all accepted), that are both accurately and precisely measured, we have evidence that testing samples in duplicate will give us results representative of the true concentration. For analytes with poor recovery but good precision, like GM-CSF, IL-17A and MIP-1β (all rejected), we are unlikely to get an accurate result no matter how many times we repeat it. For cytokines that have good recovery but poor precision (not identified in this

study) it would be desirable to test samples many times to get average results that approximate true concentrations. Finally, cytokines with poor recovery and poor precision (not represented in this study) should not be measured with this method as they have been demonstrated to be unreliable (not fit for purpose) and uninterpretable.

In diagnostic service and research environments, it is critical to define the inherent error and limitations of the methods performed in order to make recommendations for their appropriate use. In the current study, results indicate that some cytokines in the commercial panel can be reliably measured and others cannot. Once those limitations are understood, it may be possible to improve the performance by optimizing the method further, either within the laboratory of use or at the manufacturer level. Manufacturers may be able to improve method performance based on the independent findings and feedback of product users. For instance, manufacturers might respond by determining if native proteins have any advantage over recombinant proteins when trying to reduce variability of results.

Measuring cytokines in bodily fluids is complicated because cytokines degrade quickly; cytokines in urine disappear within hours when kept at room temperature and are only stable for several weeks when maintained at -80^{o}C [88]. Additional protease inhibitors can be added to samples (especially for very labile cytokines, like TNFα) to mitigate issues with protein instability. The interference of other proteins in the sample matrix can often be a problem in ligand-binding assays but options exist for combating this. A

2009 study by Julie Doucet et al. reported on the successful reduction of matrix interferences in an ELISA for monoclonal antibodies in cynomolgus monkeys. This was achieved by adding a specific mixture of dimethyl sulfoxide (DMSO) and a chaotropic salt (a salt that dissociates and disrupts hydrogen bonding, such as NaSCN or $MgCl_2$) to aid in reducing the number of non-specific (weak) interactions while leaving specific (strong) interactions intact [89]. Again, given that there are numerous cytokine binding events happening in the same well, this method may work better for some cytokines than for others.

Data reported here and in the literature point to a difficulty in obtaining consistent and accurate results using the multiplex method. Although manufacturers may report validation data for multiplexed assay kits, the research community cannot assume that the assay will work the same way in all settings and must determine these characteristics in their own laboratories.

Aim 2: Measurement of Cytokines in Non-Human Primate Serum

The second aim of this study was to determine if the methods tested in Aim One would be useful for the measurement of cytokines in healthy cynomolgus macaques. As mentioned in the Introduction, many of the cytokines studied here are involved in inflammatory and immune response pathways and are not expected to be present at high concentrations in peripheral blood of healthy animals. Under these conditions, sensitive

methods are required to accurately detect low concentrations. Therefore, method sensitivity and additional performance measures were evaluated in Aim One. Aim Two was designed to evaluate whether the multiplex method can detect baseline cytokine activity in healthy cynomolgus macaques. Namely, Aim One was designed to evaluate the technical performance of the method and Aim Two was designed to evaluate the practical performance of the method for use with samples from healthy cynomolgus macaques.

The 15 cytokines that were considered either acceptable or marginal were reviewed to determine if the values observed in the cohort of 36 healthy NHP were generally comparable to what has been reported in healthy humans using the Millipore multiplex method. Available data reflecting human cytokine values are summarized in Table 9. The cytokine, IL-10, which was not detectable in cynomolgus macaque serum using the polystyrene-bead assay kit, was excluded from this review.

The NHP cytokine data observed for IL-2, IL-4, IL-6, IL-15, and TGFα are similar to the human data reported for these cytokines (reported by Millipore using data obtained from their human cytokine multiplex assay; see Table 9).

Low or systemically undetectable levels of G-CSF are expected in healthy NHPs, and low values were observed (<0.64 – 18.4 pg/mL; 68% detectable) in this study. In addition, a similar percentage of samples had detectable concentration in the 20 people

tested by Millipore (50%) as had detectable concentration of G-CSF in the 36 NHPs (68%) sampled in this study. Because G-CSF is an activator of neutrophilic activity and neutrophil turnover is so high (about 1.6 billion cells per kilogram per day, in humans) [90], some neutrohpilic stimulation is necessary for normal immune function. However, neutrophil stimulation via G-CSF is not expected to be high in healthy human adults except during times of immunological challenge (e.g. bacterial infection) [91]. The highest G-CSF concentration reported by Millipore, using their human cytokine multiplex kit, is 756 pg/mL, which corresponds to a reported average G-CSF concentration found in people in the acute stage of infection (732 pg/mL) [91]. It is possible that one of the subjects tested by Millipore may not have been healthy.

IFNγ is critical for viral immunity but is expected to be low in healthy individuals. Millipore reported that 60% of human samples tested for IFNγ were between 3.2 and 22.2 pg/mL while the remaining 40% were undetectable. In the present study, eight (22%) of the NHPs had undetectable IFNγ concentration while the remaining 78% had IFNγ concentrations from 0.76 pg/mL to 97.6 pg/mL. The animals included in this study are not guaranteed to be pathogen free and some are known carriers of *Campylobacter*, *Shigella*, *Salmonella*, and/or *E. coli* bacteria. In addition, temporary leukocytosis is commonly observed in many of these NHPs, possibly due to reactivation of latent viral infections (e.g. Epstein-Barr virus, cytomegalovirus). Although all animals were clinically asymptomatic, this could also cause increased IFNγ concentrations.

IL-5 is important for promoting the growth of eosinophils which play an important role in helminth immunity, allergies and asthma. Likewise, IL-13 plays an important role in allergic conditions due to its role in immunoglobulin isotype switching to IgE [92]. IL-5 and IL-13 are therefore not expected to be high in healthy individuals [93]. IL-13 was found to range from undetectable levels up to 13 pg/mL in the NHPs but the human range was reported (by Millipore) to be <3.2 – 362 pg/mL. Only 10 of the 36 NHPs (28%) had IL-5 concentration above the detection level and then the highest value was only 15 pg/mL, whereas 70% of the humans tested showed detectable levels, up to 403 pg/mL. In a study from 1994 comparing IL-5 levels in asthmatic patients and their healthy counterparts, the healthy patients all showed IL-5 levels less than 40 pg/mL, whereas patients with asthma had a range of IL-5 from 40 – 690 pg/mL [93]. The IL-5 levels observed in the NHPs seem to fit the expected ranges for healthy subjects while the IL-5 levels reported by Millipore for humans suggest that some of the subjects may have allergies or asthma; the high level of IL-13 reported by Millipore in their subjects supports this hypothesis.

Il-8 is a proinflammatory cytokine that acts as a neutrophil chemoattractant [94]. Sun et al. reported low levels of IL-8 in healthy adults (5 – 11 pg/mL), while their diseased counterparts (patients with oral lichen planus, an inflammatory disease) showed IL-8 levels of 5 – 449 pg/mL [95]. In a study by Zimmermann et al., people with end-stage liver disease were shown to have serum IL-8 concentrations as high as 1500 pg/mL [96]. The range of IL-8 reported by Millipore in human subjects was <3.2 – 156 pg/mL, again

suggesting the possibility that at least one subject tested was unhealthy, based on the literature cited here. In contrast, Barra et al. reported that 27 healthy controls showed IL-8 ranging from 35 - >20,000 pg/mL. This wide range fits more closely with the range of IL-8 observed in the NHPs in this study (which ranged from 271 - >10,000 pg/mL). The validation data obtained for IL-8 showed an average recovery of 91% demonstrating that the assay can accurately detect this cytokine; therefore a biological explanation is needed. Again, asymptomatic bacterial activation could account for the high concentrations of this proinflammatory cytokine but further study of this interesting result is needed to make conclusions.

The shared p40 portion of IL-12 and IL-23 can act as its own cytokine, blocking the binding of IL-12 and IL-23, thereby limiting inflammation [97]. Millipore has reported that human IL-12/IL-23(p40) ranged from <3.2 – 28 pg/mL and only 30% of subjects had levels over 3.2 pg/mL. Human blood stimulated with LPS for 24 hours has been found to have p40 concentrations up to 334 pg/mL, when tested using a Millipore multiplex assay [98]. These data suggest that low (<30 pg/mL) levels may be normal while elevated levels could indicate a disease state. However, Wang et al. reported in 2013 that p40 levels in osteosarcoma patients were 202 pg/mL while the healthy controls were observed to have higher average p40 levels at 269 pg/mL [99]. The average concentration of IL-12/23(p40) in the NHPs was 130 pg/mL (median 88 pg/mL) and the three highest IL-12/23(p40) concentrations measured were found to be outliers (at 261, 473 and 584

pg/mL, using the Tukey method). Until more data can be obtained, these results represent the range of variation in healthy cynomolgus macaques.

MCP-1 attracts monocytes to inflammation sites and was reported by Millipore to be undetectable in eight people while the remaining 12 people showed levels up to 82 pg/mL. This is considerably lower than the range observed in the NHPs tested (102-632 pg/mL). Reactivation of latent viral infection could account for the higher concentrations observed in NHPs than in humans, or perhaps this cytokine is regulated differently in NHPs than in humans.

MIP-1α is important in attracting monocytes and B cells and contributes to local inflammation, in vivo. Therefore, it is not expected that this cytokine would be elevated in healthy individuals. In the NHPs tested here, the median MIP-1α level was determined to be 8 pg/mL and the range of results spanned from undetectable levels to 256 pg/mL (which was statistically identified as an outlier). Reported values in humans, however, are higher than these NHP values: Millipore noted a range of 64 – 439 pg/mL and Ding et al. showed an average value for healthy individuals as 118 pg/mL [100]. It is possible that the human subjects tested for MIP-1α could have been experiencing some inflammatory event without showing outward symptoms of disease. Interestingly, depression has been associated with chronic inflammation [101] and a 2004 study showed that at least some depressed patients had elevated MIP-1α levels [102], which could account for some of the elevated levels reported in humans.

The soluble CD40 ligand (sCD40L) is a platelet-derived protein that is involved in inflammation. While Millipore reports a human range of this cytokine from <3.2 – 6.4 pg/mL, all of the NHPs showed elevated levels, from 188 - > 10,000 pg/mL. Wolf et al. reported significantly increased plasma concentration of this cytokine in HIV infected patients (mean 2,300 pg/mL) than in control subjects (mean 700 pg/mL), due to activation of platelets (which produce sCD40L) [103]. It has also been shown that Epstein-Barr virus (EBV) activates platelets [104]. Therefore, if any of the NHPs were experiencing asymptomatic reactivation of EBV or a similar virus, it could account for the high sCD40L concentrations observed in this study.

Vascular endothelial growth factor (VEGF) is important in regulating the formation of new blood vessels and is necessary for wound healing. While Millipore reported low to moderate levels of VEGF in humans (<3.2 – 118 pg/mL), the NHPs generated an overall range of <0.64 – 1,798 pg/mL, and a median value of 10 pg/mL. All of the animals studied had vascular access ports (VAP) surgically implanted, possibly leading one to posit this as the cause of the higher VEGF concentrations in these animals. However, VAP placement date did not correlate with higher VEGF concentrations; for the five animals with VEGF concentrations determined to be outliers (from 202-1,798 pg/mL), VAP placement ranged from 1 day to 14 months prior. The cause of the wide range of VEGF concentrations in NHPs and the discrepancy with human VEGF concentrations will require further study.

It may be tempting to use the data generated from these 36 (34 male, 2 female) healthy NHPs to suggest a normal reference interval for the cytokines found to be accurately measured in cynomolgus macaques. However, the recommended best practice for establishing a *de novo* reference interval (RI) is to enroll at least 120 reference individuals [69] and construct the RI using a simple nonparametric method. When 120 subjects who meet inclusion criteria cannot be obtained, other statistical approaches to determining a reference interval may be used and include traditional parametric, robust and bootstrap methods. The Clinical and Laboratory Standards Institute recommends the use of the robust method, when less than 120 subjects can be obtained, as it does not require as many reference subjects as the nonparametric method and does not require the observed values to follow a Gaussian distribution. There is no recommended minimum number of reference subjects when using the robust method but as the sample size decreases, the confidence intervals around the reference interval limits will increase, introducing uncertainty. Therefore, if 120 observations are not possible, as many observations as possible should be collected to minimize uncertainty. The Clinical Laboratory and Standards Institute is "hesitant to recommend" using even the robust method when less than 80 reference individuals can be obtained. A study supporting this recommendation was published in 2009 by Geffre et al [105]. In that study, twenty randomly selected subsets (with n=27) were selected out of a larger population of 1439 canines and RIs were calculated (for plasma creatinine) using the recommended robust method after transformation to a Gaussian distribution. These small, subset RIs were

compared to the RI for the complete set and were sometimes found to be "grossly erroneous." For this reason, the authors concluded that all values should be reported graphically (in a dot plot or histogram) for small sample sizes.

Another reason reference intervals were not established in the current study is because no attempt was made at true randomization of samples collected from the healthy cynomolgus macaques; instead we maximized the contribution of healthy animals already screened and enrolled in an islet transplantation study.

Data presented here suggest that intra animal variability in cytokine concentration may be significant. Replicate samples were analyzed, when present (n=11 animals), to determine the average CV and average absolute fold-change per animal. The intra-animal CVs ranged from 16% to 86% and the fold change ranged from 1.3 to 19.1. These data are provided in Appendix E for illustration purposes. This variability exceeds the average inter-assay CV (<15% for all cytokines except TNFα with an intra-assay CV of 45%), leading to the conclusion that the variations seen in monkeys tested over time are not simply an artifact of normal assay variation, but rather stem from biological variation. A sample size of only two observations per animal gives a general confidence level of 62-80%. To determine true intra-animal variability with 95% confidence, a minimum of 15 observations per animal would be needed (using broad assumptions of a population of 1000 and a confidence interval of 25 pg/mL). A large degree of inter-individual [106] and intra-individual [81] variability in cytokines among healthy populations has been

reported (IL-1ra ranging from undetectable levels to >1,500 pg/mL in similarly aged people, for example [106]). Klein et al [107] suggests that meaningful results may need to be individualized (establishing a baseline "normal" for each individual before comparing to diseased states) because of the large amount of variation seen between individuals. If such variability is normally present, population based RI will be very wide and potentially useless for detecting significant changes. If a reference interval is large enough to encompass the normal inter-individual variability it may be insensitive to significant intra-individual variations [108]. Therefore, a test result within the population-based reference interval may be abnormal for that particular individual. In cases like this, a subject-based RI may be more appropriate. It is possible to determine mathematically whether a population- based or subject- based RI should be used for specific analytes by calculating the 'index of individuality' [109]. The index of individuality is determined using a simple calculation:

$$(CV_I^2 + CV_A^2)^{1/2} / CV_G$$

Where CV_I is the individual, or, intra-subject variation (as a coefficient of variation), CV_A is the analytical variation and CV_G and is the group, or, inter-subject variation. In the study reported here, the calculation of analytical variation has been performed. The important next step is to investigate the intra- and inter-animal variability of these cytokines in healthy cynomolgus macaques to complete the equations that will help determine whether population-or subject-based RI are the most effective way to evaluate cytokine data in this species.

Because many cytokines act in an autocrine or paracrine manner, it may be of interest to understand cytokine concentrations in specific microenvironments. For example, pancreatic islet cells are sometimes placed under the renal capsule of transplant recipients. It could be useful to understand how cytokine concentrations may be affecting the outcome of graft placement in that specific location. Likewise, peritoneal fluid, bone marrow or liver may also be sites of interest. Procuring usable samples from these sites would, of course, be more difficult than a simple blood draw. However, it is possible that less variability would be observed in microenvironments than has been shown systemically since fluctuations at other sites throughout the body would have a limited impact. It would be important to validate the assay for use in any of these microenvironments as the different matrices may alter assay results. Baseline sampling of these microenvironments would be necessary to interpret changes after transplantation.

Limitations of the study

There are study limitations to consider in this work. In order to maximize useful data from an animal cohort enrolled in an islet transplantation study, true randomization of animals was not possible. Therefore, the cytokine distributions presented represent a well-defined cohort of captive cynomolgus macaques rather than a description of the population. As mentioned previously, a magnetic bead kit was used for the validation studies while a polystyrene bead kit was used to test samples from the 36 NHPs. However, 22 of the 23 antibody pairs were identical (other than IL-10) in both kits so it is

likely that similar results would be obtained using both kits. Previous studies have suggested that the best use of multiplexed kits, where a loss of accuracy is generally offset by an abundance of data, may be in monitoring trends rather than describing absolute values. If that is the case then the best use of this kit may be in comparing healthy to diseased states of animals. The data collected in this study was from healthy cynomolgus macaques so conclusions cannot be made about the performance of this kit in comparing changes from healthy to diseased states, or vice versa. Rather the conclusions found in this study are related only to the performance of this kit in serum from healthy cynomolgus macaques. Lastly, although the data presented here suggest that intra animal variability in cytokine concentration may be significant, it was not possible to make confident conclusions related to this within-animal cytokine variability due to the limited number of replicate samples from individual animals.

Conclusions

The results reported here demonstrate that the Millipore multiplex method can be used to reliably measure the following cytokines: TGFα, IFNγ, IL-2, IL-15, IL-13, IL-5, IL-6, IL-8, MCP-1, IL-12/23(p40) and VEGF. It can also be used for the cytokines that were found to be 'marginally accepted' (G-CSF, IL-10, sCD40L, IL-4 and MIP-1α) if the consistency limitations are kept in mind. Data for seven cytokines has been found to be unreliable based on their inability to meet acceptance criteria (GM-CSF, IL-17A, IL-1ra,

IL-1β, TNFα, MIP-1β and IL-18). The ability of the method to detect cytokine concentrations in healthy cynomolgus macaques has also been confirmed.

This thesis described the error associated with the Millipore Milliplex NHP cytokine assay which demonstrated that seven out of 23 cytokines included in the panel did not meet pre-established acceptance criteria. This validation ensures that the collection of inaccurate data, the determination of erroneous research conclusions, and the dissemination of unreliable research are minimized. Data presented here will be communicated to researchers using this kit to monitor cytokine status in healthy cynomolgus macaque serum. However, our role is to establish the error associated with this assay and communicate these results, not to enforce how a researcher decides to use this information in her own study. Limited data describing serum cytokine concentrations are available in healthy cynomolgus macaques. However, this thesis included the reliable measurement of 16 cytokines in serum from 36 healthy animals. This work determined that significant cytokine variability exists within and between individual animals and provides a stepping stone for future work in establishing population- or subject-based reference intervals for cynomolgus macaques.

Recommendations for using this method include:

- Each animal should act as its own control since there appears to be a high degree of variability between animals.

- Repeated samples from individual animals may be of use as individual cytokines may vary widely in animals. Repeat sampling should be done at regular intervals and at the same time of day to avoid confounding diurnal variation.
- Validation should be completed for each sample type (serum, peritoneal fluid, synovial fluid, etc.) that will be tested in order to account for the potential of significant matrix effects.

Thesis Summary

Cytokines are a complex network of highly related proteins that regulate processes such as inflammation, basal immune function and cell growth and development. If cytokines are to be measured and used to make conclusions on research studies, we must understand the variability around the measurement and the potential for errors associated with laboratory testing. As methods become available we must evaluate their performance before we use them to make research conclusions. Method validation procedures generate data that reflect assay performance and identify the inherent measurement uncertainty, allowing data to be accurately interpreted. This study included the design and performance of a multiplex assay validation assessment for the simultaneous measurement of 23 cytokines in NHP serum. The validation included analysis of serum samples from 36 healthy cynomolgus macaques in order to determine if the method could be used to detect cytokine concentrations in healthy NHP. Although sensitivity and precision data were generally acceptable, recovery and linearity data were

highly variable. Of the 23 cytokines tested, 11 met acceptance criteria, 5 were marginal and 7 were rejected. If the validation studies had not been performed, false conclusions could be made (overestimated or underestimated cytokine values) contributing to errors within the Total Testing Process. It is clear after performing this validation study, that clinical and research investment in multi-analyte methods is not warranted without demonstration of method validity for each analyte of interest.

The validity of the multiplex method for analysis of specific cytokines in serum has been confirmed and the ability of the method to detect cytokine concentrations in healthy cynomolgus macaques has also been demonstrated. Therefore, the potential for analytical error due to use of an un-validated method has been reduced for these analytes. It would be appropriate to continue to evaluate the performance of this method in alternate matrices (saliva, urine, peritoneal fluid, bone marrow) or in the current matrix (serum) with a protease inhibitor. In addition, studies should be designed to evaluate the impact of pre-analytical error (e.g. storage time and temperature, blood collection and centrifugation steps) on the measurement of cytokines in serum from cynomolgus macaques. Because this work demonstrates significant variability among measurements within and between animals, other important next steps include the description of typical variation levels within and between animals, followed by the generation of appropriate (population- or subject-based) reference intervals for healthy animals.

Although there are clearly some issues to be considered when validating and using multiplex assays, they are likely to remain an attractive option for researchers so it is in our best interest, as a scientific community, to understand and share data on assay performance.

Bibliography

[1] H.-F. Chen, “Defective Production of Interleukin-11 by Decidua and Chorionic Villi in Human Anembryonic Pregnancy,” *J. Clin. Endocrinol. Metab.*, vol. 87, no. 5, pp. 2320–2328, May 2002.

[2] A. Kamkin and I. Kiseleva, Eds., *Mechanical Stretch and Cytokines*. London, England: Springer, 2012, p. 236.

[3] H.-L. Wong, R. M. Pfeiffer, T. R. Fears, R. Vermeulen, S. Ji, and C. S. Rabkin, “Reproducibility and correlations of multiplex cytokine levels in asymptomatic persons.,” *Cancer Epidemiol. Biomarkers Prev.*, vol. 17, no. 12, pp. 3450–6, Dec. 2008.

[4] M. B. Sporn and A. B. Roberts, “Peptide growth factors are multifunctional,” *Nature*, vol. 332, no. 6161, pp. 217–219, 1988.

[5] A. W. Thomson and M. T. Lotze, Eds., *The Cytokine Handbook*, 4th ed. Amsterdam, Netherlands: Academic Press, 2003, p. 1572.

[6] S. Cohen, P. E. Bigazzi, and T. Yoshida, “Similarities of T Cell Function In Cell-Mediated Immunity and Antibody Production,” *Cell. Immunol.*, vol. 12, no. 1, pp. 150–159, 1974.

[7] M. J. Waters, C. A. Shang, S. N. Behncken, S. P. Tam, H. Li, B. Shen, and P. E. Lobie, “Growth hormone as a cytokine.,” *Clin. Exp. Pharmacol. Physiol.*, vol. 26, no. 10, pp. 760–4, Oct. 1999.

[8] L. Ferrero-Miliani, O. H. Nielsen, P. S. Andersen, and S. E. Girardin, “Chronic inflammation: importance of NOD2 and NALP3 in interleukin-1beta generation.,” *Clin. Exp. Immunol.*, vol. 147, no. 2, pp. 227–35, Feb. 2007.

[9] B. Conti, S. Sugama, Y. Kim, C. Tinti, H. Kim, H. Baker, B. Volpe, B. Attardi, and T. Joh, “Modulation of IL-18 production in the adrenal cortex following acute ACTH or chronic corticosterone treatment.,” *Neuroimmunomodulation*, vol. 8, no. 1, pp. 1–7, Jan. 2000.

[10] E. Nagy and I. Berczi, “Immunodeficiency in hypophysectomized rats.,” *Acta Endocrinol. (Copenh).*, vol. 89, no. 3, pp. 530–7, Nov. 1978.

[11] H. Tilg and A. R. Moschen, “Adipocytokines: mediators linking adipose tissue, inflammation and immunity.,” *Nat. Rev. Immunol.*, vol. 6, no. 10, pp. 772–83, Oct. 2006.

[12] X. Chen, X. Jia, J. Qiao, Y. Guan, and J. Kang, “Adipokines in reproductive function: a link between obesity and polycystic ovary syndrome.,” *J. Mol. Endocrinol.*, vol. 50, no. 2, pp. R21–37, Apr. 2013.

[13] O. Heizmann, M. Koeller, G. Muhr, D. Oertli, and C. Schinkel, “Th1- and Th2-type cytokines in plasma after major trauma.,” *J. Trauma*, vol. 65, no. 6, pp. 1374–8, Dec. 2008.

[14] R. Bertorello, M. P. Cordone, P. Contini, P. Rossi, F. Indiveri, F. Puppo, and G. Cordone, “Increased levels of interleukin-10 in saliva of Sjögren’s syndrome patients. Correlation with disease activity.,” *Clin. Exp. Med.*, vol. 4, no. 3, pp. 148–51, Dec. 2004.

[15] K. Tashiro, I. Koyanagi, A. Saitoh, A. Shimizu, T. Shike, C. Ishiguro, M. Koizumi, K. Funabiki, S. Horikoshi, I. Shirato, and Y. Tomino, “Urinary levels of monocyte chemoattractant protein-1 (MCP-1) and interleukin-8 (IL-8), and renal injuries in patients with type 2 diabetic nephropathy.,” *J. Clin. Lab. Anal.*, vol. 16, no. 1, pp. 1–4, Jan. 2002.

[16] S. Abke, M. Neumeier, J. Weigert, G. Wehrwein, E. Eggenhofer, A. Schäffler, K. Maier, C. Aslanidis, J. Schölmerich, and C. Buechler, “Adiponectin-induced secretion of interleukin-6 (IL-6), monocyte chemotactic protein-1 (MCP-1, CCL2) and interleukin-8 (IL-8, CXCL8) is impaired in monocytes from patients with type I diabetes.,” *Cardiovasc. Diabetol.*, vol. 5, p. 17, Jan. 2006.

[17] J.-N. Sheu, M.-C. Chen, K.-H. Lue, S.-L. Cheng, I.-C. Lee, S.-M. Chen, and G. J. Tsay, “Serum and urine levels of interleukin-6 and interleukin-8 in children with acute pyelonephritis.,” *Cytokine*, vol. 36, no. 5–6, pp. 276–82, Dec. 2006.

[18] J. K. Farney, L. K. Mamedova, B. H. Godsey, and B. J. Bradford, “Technical note: validation of an ELISA for measurement of tumor necrosis factor alpha in bovine plasma.,” *J. Dairy Sci.*, vol. 94, no. 7, pp. 3504–9, Jul. 2011.

[19] C. Nirmala, R. Trivedi, P. V. Prasanthi, R. Mullangi, and N. R. Srinivas, “Development and validation of a sensitive ELISA for quantitation of Grastim (rhG-CSF) in rat plasma: Application to a pharmacokinetic study.,” *Biomed. Chromatogr.*, vol. 20, no. 9, pp. 945–50, Sep. 2006.

[20] K. Suzuki and A. Kubo, “Validation and application of an enzyme-linked immunoassay for human erythropoietin to canine serum.,” *Res. Vet. Sci.*, vol. 61, no. 1, pp. 13–6, Jul. 1996.

[21] C. Janot-Sardet, B. Assouline, R. Cheynier, M. Morre, and S. Beq, “A validated assay to measure soluble IL-7 receptor shows minimal impact of IL-7 treatment.,” *J. Immunol. Methods*, vol. 353, no. 1–2, pp. 115–23, Feb. 2010.

[22] N. Aziz, P. Nishanian, R. Mitsuyasu, R. Detels, and J. L. Fahey, “Variables that affect assays for plasma cytokines and soluble activation markers.,” *Clin. Diagn. Lab. Immunol.*, vol. 6, no. 1, pp. 89–95, Jan. 1999.

[23] P. Kern, C. J. Hemmer, J. Van Damme, H. J. Gruss, and M. Dietrich, “Elevated tumor necrosis factor alpha and interleukin-6 serum levels as markers for complicated Plasmodium falciparum malaria.,” *Am. J. Med.*, vol. 87, no. 2, pp. 139–43, Aug. 1989.

[24] V. Basic Kes, A.-M. Simundic, N. Nikolac, E. Topic, and V. Demarin, “Pro-inflammatory and anti-inflammatory cytokines in acute ischemic stroke and their relation to early neurological deficit and stroke outcome.,” *Clin. Biochem.*, vol. 41, no. 16–17, pp. 1330–4, Nov. 2008.

[25] E. Vairaktaris, C. Yapijakis, Z. Serefoglou, D. Avgoustidis, E. Critselis, S. Spyridonidou, A. Vylliotis, S. Derka, S. Vassiliou, E. Nkenke, and E. Patsouris, “Gene expression polymorphisms of interleukins-1 beta, -4, -6, -8, -10, and tumor necrosis factors-alpha, -beta: regression analysis of their effect upon oral squamous cell carcinoma.,” *J. Cancer Res. Clin. Oncol.*, vol. 134, no. 8, pp. 821–32, Aug. 2008.

[26] M. R. Walker, D. A. Makropoulos, R. Achuthanandam, S. Van Arsdell, and P. J. Bugelski, “Development of a human whole blood assay for prediction of cytokine release similar to anti-CD28 superagonists using multiplex cytokine and hierarchical cluster analysis.,” *Int. Immunopharmacol.*, vol. 11, no. 11, pp. 1697–705, Nov. 2011.

[27] K. L. Hutchinson, F. Villinger, M. E. Miranda, T. G. Ksiazek, C. J. Peters, and P. E. Rollin, “Multiplex analysis of cytokines in the blood of cynomolgus macaques naturally infected with Ebola virus (Reston serotype).,” *J. Med. Virol.*, vol. 65, no. 3, pp. 561–6, Nov. 2001.

[28] A. Biancotto, X. Feng, M. Langweiler, N. S. Young, and J. P. McCoy, “Effect of anticoagulants on multiplexed measurement of cytokine/chemokines in healthy subjects.,” *Cytokine*, vol. 60, no. 2, pp. 438–46, Nov. 2012.

[29] D. J. Stearns-Kurosawa, V. Collins, S. Freeman, V. L. Tesh, and S. Kurosawa, "Distinct physiologic and inflammatory responses elicited in baboons after challenge with Shiga toxin type 1 or 2 from enterohemorrhagic Escherichia coli.," *Infect. Immun.*, vol. 78, no. 6, pp. 2497–504, Jun. 2010.

[30] S. Lawson, J. Lunney, F. Zuckermann, F. Osorio, E. Nelson, C. Welbon, T. Clement, Y. Fang, S. Wong, K. Kulas, and J. Christopher-Hennings, "Development of an 8-plex Luminex assay to detect swine cytokines for vaccine development: assessment of immunity after porcine reproductive and respiratory syndrome virus (PRRSV) vaccination.," *Vaccine*, vol. 28, no. 32, pp. 5356–64, Jul. 2010.

[31] L. D. Giavedoni, "Simultaneous detection of multiple cytokines and chemokines from nonhuman primates using luminex technology.," *J. Immunol. Methods*, vol. 301, no. 1–2, pp. 89–101, Jun. 2005.

[32] J. M. Tarrant, "Blood cytokines as biomarkers of in vivo toxicity in preclinical safety assessment: considerations for their use.," *Toxicol. Sci.*, vol. 117, no. 1, pp. 4–16, Oct. 2010.

[33] P. Horan and L. Wheeless Jr, "Quantitative single cell analysis and sorting," *Science*, vol. 198, no. 4313, pp. 149–57, 1977.

[34] B. P. Hanley, L. Xing, and R. H. Cheng, "Variance in multiplex suspension array assays: microsphere size variation impact.," *Theor. Biol. Med. Model.*, vol. 4, p. 31, Jan. 2007.

[35] A. A. Ellington, I. J. Kullo, K. R. Bailey, and G. G. Klee, "Antibody-based protein multiplex platforms: technical and operational challenges," *Clin. Chem.*, vol. 56, no. 2, pp. 186–193, 2010.

[36] W. de Jager, H. te Velthuis, B. J. Prakken, W. Kuis, and G. T. Rijkers, "Simultaneous detection of 15 human cytokines in a single sample of stimulated peripheral blood mononuclear cells.," *Clin. Diagn. Lab. Immunol.*, vol. 10, no. 1, pp. 133–9, Jan. 2003.

[37] A. M. Ring, J.-X. Lin, D. Feng, S. Mitra, M. Rickert, G. R. Bowman, V. S. Pande, P. Li, I. Moraga, R. Spolski, E. Ozkan, W. J. Leonard, and K. C. Garcia, "Mechanistic and structural insight into the functional dichotomy between IL-2 and IL-15.," *Nat. Immunol.*, vol. 13, no. 12, pp. 1187–95, Dec. 2012.

[38] M. Y. Liu, A. M. Xydakis, R. C. Hoogeveen, P. H. Jones, E. O. Smith, K. W. Nelson, and C. M. Ballantyne, "Multiplexed analysis of biomarkers related to

obesity and the metabolic syndrome in human plasma, using the Luminex-100 system.," *Clin. Chem.*, vol. 51, no. 7, pp. 1102–9, Jul. 2005.

[39] S. S. Khan, M. S. Smith, D. Reda, A. F. Suffredini, and J. P. McCoy, "Multiplex bead array assays for detection of soluble cytokines: comparisons of sensitivity and quantitative values among kits from multiple manufacturers.," *Cytometry B. Clin. Cytom.*, vol. 61, no. 1, pp. 35–9, Sep. 2004.

[40] M. F. Elshal and J. P. McCoy, "Multiplex bead array assays: performance evaluation and comparison of sensitivity to ELISA.," *Methods*, vol. 38, no. 4, pp. 317–23, Apr. 2006.

[41] C. a Ray, R. R. Bowsher, W. C. Smith, V. Devanarayan, M. B. Willey, J. T. Brandt, and R. a Dean, "Development, validation, and implementation of a multiplex immunoassay for the simultaneous determination of five cytokines in human serum.," *J. Pharm. Biomed. Anal.*, vol. 36, no. 5, pp. 1037–44, Jan. 2005.

[42] J. L. Richens, R. a Urbanowicz, R. Metcalf, J. Corne, P. O'Shea, and L. Fairclough, "Quantitative validation and comparison of multiplex cytokine kits.," *J. Biomol. Screen.*, vol. 15, no. 5, pp. 562–8, Jun. 2010.

[43] L. Dossus, S. Becker, D. Achaintre, R. Kaaks, and S. Rinaldi, "Validity of multiplex-based assays for cytokine measurements in serum and plasma from 'non-diseased' subjects: comparison with ELISA.," *J. Immunol. Methods*, vol. 350, no. 1–2, pp. 125–32, Oct. 2009.

[44] E. C. Breen, S. M. Reynolds, C. Cox, L. P. Jacobson, L. Magpantay, C. B. Mulder, O. Dibben, J. B. Margolick, J. H. Bream, E. Sambrano, O. Martínez-Maza, E. Sinclair, P. Borrow, A. L. Landay, C. R. Rinaldo, and P. J. Norris, "Multisite comparison of high-sensitivity multiplex cytokine assays.," *Clin. Vaccine Immunol.*, vol. 18, no. 8, pp. 1229–42, Aug. 2011.

[45] G. D. Lundberg, "Acting on significant laboratory results.," *JAMA*, vol. 245, no. 17, pp. 1762–3, May 1981.

[46] P. Riches, R. Gooding, B. C. Millar, and A. W. Rowbottom, "Influence of collection and separation of blood samples on plasma IL-1, IL-6 and TNF-alpha concentrations.," *J. Immunol. Methods*, vol. 153, no. 1–2, pp. 125–31, Aug. 1992.

[47] R. P. Jackman, G. H. Utter, J. W. Heitman, D. F. Hirschkorn, J. P. Law, N. Gefter, M. P. Busch, and P. J. Norris, "Effects of blood sample age at time of separation on measured cytokine concentrations in human plasma.," *Clin. Vaccine Immunol.*, vol. 18, no. 2, pp. 318–26, Feb. 2011.

[48] E. Hooijberg, E. Leidinger, and K. P. Freeman, “An error management system in a veterinary clinical laboratory.,” *J. Vet. Diagn. Invest.*, vol. 24, no. 3, pp. 458–68, May 2012.

[49] R. Hawkins, “Managing the pre- and post-analytical phases of the total testing process.,” *Ann. Lab. Med.*, vol. 32, no. 1, pp. 5–16, Jan. 2012.

[50] M. Koulmanda, A. Qipo, Z. Fan, N. Smith, H. Auchincloss, X. X. Zheng, and T. B. Strom, “Prolonged survival of allogeneic islets in cynomolgus monkeys after short-term triple therapy.,” *Am. J. Transplant*, vol. 12, no. 5, pp. 1296–302, May 2012.

[51] M. Arikata, Y. Itoh, M. Okamatsu, T. Maeda, T. Shiina, K. Tanaka, S. Suzuki, M. Nakayama, Y. Sakoda, H. Ishigaki, A. Takada, H. Ishida, K. Soda, V. L. Pham, H. Tsuchiya, S. Nakamura, R. Torii, T. Shimizu, H. Inoko, I. Ohkubo, H. Kida, and K. Ogasawara, “Memory immune responses against pandemic (H1N1) 2009 influenza virus induced by a whole particle vaccine in cynomolgus monkeys carrying Mafa-A1*052:02.,” *PLoS One*, vol. 7, no. 5, p. e37220, Jan. 2012.

[52] L. Kari, W. M. Whitmire, D. D. Crane, N. Reveneau, J. H. Carlson, M. M. Goheen, E. M. Peterson, S. Pal, L. M. de la Maza, and H. D. Caldwell, “Chlamydia trachomatis native major outer membrane protein induces partial protection in nonhuman primates: implication for a trachoma transmission-blocking vaccine.,” *J. Immunol.*, vol. 182, no. 12, pp. 8063–70, Jun. 2009.

[53] J. W. Lee, V. Devanarayan, Y. C. Barrett, R. Weiner, J. Allinson, S. Fountain, S. Keller, I. Weinryb, M. Green, L. Duan, J. a Rogers, R. Millham, P. J. O’Brien, J. Sailstad, M. Khan, C. Ray, and J. a Wagner, “Fit-for-purpose method development and validation for successful biomarker measurement.,” *Pharm. Res.*, vol. 23, no. 2, pp. 312–28, Feb. 2006.

[54] M. Wadman, “NIH mulls rules for validating key results.,” *Nature*, vol. 500, no. 7460, pp. 14–6, Aug. 2013.

[55] J. F. Djoba Siawaya, T. Roberts, C. Babb, G. Black, H. J. Golakai, K. Stanley, N. B. Bapela, E. Hoal, S. Parida, P. van Helden, and G. Walzl, “An evaluation of commercial fluorescent bead-based luminex cytokine assays.,” *PLoS One*, vol. 3, no. 7, p. e2535, Jan. 2008.

[56] U. Prabhakar, E. Eirikis, M. Reddy, E. Silvestro, S. Spitz, C. Pendley, H. M. Davis, and B. E. Miller, “Validation and comparative analysis of a multiplexed assay for the simultaneous quantitative measurement of Th1/Th2 cytokines in

human serum and human peripheral blood mononuclear cell culture supernatants.," *J. Immunol. Methods*, vol. 291, no. 1–2, pp. 27–38, Aug. 2004.

[57] F. Chowdhury, A. Williams, and P. Johnson, "Validation and comparison of two multiplex technologies, Luminex and Mesoscale Discovery, for human cytokine profiling.," *J. Immunol. Methods*, vol. 340, no. 1, pp. 55–64, Jan. 2009.

[58] J. O. Westgard, *Basic Method Validation*, 3rd ed. Madison, WI: Westgard QC, Inc., 2008, p. 320.

[59] "Guidance for Industry Bioanalytical Method Validation," 2013.

[60] A. Malekzadeh, V. de Groot, H. Beckerman, B. W. van Oosten, M. a Blankenstein, and C. Teunissen, "Challenges in multi-plex and mono-plex platforms for the discovery of inflammatory profiles in neurodegenerative diseases.," *Methods*, vol. 56, no. 4, pp. 508–13, May 2012.

[61] C. M. Helms, I. Messaoudi, S. Jeng, W. M. Freeman, K. E. Vrana, and K. A. Grant, "A longitudinal analysis of circulating stress-related proteins and chronic ethanol self-administration in cynomolgus macaques.," *Alcohol. Clin. Exp. Res.*, vol. 36, no. 6, pp. 995–1003, Jun. 2012.

[62] W. M. Freeman, A. C. Salzberg, S. W. Gonzales, K. A. Grant, and K. E. Vrana, "Classification of alcohol abuse by plasma protein biomarkers.," *Biol. Psychiatry*, vol. 68, no. 3, pp. 219–22, Aug. 2010.

[63] C. He, P. K. Narayanan, and M. M. Fort, "Assessment of the Performance of Three Multiplex Array Panels for the Detection of Circulating Cytokines and Chemokines in Naive, LPS, and SEB-treated Cynomolgus Macaques.," *Toxicol. Pathol.*, Nov. 2013.

[64] J. W. Lee, R. S. Weiner, J. M. Sailstad, R. R. Bowsher, D. W. Knuth, P. J. O'Brien, J. L. Fourcroy, R. Dixit, L. Pandite, R. G. Pietrusko, H. D. Soares, V. Quarmby, O. L. Vesterqvist, D. M. Potter, J. L. Witliff, H. A. Fritche, T. O'Leary, L. Perlee, S. Kadam, and J. A. Wagner, "Method validation and measurement of biomarkers in nonclinical and clinical samples in drug development: a conference report.," *Pharm. Res.*, vol. 22, no. 4, pp. 499–511, Apr. 2005.

[65] "Product Information Sheet MagPlex Microshperes," 2012. [Online]. Available: http://www.luminexcorp.com/prod/groups/public/documents/lmnxcorp/89-60000-00-049.pdf. [Accessed: 23-Apr-2013].

[66] “Luminex xMAP Technology,” 2012. [Online]. Available: http://www.luminexcorp.com/TechnologiesScience/xMAPTechnology/. [Accessed: 07-Jun-2012].

[67] T. Hendrich, “Personal communication.” Email, 2012.

[68] M. L. Graham, E. F. Rieke, M. Dunning, L. a Mutch, A. M. Craig, E. K. Zolondek, B. J. Hering, H.-J. Schuurman, and R. W. Bianco, “A novel alternative placement site and technique for totally implantable vascular access ports in non-human primates.,” *J. Med. Primatol.*, vol. 38, no. 3, pp. 204–12, Jun. 2009.

[69] “Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline - Third Edition,” 2008.

[70] R. Thiébaut and H. Jacqmin-Gadda, “Mixed models for longitudinal left-censored repeated measures.,” *Comput. Methods Programs Biomed.*, vol. 74, no. 3, pp. 255–60, Jun. 2004.

[71] V. E. C. J. Schijns and M. C. Horzinek, Eds., *Cytokines in Veterinary Medicine*. Wallingford, UK: CABI, 1997, p. 352.

[72] M. Banerjee, “Personal communication.” Email, 2013.

[73] Millipore, “Milliplex Human Cytokine / Chemokine 96-Well Plate Assay.” Kit insert.

[74] S. X. Leng, J. E. McElhaney, J. D. Walston, D. Xie, N. S. Fedarko, and G. A. Kuchel, “ELISA and multiplexing technologies for cytokine measurements in inflammation and aging research,” *journals Gerontol. Ser. A, Biol. Sci. Med. Sci.*, vol. 63, no. 8, pp. 879–884, 2008.

[75] K. Kofoed, U. V. Schneider, T. Scheel, O. Andersen, and J. Eugen-Olsen, “Development and validation of a multiplex add-on assay for sepsis biomarkers using xMAP technology.,” *Clin. Chem.*, vol. 52, no. 7, pp. 1284–93, Jul. 2006.

[76] J. A. Bastarache, T. Koyama, N. E. Wickersham, D. B. Mitchell, R. L. Mernaugh, and L. B. Ware, “Accuracy and reproducibility of a multiplex immunoassay platform: A validation study,” *J. Immunol. Methods*, vol. 367, no. 1–2, pp. 33–39, Mar. 2011.

[77] K. James, I. Milne, A. Cunningham, and S. F. Elliott, “The effect of alpha 2 macroglobulin in commercial cytokine assays.,” *J. Immunol. Methods*, vol. 168, no. 1, pp. 33–7, Jan. 1994.

[78] M. B. Hansen, M. Svenson, M. Diamant, and K. Bendtzen, “High-affinity IgG autoantibodies to IL-6 in sera of normal individuals are competitive inhibitors of IL-6 in vitro.,” *Cytokine*, vol. 5, no. 1, pp. 72–80, Jan. 1993.

[79] N. Mae, D. J. Liberato, R. Chizzonite, and H. Satoh, “Identification of high-affinity anti-IL-1 alpha autoantibodies in normal human serum as an interfering substance in a sensitive enzyme-linked immunosorbent assay for IL-1 alpha.,” *Lymphokine Cytokine Res.*, vol. 10, no. 1–2, pp. 61–8, Apr. 1991.

[80] P. Seckinger, J. H. Zhang, B. Hauptmann, and J. M. Dayer, “Characterization of a tumor necrosis factor alpha (TNF-alpha) inhibitor: evidence of immunological cross-reactivity with the TNF receptor.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 87, no. 13, pp. 5188–92, Jul. 1990.

[81] A.-M. Tsimberidou, T. Waddelow, H. M. Kantarjian, M. Albitar, and F. J. Giles, “Pilot study of recombinant human soluble tumor necrosis factor (TNF) receptor (p75) fusion protein (TNFR:Fc; Enbrel) in patients with refractory multiple myeloma: increase in plasma TNFα levels during treatment,” *Leuk. Res.*, vol. 27, no. 5, pp. 375–380, May 2003.

[82] J. D. Eason, M. Pascual, S. Wee, M. Farrell, J. Phelan, S. Boskovic, C. Blosch, K. M. Mohler, and A. B. Cosimi, “Evaluation of recombinant human soluble dimeric tumor necrosis factor receptor for prevention of OKT3-associated acute clinical syndrome.,” *Transplantation*, vol. 61, no. 2, pp. 224–8, Jan. 1996.

[83] A. R. Rudnicka, A. Rumley, G. D. O. Lowe, and D. P. Strachan, “Diurnal, seasonal, and blood-processing patterns in levels of circulating fibrinogen, fibrin D-dimer, C-reactive protein, tissue plasminogen activator, and von Willebrand factor in a 45-year-old population.,” *Circulation*, vol. 115, no. 8, pp. 996–1003, Feb. 2007.

[84] R. P. Doe, J. A. Vennes, and E. B. Flink, “Diurnal variation of 17-hydroxycorticosteroids, sodium, potassium, magnesium and creatinine in normal subjects and in cases of treated adrenal insufficiency and Cushing’s syndrome.,” *J. Clin. Endocrinol. Metab.*, vol. 20, pp. 253–65, Feb. 1960.

[85] H. K. Nielsen, P. Laurberg, K. Brixen, and L. Mosekilde, “Relations between diurnal variations in serum osteocalcin, cortisol, parathyroid hormone, and ionized calcium in normal individuals,” *Eur. J. Endocrinol.*, vol. 124, no. 4, pp. 391–398, Apr. 1991.

[86] N. Petrovsky, P. McNair, and L. C. Harrison, “Diurnal rhythms of pro-inflammatory cytokines: regulation by plasma cortisol and therapeutic implications.,” *Cytokine*, vol. 10, no. 4, pp. 307–12, Apr. 1998.

[87] F. Villinger, S. S. Brar, A. Mayne, N. Chikkala, and A. A. Ansari, “Comparative sequence analysis of cytokine genes from human and nonhuman primates.,” *J. Immunol.*, vol. 155, no. 8, pp. 3946–54, Oct. 1995.

[88] H. Northoff, C. Weinstock, and a Berg, “The cytokine response to strenuous exercise.,” *Int. J. Sports Med.*, vol. 15 Suppl 3, pp. S167–71, Oct. 1994.

[89] J. Doucet, J. Canadi, C. Kalis, M.-A. Valentin, S. Marrony, F. Deckert-Salva, F. Legay, and A. Avrameas, “Reduction of matrix interferences by the combination of chaotropic salt and DMSO in a broadly applicable target-based ELISA for pharmacokinetic studies of therapeutic monoclonal antibodies,” *J. Pharm. Biomed. Anal.*, vol. 50, no. 5, pp. 924–31, Dec. 2009.

[90] J. T. Dancey, K. A. Deubelbeiss, L. A. Harker, and C. A. Finch, “Neutrophil kinetics in man.,” *J. Clin. Invest.*, vol. 58, no. 3, pp. 705–15, Sep. 1976.

[91] M. Kawakami, H. Tsutsumi, T. Kumakawa, H. Abe, M. Hirai, S. Kurosawa, M. Mori, and M. Fukushima, “Levels of serum granulocyte colony-stimulating factor in patients with infections.,” *Blood*, vol. 76, no. 10, pp. 1962–4, Nov. 1990.

[92] G. J. McKenzie, P. G. Fallon, C. L. Emson, R. K. Grencis, and A. N. McKenzie, “Simultaneous disruption of interleukin (IL)-4 and IL-13 defines individual roles in T helper cell type 2-mediated responses.,” *J. Exp. Med.*, vol. 189, no. 10, pp. 1565–72, May 1999.

[93] A. G. Alexander, J. Barkans, R. Moqbel, N. C. Barnes, A. B. Kay, and C. J. Corrigan, “Serum interleukin 5 concentrations in atopic and non-atopic patients with glucocorticoid-dependent chronic severe asthma.,” *Thorax*, vol. 49, no. 12, pp. 1231–1233, Dec. 1994.

[94] T. Yoshimura, K. Matsushima, S. Tanaka, E. A. Robinson, E. Appella, J. J. Oppenheim, and E. J. Leonard, “Purification of a human monocyte-derived neutrophil chemotactic factor that has peptide sequence similarity to other host defense cytokines.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 84, no. 24, pp. 9233–7, Dec. 1987.

[95] A. Sun, J. T. Wang, J. S. Chia, and C. P. Chiang, “Serum interleukin-8 level is a more sensitive marker than serum interleukin-6 level in monitoring the disease

activity of oral lichen planus.," *Br. J. Dermatol.*, vol. 152, no. 6, pp. 1187–92, Jun. 2005.

[96] H. W. Zimmermann, S. Seidler, N. Gassler, J. Nattermann, T. Luedde, C. Trautwein, and F. Tacke, "Interleukin-8 is activated in patients with chronic liver diseases and associated with hepatic macrophage accumulation in human liver fibrosis.," *PLoS One*, vol. 6, no. 6, p. e21381, Jan. 2011.

[97] F. Mattner, L. Ozmen, F. J. Podlaski, V. L. Wilkinson, D. H. Presky, M. K. Gately, and G. Alber, "Treatment with homodimeric interleukin-12 (IL-12) p40 protects mice from IL-12-dependent shock but not from tumor necrosis factor alpha-dependent shock.," *Infect. Immun.*, vol. 65, no. 11, pp. 4734–7, Nov. 1997.

[98] F. H. van Dooren, N. W. Duijvis, and A. A. te Velde, "Analysis of cytokines and chemokines produced by whole blood, peripheral mononuclear and polymorphonuclear cells.," *J. Immunol. Methods*, vol. 396, no. 1–2, pp. 128–33, Oct. 2013.

[99] J. Wang, L. Nong, Y. Wei, S. Qin, Y. Zhou, and Y. Tang, "Association of interleukin-12 polymorphisms and serum IL-12p40 levels with osteosarcoma risk.," *DNA Cell Biol.*, vol. 32, no. 10, pp. 605–10, Oct. 2013.

[100] L. Ding, B. Li, Y. Zhao, Y.-F. Fu, E.-L. Hu, Q.-G. Hu, Y.-H. Ni, and Y.-Y. Hou, "Serum CCL2 and CCL3 as potential biomarkers for the diagnosis of oral squamous cell carcinoma," *Tumour Biol.*, Jul. 2014.

[101] R. Dantzer, J. C. O'Connor, G. G. Freund, R. W. Johnson, and K. W. Kelley, "From inflammation to sickness and depression: when the immune system subjugates the brain.," *Nat. Rev. Neurosci.*, vol. 9, no. 1, pp. 46–56, Jan. 2008.

[102] R. A. Merendino, G. Di Pasquale, F. De Luca, L. Di Pasquale, E. Ferlazzo, G. Martino, M. C. Palumbo, F. Morabito, and S. Gangemi, "Involvement of fractalkine and macrophage inflammatory protein-1 alpha in moderate-severe depression.," *Mediators Inflamm.*, vol. 13, no. 3, pp. 205–7, Jun. 2004.

[103] K. Wolf, D. A. Tsakiris, R. Weber, P. Erb, and M. Battegay, "Antiretroviral therapy reduces markers of endothelial and coagulation activation in patients infected with human immunodeficiency virus type 1.," *J. Infect. Dis.*, vol. 185, no. 4, pp. 456–62, Feb. 2002.

[104] A. Ahmad and J. Menezes, "Binding of the Epstein-Barr virus to human platelets causes the release of transforming growth factor-beta.," *J. Immunol.*, vol. 159, no. 8, pp. 3984–8, Oct. 1997.

[105] A. Geffré, J. P. Braun, C. Trumel, and D. Concordet, “Estimation of reference intervals from small samples: an example using canine plasma creatinine.,” *Vet. Clin. Pathol.*, vol. 38, no. 4, pp. 477–84, Dec. 2009.

[106] U. Sack, U. Burkhardt, M. Borte, H. Schädlich, K. Berg, and F. Emmrich, “Age-dependent levels of select immunological mediators in sera of healthy children.,” *Clin. Diagn. Lab. Immunol.*, vol. 5, no. 1, pp. 28–32, Jan. 1998.

[107] R. Klein, M. Schwenk, and D. M. Templeton, “Cytokine profiles in human exposure to metals (IUPAC Technical Report),” *Pure Appl. Chem.*, vol. 78, no. 11, pp. 2155–2168, 2006.

[108] S. Andersen, K. M. Pedersen, N. H. Bruun, and P. Laurberg, “Narrow individual variations in serum T(4) and T(3) in normal subjects: a clue to the understanding of subclinical thyroid disease.,” *J. Clin. Endocrinol. Metab.*, vol. 87, no. 3, pp. 1068–72, Mar. 2002.

[109] R. M. Walton, “Subject-based reference values: biological variation, individuality, and reference change values.,” *Vet. Clin. Pathol.*, vol. 41, no. 2, pp. 175–81, Jun. 2012.

[110] T. A. Waldmann, “The multi-subunit interleukin-2 receptor.,” *Annu. Rev. Biochem.*, vol. 58, pp. 875–911, Jan. 1989.

[111] M. Lenardo, K. M. Chan, F. Hornung, H. McFarland, R. Siegel, J. Wang, and L. Zheng, “Mature T lymphocyte apoptosis--immune regulation in a dynamic and unpredictable antigenic environment.,” *Annu. Rev. Immunol.*, vol. 17, pp. 221–53, Jan. 1999.

[112] B. Sadlack, H. Merz, H. Schorle, A. Schimpl, A. C. Feller, and I. Horak, “Ulcerative colitis-like disease in mice with a disrupted interleukin-2 gene,” *Cell*, vol. 75, no. 2, pp. 253–261, Oct. 1993.

[113] N. F. Adkinson Jr., B. S. Bochner, W. M. Busse, S. T. Holgate, R. F. Lemanske Jr., and F. E. R. Simons, Eds., *Middleton’s Allergy: Principles & Practice*, 7th ed. Philadelphia, PA: Mosby Elsevier, 2009.

[114] A. Waage, P. Brandtzaeg, A. Halstensen, P. Kierulf, and T. Espevik, “The complex pattern of cytokines in serum from patients with meningococcal septic shock. Association between interleukin 6, interleukin 1, and fatal outcome,” *J. Exp. Med.*, vol. 169, no. January, pp. 2–7, 1989.

[115] G. Agbanoma, C. Li, D. Ennis, A. C. Palfreeman, L. M. Williams, and F. M. Brennan, “Production of TNF-α in macrophages activated by T cells, compared with lipopolysaccharide, uses distinct IL-10-dependent regulatory mechanism.,” *J. Immunol.*, vol. 188, no. 3, pp. 1307–17, Feb. 2012.

[116] W. Sun, I. Tadmori, L. Yang, M. Delgado, and D. Ganea, “Vasoactive intestinal peptide (VIP) inhibits TGF-beta1 production in murine macrophages.,” *J. Neuroimmunol.*, vol. 107, no. 1, pp. 88–99, Jul. 2000.

[117] A. L. Moreira, E. P. Sampaio, S. A. Zmuidzinas, P. Frindt, K. A. Smith, and G. Kaplan, “Thalidomide Exerts Its Inhibitory Action on Tumor Necrosis Factor oe by Enhancing m RNA Degradation,” *J. Exp. Med.*, vol. 177, no. June, pp. 1675–80, 1993.

[118] E. Yang, E. van Nimwegen, M. Zavolan, N. Rajewsky, M. Schroeder, M. Magnasco, and J. E. Darnell, “Decay rates of human mRNAs: correlation with functional characteristics and sequence attributes.,” *Genome Res.*, vol. 13, no. 8, pp. 1863–72, Aug. 2003.

[119] P. W. Thavasu, S. Longhurst, S. P. Joel, M. L. Slevin, and F. R. Balkwill, “Measuring cytokine levels in blood. Importance of anticoagulants, processing, and storage conditions,” *J. Immunol. Methods*, vol. 153, no. 1–2, pp. 115–24, Aug. 1992.

[120] J. C. Oliver, L. A. Bland, C. W. Oettinger, M. J. Arduino, S. K. Mcallister, S. M. Aguero, and M. S. Favero, “Cytokine kinetics in an in vitro whole blood model following an endotoxin challenge,” *Lymphokine Cytokine Res.*, vol. 12, no. 2, pp. 115–120, 1993.

[121] K. Nakanishi, T. Yoshimoto, H. Tsutsui, and H. Okamura, “Interleukin-18 regulates both Th1 and Th2 responses.,” *Annu. Rev. Immunol.*, vol. 19, pp. 423–74, Jan. 2001.

[122] T. Yoshimoto, K. Takeda, T. Tanaka, K. Ohkusu, S. Kashiwamura, H. Okamura, S. Akira, and K. Nakanishi, “IL-12 up-regulates IL-18 receptor expression on T cells, Th1 cells, and B cells: synergism with IL-18 for IFN-gamma production.,” *J. Immunol.*, vol. 161, no. 7, pp. 3400–7, Oct. 1998.

[123] N. Ferrara, H. Chen, T. Davis-Smyth, H. P. Gerber, T. N. Nguyen, D. Peers, V. Chisholm, K. J. Hillan, and R. H. Schwall, “Vascular endothelial growth factor is essential for corpus luteum angiogenesis.,” *Nat. Med.*, vol. 4, no. 3, pp. 336–40, Mar. 1998.

[124] K. Gogat, L. Le Gat, L. Van Den Berghe, D. Marchant, A. Kobetz, S. Gadin, B. Gasser, I. Quéré, M. Abitbol, and M. Menasche, “VEGF and KDR gene expression during human embryonic and fetal eye development.,” *Invest. Ophthalmol. Vis. Sci.*, vol. 45, no. 1, pp. 7–14, Jan. 2004.

[125] E. Zelzer, R. Mamluk, N. Ferrara, R. S. Johnson, E. Schipani, and B. R. Olsen, “VEGFA is necessary for chondrocyte survival during bone development.,” *Development*, vol. 131, no. 9, pp. 2161–71, May 2004.

[126] D. J. Peacock, M. L. Banquerigo, and E. Brahn, “Angiogenesis inhibition suppresses collagen arthritis.,” *J. Exp. Med.*, vol. 175, no. 4, pp. 1135–8, Apr. 1992.

[127] A. P. Adamis, J. W. Miller, M. T. Bernal, D. J. D’Amico, J. Folkman, T. K. Yeo, and K. T. Yeo, “Increased vascular endothelial growth factor levels in the vitreous of eyes with proliferative diabetic retinopathy.,” *Am. J. Ophthalmol.*, vol. 118, no. 4, pp. 445–50, Oct. 1994.

[128] M. Detmar, L. F. Brown, K. P. Claffey, K. T. Yeo, O. Kocher, R. W. Jackman, B. Berse, and H. F. Dvorak, “Overexpression of vascular permeability factor/vascular endothelial growth factor and its receptors in psoriasis.,” *J. Exp. Med.*, vol. 180, no. 3, pp. 1141–6, Sep. 1994.

[129] N. S. Levy, S. Chung, H. Furneaux, and A. P. Levy, “Hypoxic stabilization of vascular endothelial growth factor mRNA by the RNA-binding protein HuR.,” *J. Biol. Chem.*, vol. 273, no. 11, pp. 6417–23, Mar. 1998.

[130] L. S. Angelo and R. Kurzrock, “Vascular endothelial growth factor and its relationship to inflammatory mediators.,” *Clin. Cancer Res.*, vol. 13, no. 10, pp. 2825–30, May 2007.

[131] K. Matsushima, K. Morishita, T. Yoshimura, S. Lavu, Y. Kobayashi, W. Lew, E. Appella, H. F. Kung, E. J. Leonard, and J. J. Oppenheim, “Molecular cloning of a human monocyte-derived neutrophil chemotactic factor (MDNCF) and the induction of MDNCF mRNA by interleukin 1 and tumor necrosis factor.,” *J. Exp. Med.*, vol. 167, no. 6, pp. 1883–93, Jun. 1988.

[132] E. DiMango, H. J. Zar, R. Bryan, and A. Prince, “Diverse Pseudomonas aeruginosa gene products stimulate respiratory epithelial cells to produce interleukin-8.,” *J. Clin. Invest.*, vol. 96, no. 5, pp. 2204–10, Nov. 1995.

[133] J. E. Crabtree, Z. Xiang, I. J. Lindley, D. S. Tompkins, R. Rappuoli, and A. Covacci, “Induction of interleukin-8 secretion from gastric epithelial cells by a

cagA negative isogenic mutant of Helicobacter pylori.," *J. Clin. Pathol.*, vol. 48, no. 10, pp. 967–9, Oct. 1995.

[134] X. Xu, H. Zhang, Q. Zhang, Y. Huang, J. Dong, Y. Liang, H.-J. Liu, and D. Tong, "Porcine epidemic diarrhea virus N protein prolongs S-phase cell cycle, induces endoplasmic reticulum stress, and up-regulates interleukin-8 expression.," *Vet. Microbiol.*, vol. 164, no. 3–4, pp. 212–21, Jun. 2013.

[135] T. Pollicino, L. Bellinghieri, A. Restuccia, G. Raffa, C. Musolino, A. Alibrandi, D. Teti, and G. Raimondo, "Hepatitis B virus (HBV) induces the expression of interleukin-8 that in turn reduces HBV sensitivity to interferon-alpha.," *Virology*, vol. 444, no. 1–2, pp. 317–28, Sep. 2013.

[136] E. K. Rydberg, L. Salomonsson, L. M. Hultén, K. Norén, G. Bondjers, O. Wiklund, T. Björnheden, and B. G. Ohlsson, "Hypoxia increases 25-hydroxycholesterol-induced interleukin-8 protein secretion in human macrophages," *Atherosclerosis*, vol. 170, no. 2, pp. 245–252, Oct. 2003.

[137] A. Simonini, M. Moscucci, D. W. Muller, E. R. Bates, F. D. Pagani, M. D. Burdick, and R. M. Strieter, "IL-8 is an angiogenic factor in human coronary atherectomy tissue.," *Circulation*, vol. 101, no. 13, pp. 1519–26, Apr. 2000.

[138] S. L. Taylor, B. R. Renshaw, K. E. Garka, D. E. Smith, and J. E. Sims, "Genomic organization of the interleukin-1 locus.," *Genomics*, vol. 79, no. 5, pp. 726–33, May 2002.

[139] V. Speirs, M. J. Kerin, C. J. Newton, D. S. Walton, A. R. Green, S. B. Desai, and S. L. Atkin, "Evidence for transcriptional activation of ERalpha by IL-1beta in breast cancer cells.," *Int. J. Oncol.*, vol. 15, no. 6, pp. 1251–4, Dec. 1999.

[140] C. H. Hannum, C. J. Wilcox, W. P. Arend, F. G. Joslin, D. J. Dripps, P. L. Heimdal, L. G. Armes, A. Sommer, S. P. Eisenberg, and R. C. Thompson, "Interleukin-1 receptor antagonist activity of a human interleukin-1 inhibitor.," *Nature*, vol. 343, no. 6256, pp. 336–40, Jan. 1990.

[141] S. B. Abramson and A. Amin, "Blocking the effects of IL-1 in rheumatoid arthritis protects bone and cartilage.," *Rheumatology (Oxford).*, vol. 41, no. 9, pp. 972–80, Sep. 2002.

[142] G. Zurawski and J. E. de Vries, "Interleukin 13, an interleukin 4-like cytokine that acts on monocytes and B cells, but not on T cells.," *Immunol. Today*, vol. 15, no. 1, pp. 19–26, Jan. 1994.

[143] A.-L. Andrews, J. W. Holloway, S. M. Puddicombe, S. T. Holgate, and D. E. Davies, “Kinetic analysis of the interleukin-13 receptor complex.,” *J. Biol. Chem.*, vol. 277, no. 48, pp. 46073–8, Nov. 2002.

[144] G. J. McKenzie, A. Bancroft, R. K. Grencis, and A. N. McKenzie, “A distinct role for interleukin-13 in Th2-cell-mediated immune responses.,” *Curr. Biol.*, vol. 8, no. 6, pp. 339–42, Mar. 1998.

[145] B. A. Pippin, M. Rosenstein, W. F. Jacob, Y. Chiang, and M. T. Lotze, “Local IL-4 delivery enhances immune reactivity to murine tumors: gene therapy in combination with IL-2.,” *Cancer Gene Ther.*, vol. 1, no. 1, pp. 35–42, Mar. 1994.

[146] G. Noffz, Z. Qin, M. Kopf, and T. Blankenstein, “Neutrophils but not eosinophils are involved in growth suppression of IL-4-secreting tumors.,” *J. Immunol.*, vol. 160, no. 1, pp. 345–50, Jan. 1998.

[147] W. M. Stadler, M. E. Rybak, and N. J. Vogelzang, “A phase II study of subcutaneous recombinant human interleukin-4 in metastatic renal cell carcinoma.,” *Cancer*, vol. 76, no. 9, pp. 1629–33, Nov. 1995.

[148] C. Stremmel, E. A. Greenfield, E. Howard, G. J. Freeman, and V. K. Kuchroo, “B7-2 expressed on EL4 lymphoma suppresses antitumor immunity by an interleukin 4-dependent mechanism.,” *J. Exp. Med.*, vol. 189, no. 6, pp. 919–30, Mar. 1999.

[149] M. Kobayashi, H. Kobayashi, R. B. Pollard, and F. Suzuki, “A pathogenic role of Th2 cells and their cytokine products on the pulmonary metastasis of murine B16 melanoma.,” *J. Immunol.*, vol. 160, no. 12, pp. 5869–73, Jun. 1998.

[150] S. Gillessen, D. Carvajal, P. Ling, F. J. Podlaski, D. L. Stremlo, P. C. Familletti, U. Gubler, D. H. Presky, A. S. Stern, and M. K. Gately, “Mouse interleukin-12 (IL-12) p40 homodimer: a potent IL-12 antagonist.,” *Eur. J. Immunol.*, vol. 25, no. 1, pp. 200–6, Jan. 1995.

[151] B. S. McKenzie, R. A. Kastelein, and D. J. Cua, “Understanding the IL-23-IL-17 immune pathway.,” *Trends Immunol.*, vol. 27, no. 1, pp. 17–23, Jan. 2006.

[152] O. Shimozato, S. Ugai, M. Chiyo, H. Takenobu, H. Nagakawa, A. Wada, K. Kawamura, H. Yamamoto, and M. Tagawa, “The secreted form of the p40 subunit of interleukin (IL)-12 inhibits IL-23 functions and abrogates IL-23-mediated antitumour effects.,” *Immunology*, vol. 117, no. 1, pp. 22–8, Jan. 2006.

[153] S. J. Ha, C. H. Lee, S. B. Lee, C. M. Kim, K. L. Jang, H. S. Shin, and Y. C. Sung, “A novel function of IL-12p40 as a chemotactic molecule for macrophages.,” *J. Immunol.*, vol. 163, no. 5, pp. 2902–8, Sep. 1999.

[154] S. A. Khader, S. Partida-Sanchez, G. Bell, D. M. Jelley-Gibbs, S. Swain, J. E. Pearl, N. Ghilardi, F. J. Desauvage, F. E. Lund, and A. M. Cooper, “Interleukin 12p40 is required for dendritic cell migration and T cell priming after Mycobacterium tuberculosis infection.,” *J. Exp. Med.*, vol. 203, no. 7, pp. 1805–15, Jul. 2006.

[155] J. R. Piccotti, S. Y. Chan, K. Li, E. J. Eichwald, and D. K. Bishop, “Differential effects of IL-12 receptor blockade with IL-12 p40 homodimer on the induction of CD4+ and CD8+ IFN-gamma-producing cells.,” *J. Immunol.*, vol. 158, no. 2, pp. 643–8, Jan. 1997.

[156] T. C. Jones, “Future uses of granulocyte-macrophage colony-stimulating factor (GM-CSF).,” *Stem Cells*, vol. 12 Suppl 1, pp. 229–39; discussion 239–40, Jan. 1994.

[157] W. P. Peters, “The effect of recombinant human colony-stimulating factors on hematopoietic reconstitution following autologous bone marrow transplantation.,” *Semin. Hematol.*, vol. 26, no. 2 Suppl 2, pp. 18–23, Apr. 1989.

[158] K. H. Grabstein, D. L. Urdal, R. J. Tushinski, D. Y. Mochizuki, V. L. Price, M. A. Cantrell, S. Gillis, and P. J. Conlon, “Induction of macrophage tumoricidal activity by granulocyte-macrophage colony-stimulating factor.,” *Science*, vol. 232, no. 4749, pp. 506–8, Apr. 1986.

[159] L. J. Zhou and T. F. Tedder, “CD14+ blood monocytes can differentiate into functionally mature CD83+ dendritic cells.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 93, no. 6, pp. 2588–92, Mar. 1996.

[160] K. L. Sung, Y. Li, M. Elices, J. Gang, P. Sriramarao, and D. H. Broide, “Granulocyte-macrophage colony-stimulating factor regulates the functional adhesive state of very late antigen-4 expressed by eosinophils.,” *J. Immunol.*, vol. 158, no. 2, pp. 919–27, Jan. 1997.

[161] J. Nemunaitis, S. N. Rabinowe, J. W. Singer, P. J. Bierman, J. M. Vose, A. S. Freedman, N. Onetto, S. Gillis, D. Oette, and M. Gold, “Recombinant granulocyte-macrophage colony-stimulating factor after autologous bone marrow transplantation for lymphoid cancer.,” *N. Engl. J. Med.*, vol. 324, no. 25, pp. 1773–8, Jun. 1991.

[162] J. Nemunaitis, “A comparative review of colony-stimulating factors.,” *Drugs*, vol. 54, no. 5, pp. 709–29, Nov. 1997.

[163] S. Pestka, J. A. Langer, K. C. Zoon, and C. E. Samuel, “Interferons and their actions.,” *Annu. Rev. Biochem.*, vol. 56, pp. 727–77, Jan. 1987.

[164] U. Boehm, T. Klamp, M. Groot, and J. C. Howard, “Cellular responses to interferon-gamma.,” *Annu. Rev. Immunol.*, vol. 15, pp. 749–95, Jan. 1997.

[165] H. Ikeda, L. J. Old, and R. D. Schreiber, “The roles of IFN gamma in protection against tumor development and cancer immunoediting,” *Cytokine Growth Factor Rev.*, vol. 13, no. 2, pp. 95–109, Apr. 2002.

[166] H. F. Rosenberg, S. Phipps, and P. S. Foster, “Eosinophil trafficking in allergy and asthma.,” *J. Allergy Clin. Immunol.*, vol. 119, no. 6, pp. 1303–10; quiz 1311–2, Jun. 2007.

[167] E. J. Clutterbuck, E. M. Hirst, and C. J. Sanderson, “Human interleukin-5 (IL-5) regulates the production of eosinophils in human bone marrow cultures: comparison and interaction with IL-1, IL-3, IL-6, and GMCSF.,” *Blood*, vol. 73, no. 6, pp. 1504–12, May 1989.

[168] Y. Hitoshi, N. Yamaguchi, S. Mita, E. Sonoda, S. Takaki, A. Tominaga, and K. Takatsu, “Distribution of IL-5 Receptor-Positive B Cells: Expression of IL-5 Receptor on Ly-1(CD5)+ B Cells,” *J. Immunol.*, vol. 144, no. 11, pp. 4218–4225, 1990.

[169] W. Shearer, “Biology of common β receptor–signaling cytokines IL-3, IL-5, and GM-CSF,” *J. Allergy Clin. Immunol.*, vol. 112, no. 4, pp. 653–665, Oct. 2003.

[170] E. Clutterbuck, J. G. Shields, J. Gordon, S. H. Smith, A. Boyd, R. E. Callard, H. D. Campbell, I. G. Young, and C. J. Sanderson, “Recombinant human interleukin 5 is an eosinophil differentiation factor but has no activity in standard human B cell growth factor assays.,” *Eur. J. Immunol.*, vol. 17, no. 12, pp. 1743–50, Dec. 1987.

[171] R. J. Bende, G. J. Jochems, T. H. Frame, M. R. Klein, R. V van Eijk, R. A. van Lier, and W. P. Zeijlemaker, “Effects of IL-4, IL-5, and IL-6 on growth and immunoglobulin production of Epstein-Barr virus-infected human B cells.,” *Cell. Immunol.*, vol. 143, no. 2, pp. 310–23, Sep. 1992.

[172] H. Kikutani, T. Taga, S. Akira, H. Kishi, Y. Miki, O. Saiki, Y. Yamamura, and T. Kishimoto, “Effect of B cell differentiation factor (BCDF) on biosynthesis and

secretion of immunoglobulin molecules in human B cell lines," *J. Immunol.*, vol. 143, no. 2, pp. 990–995, 1985.

[173] F. Houssiau, J. Devogelaer, J. Van Damme, C. de Deuxchaisnes, and J. Van Snick, "Interleukin-6 in synovial fluid and serum of patients whith rheumatoid arthritis and other inflammatory arthritides," *Arthritis Rheum.*, vol. 31, no. 6, pp. 784–8, 1988.

[174] T. Tanaka, M. Narazaki, and T. Kishimoto, "Anti-interleukin-6 receptor antibody, tocilizumab, for the treatment of autoimmune diseases," *FEBS Lett.*, vol. 585, no. 23, pp. 3699–708, 2011.

[175] J. Gauldie, C. Richards, D. Harnish, P. Lansdrop, and H. Baumann, "Interferon beta 2/B-cell stimulatory factor type 2 shares identity with monocyte-derived hepatocyte-stimulating factor and regulates the major acute phase protein response in liver cells," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 84, no. 20, pp. 7251–5, 1987.

[176] J. Fukata, T. Usui, Y. Naitoh, Y. Nakai, and H. Imura, "Effects of recombinant human interleukin-1 alpha, -1 beta, 2 and 6 on ACTH synthesis and release in the mouse pituitary tumour cell line AtT-20," *J. Endocrinol.*, vol. 122, no. 1, pp. 33–9, 1989.

[177] B. Spangelo, A. Judd, P. Isakson, and R. MacLeod, "Interleukin-6 stimulates anterior pituitary hormone release in vitro," *Endocrinology*, vol. 125, no. 1, pp. 575–7, 1989.

[178] J. P. DiSanto, "Cytokines: Shared receptors, distinct functions," *Curr. Biol.*, vol. 7, no. 7, pp. R424–R426, Jul. 1997.

[179] R. J. Armitage, B. M. Macduff, J. Eisenman, R. Paxton, and K. H. Grabstein, "IL-15 has stimulatory activity for the induction of B cell proliferation and differentiation.," *J. Immunol.*, vol. 154, no. 2, pp. 483–90, Jan. 1995.

[180] D. V Jovanovic, J. A. Di Battista, J. Martel-Pelletier, F. C. Jolicoeur, Y. He, M. Zhang, F. Mineau, and J. P. Pelletier, "IL-17 stimulates the production and expression of proinflammatory cytokines, IL-beta and TNF-alpha, by human macrophages.," *J. Immunol.*, vol. 160, no. 7, pp. 3513–21, Apr. 1998.

[181] K. Hirota, H. Ahlfors, J. H. Duarte, and B. Stockinger, "Regulation and function of innate and adaptive interleukin-17-producing cells.," *EMBO Rep.*, vol. 13, no. 2, pp. 113–20, Feb. 2012.

[182] C. R. Crowe, K. Chen, D. A. Pociask, J. F. Alcorn, C. Krivich, R. I. Enelow, T. M. Ross, J. L. Witztum, and J. K. Kolls, “Critical role of IL-17RA in immunopathology of influenza infection.,” *J. Immunol.*, vol. 183, no. 8, pp. 5301–10, Oct. 2009.

[183] C. Van Kooten, J. G. Boonstra, M. E. Paape, F. Fossiez, J. Banchereau, S. Lebecque, J. A. Bruijn, J. W. De Fijter, L. A. Van Es, and M. R. Daha, “Interleukin-17 activates human renal epithelial cells in vitro and is expressed during renal allograft rejection.,” *J. Am. Soc. Nephrol.*, vol. 9, no. 8, pp. 1526–34, Aug. 1998.

[184] L. Schandené, C. Alonso-Vega, F. Willems, C. Gérard, A. Delvaux, T. Velu, R. Devos, M. de Boer, and M. Goldman, “B7/CD28-dependent IL-5 production by human resting T cells is inhibited by IL-10.,” *J. Immunol.*, vol. 152, no. 9, pp. 4368–74, May 1994.

[185] D. F. Fiorentino, A. Zlotnik, T. R. Mosmann, M. Howard, and A. O’Garra, “IL-10 inhibits cytokine production by activated macrophages.,” *J. Immunol.*, vol. 147, no. 11, pp. 3815–22, Dec. 1991.

[186] F. Rousset, S. Peyrol, E. Garcia, N. Vezzio, M. Andujar, J. A. Grimaud, and J. Banchereau, “Long-term cultured CD40-activated B lymphocytes differentiate into plasma cells in response to IL-10 but not IL-4.,” *Int. Immunol.*, vol. 7, no. 8, pp. 1243–53, Aug. 1995.

[187] Y. Levy and J. C. Brouet, “Interleukin-10 prevents spontaneous death of germinal center B cells by induction of the bcl-2 protein.,” *J. Clin. Invest.*, vol. 93, no. 1, pp. 424–8, Jan. 1994.

[188] J. E. de Larco and G. J. Todaro, “Growth factors from murine sarcoma virus-transformed cells.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 75, no. 8, pp. 4001–5, Aug. 1978.

[189] R. J. Coffey, M. Romano, W. H. Polk, and P. J. Dempsey, “Roles for transforming growth factor-alpha in gastric physiology and pathophysiology.,” *Yale J. Biol. Med.*, vol. 65, no. 6, pp. 693–704; discussion 621–3.

[190] N. Tørring, P. E. Jørgensen, B. S. Sørensen, and E. Nexø, “Increased expression of heparin binding EGF (HB-EGF), amphiregulin, TGF alpha and epiregulin in androgen-independent prostate cancer cell lines.,” *Anticancer Res.*, vol. 20, no. 1A, pp. 91–5.

[191] M. L. Rand, J. S. Warren, M. K. Mansour, W. Newman, and D. J. Ringler, “Inhibition of T cell recruitment and cutaneous delayed-type hypersensitivity-induced inflammation with antibodies to monocyte chemoattractant protein-1.,” *Am. J. Pathol.*, vol. 148, no. 3, pp. 855–64, Mar. 1996.

[192] J. H. Gong, L. G. Ratkay, J. D. Waterfield, and I. Clark-Lewis, “An antagonist of monocyte chemoattractant protein 1 (MCP-1) inhibits arthritis in the MRL-lpr mouse model.,” *J. Exp. Med.*, vol. 186, no. 1, pp. 131–7, Jul. 1997.

[193] D. A. Zisman, S. L. Kunkel, R. M. Strieter, W. C. Tsai, K. Bucknell, J. Wilkowski, and T. J. Standiford, “MCP-1 protects mice in lethal endotoxemia.,” *J. Clin. Invest.*, vol. 99, no. 12, pp. 2832–6, Jun. 1997.

[194] T. J. Schall, K. Bacon, R. D. Camp, J. W. Kaspari, and D. V Goeddel, “Human macrophage inflammatory protein alpha (MIP-1 alpha) and MIP-1 beta chemokines attract distinct populations of lymphocytes.,” *J. Exp. Med.*, vol. 177, no. 6, pp. 1821–6, Jun. 1993.

[195] R. Alam, P. A. Forsythe, S. Stafford, M. A. Lett-Brown, and J. A. Grant, “Macrophage inflammatory protein-1 alpha activates basophils and mast cells.,” *J. Exp. Med.*, vol. 176, no. 3, pp. 781–6, Sep. 1992.

[196] G. J. Graham, E. G. Wright, R. Hewick, S. D. Wolpe, N. M. Wilkie, D. Donaldson, S. Lorimore, and I. B. Pragnell, “Identification and characterization of an inhibitor of haemopoietic stem cell proliferation.,” *Nature*, vol. 344, no. 6265, pp. 442–4, Mar. 1990.

[197] N. Kiguchi, T. Maeda, Y. Kobayashi, Y. Fukazawa, and S. Kishioka, “Leptin enhances CC-chemokine ligand expression in cultured murine macrophage.,” *Biochem. Biophys. Res. Commun.*, vol. 384, no. 3, pp. 311–5, Jul. 2009.

[198] D. Graf, S. Müller, U. Korthäuer, C. van Kooten, C. Weise, and R. A. Kroczek, “A soluble form of TRAP (CD40 ligand) is rapidly released after T cell activation.,” *Eur. J. Immunol.*, vol. 25, no. 6, pp. 1749–54, Jun. 1995.

[199] L. D. Notarangelo and A. R. Hayward, “X-linked immunodeficiency with hyper-IgM (XHIM).,” *Clin. Exp. Immunol.*, vol. 120, no. 3, pp. 399–405, Jun. 2000.

Appendix A: Definitions and Summaries of the 23 Cytokines Studied

One cytokine subset, interleukins, was originally named for their ability to mediate actions between leukocytes. Interleukins have since been found to be produced by and to act on many different cell types beyond leukocytes. Interleukin 2 (IL-2) is produced by and acts primarily on T lymphocyte cells (T cells) produced in the thymus. After T cells are antigenically stimulated they release IL-2 which aids in generating cytotoxic and regulatory T cells, helping to prevent overstimulation of the immune response [110]. IL-2 is necessary for the activation-induced cell death of T cells via the Fas ligand. The activation of this pathway induces apoptosis, eliminating autoreactive T cells, thus aiding in the prevention of autoimmune diseases [111]. It has been observed that 50% of IL-2 deficient mice die within 4-9 weeks of birth and the remaining mice develop an autoimmune disease resembling human ulcerative colitis, and express antibodies against self [112]. IL-2 is also stored in resting eosinophils at a concentration of 6 ± 2 picograms/ 1 million cells and can be released within minutes after eosinophil stimulation [113].

Tumor necrosis factor alpha (TNFα) was named appropriately after it was observed the protein was able to induce cell death in tumors. Low levels of TNFα are important in normal tissue remodeling but high levels of TNFα can have negative and potentially fatal effects, such as toxic shock and sepsis [113], [114]. TNFα can be produced by many different cells, both immune and non-immune: macrophages, neutrophils, natural killer cells, osteoblasts, neurons and smooth muscle cells, to name a few. The most potent stimulus for TNFα activation is lipopolysaccharide (LPS); ultraviolet light also stimulates TNFα production in epithelial cells. TNFα mRNA can be detected 15-30 minutes after stimulation even though no *de novo* synthesis is needed, suggesting that TNFα is not stored in cells but perhaps the precursors necessary for its production are [5], [115]. Cytokine messages are known for being labile [116] and the half-life for TNFα mRNA is a brief 30 minutes [117] (half-lives of human mRNA less than two hours have been described as “fast” [118]) In addition, fully formed TNFα is also particularly unstable with a reported half-life of 18.2 minutes in human blood [119], [120], for comparison, thyroid hormones have a half-life of about one week in humans and the p35 subunit of IL-12 has a half-life of two to three hours.

Interleukin 18 (IL-18) is a constitutively expressed proinflammatory cytokine that was previously known as IFNγ-inducing factor. IL-18 works synergistically with IL-12 to induce Th1 cytokines like IFNγ, leading to release of tissue damaging molecules like nitric oxide and reactive oxygen species to fight bacterial infections [121]. IL-18 or IL-12 alone will only modestly induce IFNγ, but when combined, their ability to induce IFNγ is significantly enhanced [122]. In the absence of IL-12, however, IL-18 induces a Th2 cytokine response, stimulating cytokines like IL4 and IL-13.

Vascular endothelial growth factor (VEGF) is an important regulator of angiogenesis (the formation of new blood vessels). The ability of VEGF to stimulate angiogenesis is necessary for wound healing, menstruation, normal development and growth in the corpus luteum [123], fetal eye development [124] and bone development [125], for example. VEGF also plays important roles in many diseases because of its ability to stimulate vascular invasion. Tumor cells cannot survive without adequate blood supply making angiogenesis integral to survival past a certain tumor volume threshold. Arthritis [126], diabetic retinopathy [127] and psoriasis [128] also rely on angiogenesis via VEGF for at least part of their pathogenesis. VEGF is stimulated by hypoxia, especially in tumor cells, due to increased mRNA stability under oxygen-poor conditions [129]. Inflammatory cytokines such as IL-6 and TNFα also stimulate the release VEGF [130], linking chronic inflammation with angiogenesis and cancer.

Interleukin 8 (IL-8) is a proinflammatory cytokine first observed in 1987 for its ability to act as a neutrophil chemoattractant [94]. IL-8 is produced by many different cells, including T cells, neutrophils, NK cells, fibroblasts, epithelial cells and tumor cells. IL-8 can be induced by other proinflammatory cytokines (IL-1, TNFα) [131], bacteria [132], [133] and viral infection [134], [135]. Similar to VEGF, IL-8 is also induced by hypoxia [136] and can act as an angiogenic factor [137].

There are 3 classic members of the IL-1 family: IL-1α, IL-1β and IL-1ra. Although IL-1α and IL-1β are structurally similar, bind to the same cell surface receptor and have similar functions, they only show about 25% amino acid homology [138]. IL-1β is a proinflammatory cytokine that is generally not produced in healthy human cells until challenged with bacterial or viral infections at which time it moderates the acute phase response. Even small amounts of IL-1β will cause a drop in blood pressure, raise the internal body temperature and induce production of IL-6, leading to synthesis of acute phase proteins from the liver [8]. IL-1 has been shown to be important in carcinogenesis and IL-1β was found to induce transcriptional activation of breast cancer cells via estrogen receptor alpha [139]. Interleukin 1 receptor antagonist (IL-1ra) passively inhibits the inflammatory cytokine IL-1 by binding to the IL-1 receptor and blocking IL-1 binding [140]. The production of IL-1ra from monocytes and macrophages is stimulated by IgG binding, GM-CSF and IL-1. Human recombinant IL-1ra is produced commercially under the name anakinra and is used to treat rheumatoid arthritis symptoms by decreasing inflammation [141].

Interleukin 4 (IL-4) and interleukin 13 (IL-13) share several features including immunoglobulin (Ig) isotype switching to IgE. In fact, mice deficient in both IL-4 and IL-13 had undetectable levels of circulating IgE [92]. Due to their integral role in IgE production, these cytokines are important in type I hypersensitivity and allergic conditions. Although IL-4 and IL-13 only exhibit about 30% amino acid sequence homology [142], their related functions are partly due to a shared receptor system [143]. IL-13 and IL-4 are secreted by T cells, mast cells and basophils, while IL-13 is

additionally secreted by natural killer (NK) cells, smooth muscle cells and eosinophils. IL-13 appears to be necessary for the expulsion of gastrointestinal parasites while IL-4 is not [144]. The anti-tumor effects of IL-4 are currently under examination as the cytokine seems to show conflicting roles in tumor development. Originally IL-4 seemed to show a clear anti-tumor role [145], [146], but when therapeutic IL-4 was studied in a clinical human cancer trial, the results did not support the proposed role of IL-4 [147] and indicated that IL-4 may instead be exacerbating the problem. It was later found that tumors would only grow in mice with functional IL-4 [148] and additionally, that IL-4 seemed to promote tumor metastasis [149]. Further investigation into the tumor-development role of IL-4 in will undoubtedly continue.

Interleukins 12 and 23 are both heterodimers that have a p40 subunit in common, while IL-12 has a unique p35 subunit and IL-23 has a unique p19 subunit. This shared p40 subunit is produced independently from both the p35 and p19 subunits; it can be secreted and can function independently (as a monomer, p40, or as a homodimer, p80) [150] or in conjunction with its respective partner subunit. IL-12 and IL-23 are both expressed by dendritic cells and macrophages, while IL-12 is also produced by B cells [151]. The p40 subunit is important in the inflammatory pathway and has been observed to act as an antagonist to both IL-12 and IL-23 by competitively binding to the same receptor and blocking IL-12/23 responses [97], [152]. When released as a homodimer this cytokine is a chemoattractant for macrophages [153], an inducer of dendritic cell migration to the lymph node [154], and a suspect in allograft rejection, partly due to its ability to induce IFNγ [155].

The colony stimulating factors, granulocyte-macrophage colony-stimulating factor (GM-CSF) and granulocyte colony-stimulating factor (G-CSF) are both inducers of hematopoietic cell development and migration of stem cells from the bone marrow to circulation. Both cytokines are known to be potent stimulators of neutrophilic activity [156], [157]. Unlike G-CSF, GM-CSF has also been shown to increase phagocytosis by macrophages [158], promote dendritic cell development [159], and promote eosinophil development and adhesion [160] as it is directed toward a population of progenitor cells earlier in development than G-CSF. Both cytokines can be produced by several different cell types including, macrophages, endothelial cells, fibroblasts and stromal cells, while GM-CSF has also been shown to be produced by T cells, NK cells and respiratory epithelial cells, among others [5]. Both G-CSF and GM-CSF have been approved by the FDA to treat neutropenia (low neutrophil counts) in patients, especially after bone marrow transplant or chemotherapy [161], [162].

Cytokines known as interferons (IFNs) protect against viral and bacterial infection and are classified by two types: type I include IFN-alpha (IFNα) and IFN-beta (IFNβ), and both respond to viral infection [163]; type II IFN is known as IFN-gamma (IFNγ) and is activated after NK cells or T lymphocytes are activated with a non-viral inflammatory agent [164]. IFNγ up-regulates the expression of MHC class II on antigen presenting

cells and is a Th1 cytokine that activates NK cells, cytotoxic T cells and macrophages. IFNγ is also important for anti-tumor activity via two routes; non-immunologic actions like inhibiting angiogenesis and proliferation and promoting apoptosis, and immunologic actions like activating macrophages. Furthermore, recent research suggests that IFNγ may affect certain tumor cell genes making the tumor more immunogenic [165].

Interleukin 5 (IL-5) is produced primarily by Th2 cells [166] and is known predominantly for promoting the growth of eosinophils [167]. Due to the role eosinophils play, IL-5 indirectly effects helminth immunity, allergies and asthma. IL-5 was suspected to be important for B cell development as it was observed that murine B cells in the peritoneal cavity constitutively expressed IL-5 receptors [168]. Although many studies have demonstrated the interaction between IL-5 and B cells *in vitro*, the *in vivo* effects seem to be minimal in both humans and mice. IL-5 knockout mice show very little adverse effects in regards to B cells, other than delayed B cell development [5], [169]. The role of IL-5 in human B cell development also seems to be less important than originally suspected as little or no effect of IL-5 on human B cells was seen in several experiments [170], [171].

Interleukin 6 (IL-6) was first observed in 1985 as a factor that helped stimulate and differentiate antigen-activated B cells into plasma cells [172]. IL-6 is produced in response to many different stimuli, including bacterial and viral infection, other proinflammatory cytokines, and UV exposure. After IL-6 was found to be involved in several inflammatory diseases, such as rheumatoid arthritis (RA) [173] and Chrohn's disease, an anti-IL-6-receptor antibody was approved for use in RA patients by the FDA [174]. IL-6 is important in the synthesis and release of acute phase proteins [175] and has been shown to stimulate the release of adrenocorticotrophic hormones [176], prolactin, growth hormone and luteinizing hormone [177]. IL-6 is produced by many cells including macrophages, dendritic cells, mast cells, B cells, T helper cells, fibroblasts and epithelial cells.

Interleukin 15 (IL-15) has similar effects to IL-2 and both cytokines use the same IL-2 receptor subunits for signal transduction, however, while IL-2 aids in T cell development within the thymus, IL-15 promotes development of T and NK cells outside of the thymus [178]. IL-15 is mainly produced by activated macrophages and muscle cells but is also produced to a lesser extent by keratinocytes, dendritic cells and by placental tissues, among others, but is not produced by activated T cells, which are the sole source of IL-2. IL-15 also stimulates the proliferation of activated B cells but has no effect on resting B cells [179].

Interleukin 17A (IL-17A) is a proinflammatory cytokine that recruits neutrophils and monocytes and induces the release of other cytokines such as TNFα, IL-1β and IL-6 from macrophages [180] making it important in defending against infection. IL-17A is also known as IL-17 or CTLA8 in rodents and is produced by Th17 cells [181]. Although IL-

17A is important in defending against infection, it can be harmful in high concentrations and has been implicated in lung damage after flu infection due to extensive neutrophil recruitment [182]. Furthermore, IL-17A is likely involved in transplant rejection as it has been detected in transplant environments during graft rejection where it was previously absent [183].

Since proinflammatory cytokines can cause damage (cell death) at high concentrations, interleukin 10 (IL-10) has the job of keeping those potentially harmful cytokines in check. IL-10 can inhibit the synthesis of IL-2 and IL-5 from T cells [184] and IL-1, IL-6, IL-8, IL-12 and TNFα from monocytes and macrophages [185]. Another important role for IL-10 is its ability to induce differentiation and proliferation of B cells [186]. This cytokine is produced by many cells types but most predominantly by monocytes, macrophages and T cells. B cells also produce IL-10 where it acts in an autocrine manner to prevent apoptosis [187].

Transforming growth factor alpha (TGFα) was first discovered in sarcoma virus-infected fibroblasts in 1978 [188] and was later discovered to be an embryonic growth factor that was abnormally expressed in neoplasms [189]. Aside from its role in normal embryonic and organ development, it is also important for wound repair, bone reabsorption and angiogenesis. Biological sources of TGFα include keratinocytes, airway epithelial cells, pituitary, mucosal tissues and eosinophils, among others. Despite its important role in normal development, TGFα is most well-known for its role in cancer. TGFα is thought to self-promote growth of tumor cells via the endothelial growth factor receptor (EGFR) on the tumor cell surface [190].

As the name suggests, monocyte chemotactic protein 1 (MCP-1), also known as CCL2, is recognized for its ability to attract monocytes, which becomes especially important at inflammation sites. Blocking MCP-1 has been shown to suppress delayed-type hypersensitivity [191], inflammatory arthritis [192] and shock due to endotoxins [193], delineating its role in inflammation but also pointing to its potential for negative effects. MCP-1 is produced by monocytes, fibroblasts, vascular smooth muscle cells, endothelial cells and mast cells.

Similar to MCP-1, the macrophage inflammatory proteins (MIP) are also monocyte chemoattractants. There are two members of the MIP family, MIP-1α, also known as CCL3 and MIP-1β, also known as CCL4, both are produced by macrophages. Both MIPs can attract T lymphocytes while MIP-1α can also attract B cells [194], eosinophils and basophils; MIP-1α not only attracts these cells but can cause degranulation of eosinophils and histamine release from basophils and mast cells [195]. MIP-1α has been shown to inhibit the proliferation of hematopoietic stem cells [196]. The metabolic hormone leptin can enhance MIP-1α and MIP-1β production in murine macrophages *in vitro* [197].

The interaction between the cell surface receptor CD40 and its ligand CD40L (also known as CD154) plays a critical role in the activation of B cells and subsequent release of immunoglobulins. It was originally thought that cell-cell contact via CD40 on B cells with CD40L on activated T cells was necessary for B cell activation through this receptor, however it was discovered that a soluble form of CD40L (sCD40L) can be cleaved from the membrane-bound ligand and produce the necessary activation signal [198]. Platelets can also express both a membrane-bound and a soluble version of CD40L. Deficiency of CD40L prevents B cells from undergoing class switch and results in the immunodeficiency disease hyper-IgM syndrome [199].

Table 10: Summary of the 23 cytokines measured using the Multiplex assay.
Abbreviations: B: B cell; Baso: Basophil; EOS: Eosinophil; En: Endothelial cell; Epi: Epithelial cell; DC: Dendritic cell; FB: Fibroblast; KC: Keratinocyte; T: T cell; Th2: type 2 T helper; Treg: regulatory T cell; M: Macrophage; MC: Mast cell; Mo: Monocyte; Neu: Neutrophil; NK: natural killer cell; VSM: Vascular smooth muscle

Cytokine	**Major Role(s)**	**Produced By**	**Observed Concentration Ranges in Human Sera* (n=20)**
GM-CSF	Growth and differentiation of hematopoietic cells	T, B, FB, En, M	<3.2 – 756 pg/mL (50% detectable)
TGFα	Embryonic development, wound repair, neoplasia	EOS, KC, mucosal cells	<3.2 – 53.9 pg/mL (45% detectable)
G-CSF	Growth and development of Neu	M, En, FB, stromal cells	<3.2 – 22.2 pg/mL (60% detectable)
IFNγ	Primary cytokine responsible for cell-mediated immunity	Activated T, NK	<3.2 – 32.7 pg/mL (80% detectable)
IL-2	Proliferation and regulation of T cells	$CD4^+$ T	<3.2 – 30 pg/mL (25% detectable)
IL-10	Anti-inflammatory; inhibits Mo and M	M, Th2, Treg	<3.2 – 102 pg/mL (55% detectable)
IL-15	T and NK proliferation	M, muscle cells	<3.2 – 10.7 pg/mL (75% detectable)
sCD40L	B cell activation, class switch, Ig release	CD4+ T, platelets	<3.2 – 403 pg/mL (70% detectable)
IL-17A	Proinflammatory; Neu recruitment	Th17	<3.2 – 10 pg/mL (95% detectable)
IL-1ra	Inhibits IL-1 activity	Mo and tissue M	<3.2 – 156 pg/mL (55% detectable)
IL-13	IgE production, expulsion of GI parasites	EOS, Baso; T, NK, MC, smooth muscle cells	<3.2 – 16.8 pg/mL (20% detectable)
IL-1β	Proinflammatory; moderates acute phase response	Activated M	<3.2 – 27.9 pg/mL (30% detectable)
IL-4	IgE production	T, MC, Baso	<3.2 – 362 pg/mL (55% detectable)
IL-5	EOS production	Th2, EOS, MC	<3.2 – 14.9 pg/mL (70% detectable)
IL-6	Proinflammatory; B cell differentiation; release of acute phase proteins	Many, incl. M, DC, MC, B	<3.2 – 39 pg/mL (80% detectable)
IL-8	Proinflammatory; Neu chemoattractant	Many, incl. T cells, neutrophils, NK cells, fibroblasts, epithelial cells, tumor cells	<3.2 – 81.8 pg/mL (60% detectable)

MIP-1α	Chemoattractant for Mo, T, B, EOS, Baso	M	<3.2 – 1510 pg/mL (70% detectable)
MCP-1	Mo chemoattractant	Mo, FB, VSM, En, MC	63.5 – 439 pg/mL
TNFα	Mediates inflammatory reactions	Many; Activated M, Neu, T, B, NK	<3.2 – 6.4 pg/mL (5% detectable)
MIP-1β	Mo and T chemoattractant	M	<3.2 – 118 pg/mL (85% detectable)
IL-12/23(p40)	M chemoattractant; IFNγ synthesis; IL-12 antagonist	DC, M, B	<3.2 – 406 pg/mL (30% detectable)
VEGF	Stimulates angiogenesis	Many; incl. Neu, M, adipose tissue, En, VSM	<3.2 – 118 pg/mL (75% detectable)
IL-18	Proinflammatory; T and NK maturation; IFNγ stimulation	Several; Incl. M, DC, osteoblasts, KC, intestinal epithelium	** 36.1 – 258 pg/mL (n=46)

*All human cytokine concentration data except for IL-18 provided by Millipore as obtained from their Milliplex MAP Human Cytokine/Chemokine 42-plex panel (catalog number MPXHCYTO-60K) [73].

**Human IL-18 concentration data provided by R&D Systems as obtained by their IL-18 Quantikine ELISA

Appendix B: Additional Materials and Methods

Table 11: Summary of protein and antibody sources in the Multiplex cytokine assay kit used.

Cytokine	Protein Source for Standard and Control	Species Detection antibody was raised against
IFNγ, IL-1β, IL-4, IL-5, IL-12/23(p40), IL-13, TNFα, IL-18	Recombinant NHP	Human
IL-2	Recombinant NHP	NHP
VEGF	Human	Mouse
G-CSF, GM-CSF, IL-1ra, IL-6, IL-8, IL-10, IL-15, IL17A, MCP-1, MIP-1α, MIP-1β, sCD40L, TGFα	Human	Human

Table 12: Summary of assay controls used in all NHP cytokine assays performed.

QC 1	Manufacturer provided	Reconstituted with nanopure water prior to each assay, according to manufacturer instructions (0.25mL water)
QC 2	Manufacturer provided	Reconstituted with nanopure water prior to each assay, according to manufacturer instructions (0.25mL water)
QC 3	NHP pool + diluted QC 1	QC 1 was diluted to 25% with assay buffer and mixed in equal volume with the NHP pool*
QC 4	NHP pool + diluted QC 2	QC 2 was diluted to 25% with assay buffer and mixed in equal volume with the NHP pool*

*NHP Pool was made by combining serum from 2 healthy cynomolgus macaques, collected on the same day and kept frozen at -20^{o}C until the day of assay.

Figure 7: Plate layouts used for all validation assays performed.

Validation Plate 1:

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
A	0 pg/mL Back-ground	0 pg/mL Back-ground	0 pg/mL Back-ground	QC1	QC1	QC1	Precision [high]	Precision [high]	Precision [high]	Linearity: (sample 1 + Std 7) at 20%	Linearity: (sample 1 + Std 7) at 20%	Linearity: (sample 1 + Std 7) at 20%
B	Standard 1	Standard 1	Standard 1	QC2	QC2	QC2	Precision [high]	Precision [high]	Precision [high]	Sample 2 neat	Sample 2 neat	Sample 2 neat
C	Standard 2	Standard 2	Standard 2	QC3	QC3	QC3	Precision [high]	Precision [high]	Precision [high]	Recovery: Sample 2 + Std 7	Recovery: Sample 2 + Std 7	Recovery: Sample 2 + Std 7
D	Standard 3	Standard 3	Standard 3	QC4	QC4	QC4	Precision [high]	Precision [high]	Precision [high]	Recovery: Sample 2 + Std 6	Recovery: Sample 2 + Std 6	Recovery: Sample 2 + Std 6
E	Standard 4	Standard 4	Standard 4	Precision [low]	Precision [low]	Precision [low]	Sample 1 neat	Sample 1 neat	Sample 1 neat	Linearity: (sample 2 + Std 7) at 50%	Linearity: (sample 2 + Std 7) at 50%	Linearity: (sample 2 + Std 7) at 50%
F	Standard 5	Standard 5	Standard 5	Precision [low]	Precision [low]	Precision [low]	Recovery: Sample 1 + Std 7	Recovery: Sample 1 + Std 7	Recovery: Sample 1 + Std 7	Linearity: (sample 2 + Std 7) at 20%	Linearity: (sample 2 + Std 7) at 20%	Linearity: (sample 2 + Std 7) at 20%
G	Standard 6	Standard 6	Standard 6	Precision [low]	Precision [low]	Precision [low]	Recovery: Sample 1 + Std 6	Recovery: Sample 1 + Std 6	Recovery: Sample 1 + Std 6	Sensitivity: Std 1 at 50%	Sensitivity: Std 1 at 50%	Sensitivity: Std 1 at 50%
H	Standard 7	Standard 7	Standard 7	Precision [low]	Precision [low]	Precision [low]	Linearity: (sample 1 + Std 7) at 50%	Linearity: (sample 1 + Std 7) at 50%	Linearity: (sample 1 + Std 7) at 50%	Sensitivity: Std 1 at 50%	Sensitivity: Std 1 at 50%	Sensitivity: Std 1 at 50%

Figure 7 Continued: Validation Plate 2

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
A	0 pg/mL Back-ground	0 pg/mL Back-ground	0 pg/mL Back-ground	QC1	QC1	QC1	Recovery: Sample 3 + Std 6	Recovery: Sample 3 + Std 6	Recovery: Sample 3 + Std 6	Linearity: (sample 4 + Std 7) at 50%	Linearity: (sample 4 + Std 7) at 50%	Linearity: (sample 4 + Std 7) at 50%
B	Standard 1	Standard 1	Standard 1	QC2	QC2	QC2	Linearity: sample 3 at 50%	Linearity: sample 3 at 50%	Linearity: sample 3 at 50%	Linearity: (sample 4 + Std 7) at 20%	Linearity: (sample 4 + Std 7) at 20%	Linearity: (sample 4 + Std 7) at 20%
C	Standard 2	Standard 2	Standard 2	QC3	QC3	QC3	Linearity: (sample 3 + Std 7) at 50%	Linearity: (sample 3 + Std 7) at 50%	Linearity: (sample 3 + Std 7) at 50%	Sample5 (pool) neat	Sample5 (pool) neat	Sample5 (pool) neat
D	Standard 3	Standard 3	Standard 3	QC4	QC4	QC4	Linearity: (sample 3 + Std 7) at 20%	Linearity: (sample 3 + Std 7) at 20%	Linearity: (sample 3 + Std 7) at 20%	Recovery: Sample 5 + Std 7	Recovery: Sample 5 + Std 7	Recovery: Sample 5 + Std 7
E	Standard 4	Standard 4	Standard 4	Sensitivity : Std 1 at 50%	Sensitivity : Std 1 at 50%	Sensitivity : Std 1 at 50%	Sample 4 neat	Sample 4 neat	Sample 4 neat	Recovery: Sample 5 + Std 6	Recovery: Sample 5 + Std 6	Recovery: Sample 5 + Std 6
F	Standard 5	Standard 5	Standard 5	Sensitivity : Std 1 at 50%	Sensitivity : Std 1 at 50%	Sensitivity : Std 1 at 50%	Recovery: Sample 4 + Std 7	Recovery: Sample 4 + Std 7	Recovery: Sample 4 + Std 7	Linearity: sample 5 at 50%	Linearity: sample 5 at 50%	Linearity: sample 5 at 50%
G	Standard 6	Standard 6	Standard 6	Sample 3 neat	Sample 3 neat	Sample 3 neat	Recovery: Sample 4 + Std 6	Recovery: Sample 4 + Std 6	Recovery: Sample 4 + Std 6	Linearity: (sample 5 + Std 7) at 50%	Linearity: (sample 5 + Std 7) at 50%	Linearity: (sample 5 + Std 7) at 50%
H	Standard 7	Standard 7	Standard 7	Recovery: Sample 3 + Std 7	Recovery: Sample 3 + Std 7	Recovery: Sample 3 + Std 7	Linearity: sample 4 at 50%	Linearity: sample 4 at 50%	Linearity: sample 4 at 50%	Linearity: (sample 5 + Std 7) at 20%	Linearity: (sample 5 + Std 7) at 20%	Linearity: (sample 5 + Std 7) at 20%

Figure 7 Continued: Validation Plate 3

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
A	0 pg/mL Back-ground	0 pg/mL Back-ground	0 pg/mL Back-ground	QC1	QC1	QC1	Linearity: (sample 6 + Std 7) at 20%	Linearity: (sample 6 + Std 7) at 20%	Linearity: (sample 6 + Std 7) at 20%	Recovery: Sample 8 + Std 4	Recovery: Sample 8 + Std 4	Recovery: Sample 8 + Std 4
B	Standard 1	Standard 1	Standard 1	QC2	QC2	QC2	Sample 7 neat	Sample 7 neat	Sample 7 neat	Recovery: Sample 8 + Std 3	Recovery: Sample 8 + Std 3	Recovery: Sample 8 + Std 3
C	Standard 2	Standard 2	Standard 2	QC3	QC3	QC3	Recovery: Sample 7 + Std 7	Recovery: Sample 7 + Std 7	Recovery: Sample 7 + Std 7	Linearity: (sample 8 + Std 75 at 50%	Linearity: (sample 8 + Std 75 at 50%	Linearity: (sample 8 + Std 75 at 50%
D	Standard 3	Standard 3	Standard 3	QC4	QC4	QC4	Recovery: Sample 7 + Std 6	Recovery: Sample 7 + Std 6	Recovery: Sample 7 + Std 6	Linearity: (sample 8 + Std 5) at 20%	Linearity: (sample 8 + Std 5) at 20%	Linearity: (sample 8 + Std 5) at 20%
E	Standard 4	Standard 4	Standard 4	Sample 6 neat	Sample 6 neat	Sample 6 neat	Linearity: (sample 7 + Std 7) at 50%	Linearity: (sample 7 + Std 7) at 50%	Linearity: (sample 7 + Std 7) at 50%	Sample 9 (also run on old kit)	Sample 9 (also run on old kit)	Sample 9 (also run on old kit)
F	Standard 5	Standard 5	Standard 5	Recovery: Sample 6 + Std 7	Recovery: Sample 6 + Std 7	Recovery: Sample 6 + Std 7	Linearity: (sample 7 + Std 7) at 20%	Linearity: (sample 7 + Std 7) at 20%	Linearity: (sample 7 + Std 7) at 20%	Sample 10 (also run on old kit)	Sample 10 (also run on old kit)	Sample 10 (also run on old kit)
G	Standard 6	Standard 6	Standard 6	Recovery: Sample 6 + Std 6	Recovery: Sample 6 + Std 6	Recovery: Sample 6 + Std 6	Sample 8 neat	Sample 8 neat	Sample 8 neat	Sample 11 (also run on old kit)	Sample 11 (also run on old kit)	Sample 11 (also run on old kit)
H	Standard 7	Standard 7	Standard 7	Linearity: (sample 6 + Std 7) at 50%	Linearity: (sample 6 + Std 7) at 50%	Linearity: (sample 6 + Std 7) at 50%	Recovery: Sample 8 + Std 5	Recovery: Sample 8 + Std 5	Recovery: Sample 8 + Std 5	Sample 12 (also run on old kit)	Sample 12 (also run on old kit)	Sample 12 (also run on old kit)

Figure 7 Continued: Validation Plate 4: (where (1) through (30) represent individual animal samples tested):

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
A	0 pg/mL Back-ground	0 pg/mL Back-ground	0 pg/mL Back-ground	QC1	QC1	QC1	(5)	(5)	(13)	(13)	(21)	(21)
B	Standard 1	Standard 1	Standard 1	QC2	QC2	QC2	(6)	(6)	(14)	(14)	(22)	(22)
C	Standard 2	Standard 2	Standard 2	QC3	QC3	QC3	(7)	(7)	(15)	(15)	(23)	(23)
D	Standard 3	Standard 3	Standard 3	QC4	QC4	QC4	(8)	(8)	(16)	(16)	(24)	(24)
E	Standard 4	Standard 4	Standard 4	(1)	(1)	(29)	(9)	(9)	(17)	(17)	(25)	(25)
F	Standard 5	Standard 5	Standard 5	(2)	(2)	(29)	(10)	(10)	(18)	(18)	(26)	(26)
G	Standard 6	Standard 6	Standard 6	(3)	(3)	(30)	(11)	(11)	(19)	(19)	(27)	(27)
H	Standard 7	Standard 7	Standard 7	(4)	(4)	(30)	(12)	(12)	(20)	(20)	(28)	(28)

Appendix C: Additional Validation Results

Table 13: Summary of quantitative precision data observed. For intra-assay precision n=24 in one assay. For inter-assay precision n=4 (4 assays in which each control was tested in triplicate). In order to meet assay acceptance criteria, intra-assay precision is expected to be ≤ 20% and inter-assay precision is expected to be ≤ 25%.

Cytokine	Intra-assay Precision	Outcome	Inter-assay Precision	Outcome
GM-CSF	2%	Pass	9%	Pass
TGFα	4%	Pass	4%	Pass
G-CSF	6%	Pass	16%	Pass
IFNγ	2%	Pass	7%	Pass
IL-2	1%	Pass	6%	Pass
IL-10	2%	Pass	6%	Pass
IL-15	2%	Pass	9%	Pass
sCD40L	7%	Pass	7%	Pass
IL-17A	1%	Pass	15%	Pass
IL-1ra	3%	Pass	15%	Pass
IL-13	1%	Pass	8%	Pass
IL-1β	2%	Pass	4%	Pass
IL-4	9%	Pass	7%	Pass
IL-5	3%	Pass	5%	Pass
IL-6	3%	Pass	6%	Pass
IL-8	2%	Pass	7%	Pass
MIP-1α	4%	Pass	8%	Pass
MCP-1	3%	Pass	6%	Pass
TNFα	6%	Pass	45%	Fail
MIP-1β	4%	Pass	10%	Pass
IL-12/23(p40)	2%	Pass	11%	Pass
VEGF	4%	Pass	12%	Pass
IL-18	8%	Pass	14%	Pass

Table 14: Results of quantitative recovery studies. Seven individual animal samples and one NHP serum pool were spiked with known amounts of standard at various levels to yield 17 separate spiked samples, each measured in duplicate, over 3 assays. Each high and low accuracy shown here is the average of duplicates from one of the 17 spiked samples. In order to meet assay acceptance criteria, accuracy is expected to be 75% to 125% of expected values.

Analyte	Lowest Recovery	Highest Recovery	Median	Mean	Result	Millipore Result
GM-CSF	74%	585%	294%	307%	Fail	99%
TGFα	8%	204%	81%	82%	Pass	98%
G-CSF	38%	95%	74%	71%	Fail	93%
IFNγ	36%	120%	77%	75%	Pass	92%
IL-2	55%	139%	91%	93%	Pass	90%
IL-10	23%	113%	57%	67%	Fail	101%
IL-15	53%	113%	82%	83%	Pass	93%
sCD40L	43%	108%	87%	86%	Pass	90%
IL-17A	8%	78%	57%	47%	Fail	93%
IL-1ra	27%	130%	50%	58%	Fail	90%
IL-13	32%	131%	78%	79%	Pass	90%
IL-1β	27%	115%	61%	63%	Fail	91%
IL-4	0%	67%	45%	45%	Fail	94%
IL-5	57%	136%	109%	106%	Pass	96%
IL-6	48%	237%	125%	118%	Pass	87%
IL-8	50%	143%	96%	91%	Pass	89%
MIP-1α	14%	152%	63%	65%	Fail	89%
MCP-1	63%	252%	102%	106%	Pass	90%
TNFα	17%	139%	69%	96%	Fail	96%
MIP-1β	2%	101%	53%	44%	Fail	90%
IL-12/23(p40)	36%	125%	84%	82%	Pass	95%
VEGF	55%	110%	88%	86%	Pass	70%
IL-18	2%	31%	23%	20%	Fail	91%

Table 15: Summary of quantitative linearity studies of neat samples and samples spiked with the highest assay standard.

Cytokine	Linearity of neat samples (n=3)	Lin of spiked samples (n=16)
GM-CSF	< range	166%
TGFα	91%	235%
G-CSF	62%	172%
IFNγ	112%	128%
IL-2	119%	116%
IL-10	100%	151%
IL-15	82%	120%
sCD40L	126%	148%
IL-17A	126%	193%
IL-1ra	163%	181%
IL-13	98%	95%
IL-1β	< range	139%
IL-4	248%	111%
IL-5	33%	132%
IL-6	55%	112%
IL-8	125%	110%
MIP-1α	251%	120%
MCP-1	105%	99%
TNFα	71%	162%
MIP-1β	125%	306%
IL-12/23(p40)	107%	151%
VEGF	114%	151%
IL-18	87%	225%

Appendix D: Non-Human Primate Data for Rejected Cytokines

To discourage the interpretation of data from cytokines that have proven inaccurate, data from the rejected cytokines was not included with results presented earlier. It is presented here for informational purposes only.

The range, median and percentage of NHP samples within the detection limits of the assay are shown for each cytokine in Table 16.

Table 16: Summary of cytokine ranges and percent detectable for the rejected cytokines when tested in 36 individual animals (averages were calculated for animals with multiple samples).

Analyte	Range (pg/mL)	Median	% of samples within working range of assay
GM-CSF	<0.64 - 647	15.4	72%
IL-1ra	<0.64 - 36	<0.64	17%
IL-1β	<0.64 - 3.5	<0.64	8%
IL-17A	<0.64 - 42	<0.64	50%
MIP-1β	<0.64 - 20	2.4	81%
TNFα	<0.64 - 193	2.7	75%
IL-18	<0.64 - 335	42.9	89%

Table 17 shows how the NHP data obtained in this study compares to a normal range in human serum. Figure 8 shows the percent of samples tested form humans and NHPs that are within the detectable ranges.

Table 17: Comparison of cytokine concentrations in human and NHP sera for the rejected cytokines. All human data were provided by Millipore and the NHP data were found in this study. The percent of samples within the working range of the assay is shown if any samples fell outside the detectable limits.

Cytokine	Concentration Range in Human Sera (pg/mL)*	Observed Concentration Range in NHP Sera (pg/mL)
GM-CSF	<3.2 – 53.9 (45% detectable)	<0.64-647 (72% detectable)
IL-1ra	<3.2 – 30 (25% detectable)	<0.64-36 (17% detectable)
IL-1β	<3.2 – 32.7 (80% detectable)	<0.64-3.5 (8% detectable)
IL-17A	<3.2 – 39 (80% detectable)	<0.64-42 (50% detectable)
MIP-1β	<3.2 – 1510 (70% detectable)	<0.64-20 (81% detectable)
TNFα	<3.2 – 406 (30% detectable)	<0.64-193 (75% detectable)
IL-18	** 36.1 – 258 (n=46)	<0.64-335 (89% detectable)

*All human cytokine concentration data except IL-18 provided by Millipore as obtained from their Milliplex MAP Human Cytokine/Chemokine 42-plex panel (catalog number MPXHYTO-60K) [73].
**Human IL-18 concentration data provided by R&D Systems as obtained by their IL-18 Quantikine ELISA

Figure 8: The percent of samples tested from humans and NHPs that are within the detectable ranges for the rejected cytokines.

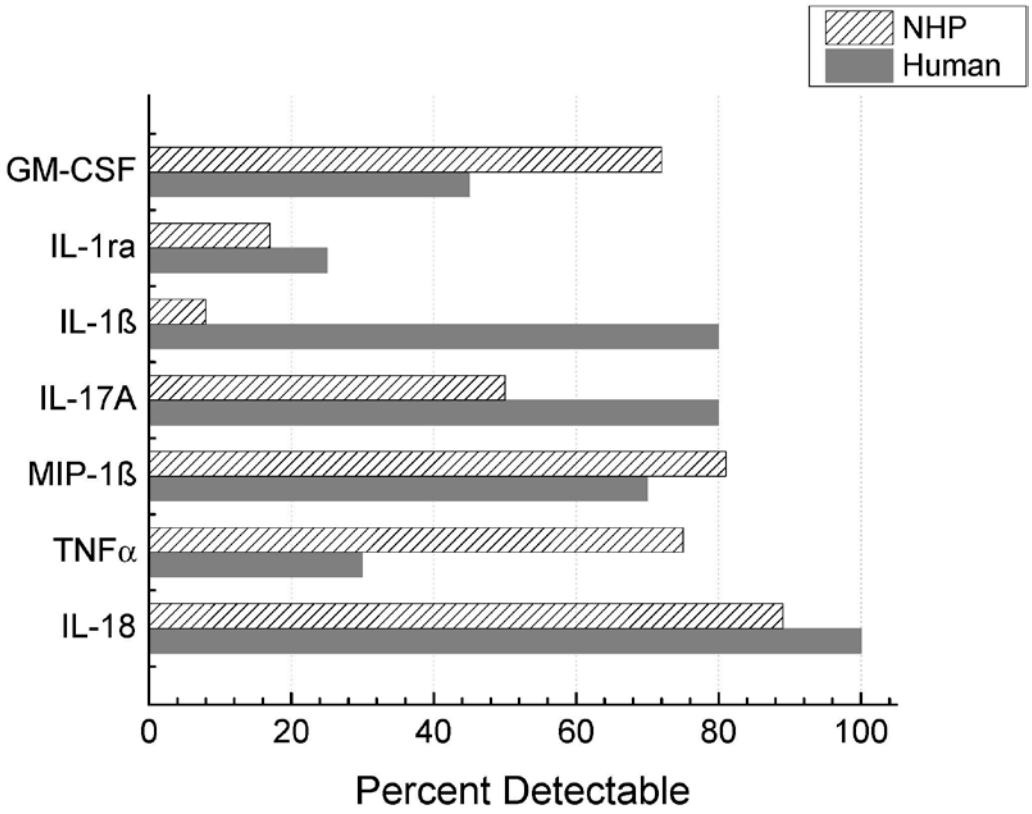
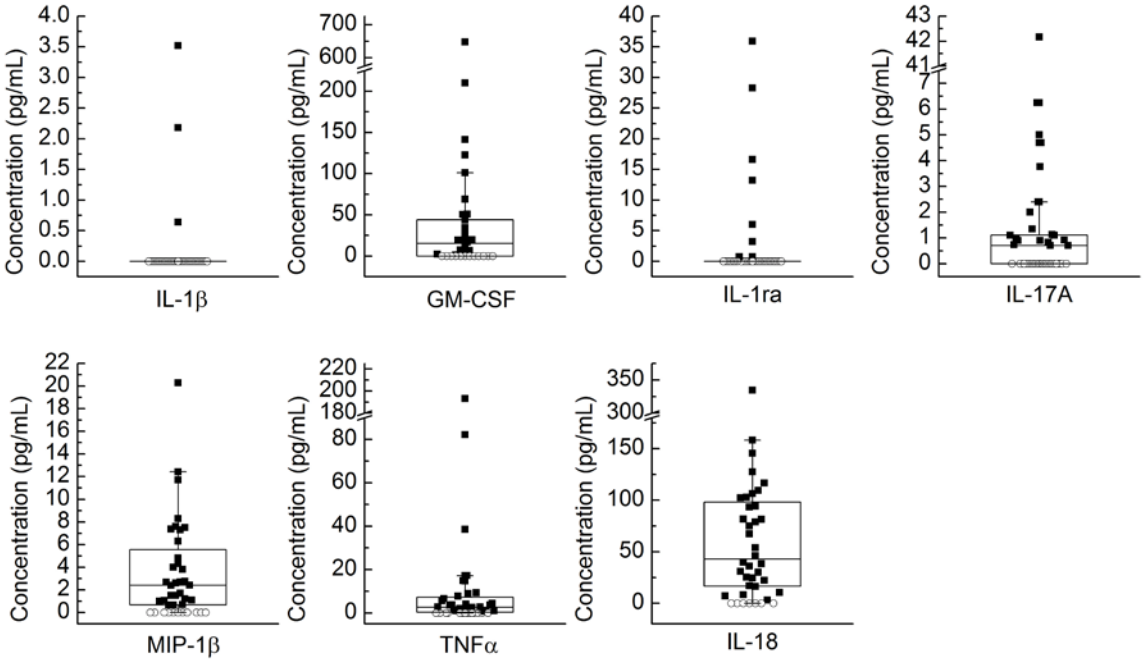


Figure 9 shows the distribution of all 36 animals for each rejected cytokine. Points representing samples outside the limit of detection are differentiated by symbol shape (open circles). For these box-and-whisker plots, the top of the box represents the top 75th percentile, the middle line is the median, the bottom of the box represents the bottom 25th percentile and the whiskers extend to the minimum and maximum (excluding outliers); outliers are shown beyond the reach of the maximum whiskers and were calculated as being $\geq$ 1.5 times the IQR outside the IQR.

Figure 9: Cytokine concentration distribution for rejected cytokines.



Appendix E: Complete NHP Cytokine Results and Fold Change Data

The cytokine concentrations observed for all animals are presented here, in Table 18. For animals that had more than one sample tested an average was taken to determine the cytokine value for that animal. Highlighted cells indicate outliers (as determined using the Tukey method).

Table 18: Cytokine concentration results for each NHP tested (averages calculated for animals with more than 1 sample tested). Outliers are shaded and were determined by the Tukey method (≥ 1.5 times the IQR above the IQR).

Animal ID	Avg of how many samples	Sex	Age (Days)	Weight (Kg)	G-CSF	GM-CSF	IFNγ	IL-1β	IL-1ra	IL-2	IL-4	IL-5	IL-6	IL-8	IL-12/23 (p40)	IL-13	IL-15	IL-17
1	1	M	2194	6.6	<0.64	<0.64	0.7	<0.64	<0.64	13.3	<0.64	<0.64	<0.64	8,468	189.2	1.4	3.8	<0.64
2	2	F	2331	6.4	18.4	209.8	41.1	<0.64	<0.64	8.0	<0.64	0.7	<0.64	622	240.3	1.6	4.5	0.9
3	1	M	2018	7.1	<0.64	<0.64	9.6	<0.64	<0.64	26.3	<0.64	2.5	1.0	2,175	261.0	1.5	10.5	<0.64
4	1	M	1700	4.5	3.1	16.9	1.3	<0.64	<0.64	7.1	<0.64	<0.64	5.6	4,126	72.7	4.0	7.3	<0.64
5	1	M	2296	5.6	6.6	3.3	2.5	<0.64	<0.64	16.6	<0.64	<0.64	21.6	2,753	234.0	12.6	4.5	42.2
6	1	M	1954	6.7	12.8	1.1	1.4	<0.64	3.2	22.2	<0.64	<0.64	<0.64	2,541	56.8	1.9	5.9	<0.64
7	1	M	1968	5.1	3.2	<0.64	<0.64	<0.64	<0.64	22.9	<0.64	<0.64	<0.64	8,215	127.4	1.7	2.7	<0.64
8	2	F	2345	5.3	11.7	647.4	46.4	<0.64	13.2	11.2	<0.64	<0.64	2.7	1,101	473.1	3.1	<0.64	<0.64
9	1	M	2176	5.9	5.7	<0.64	<0.64	<0.64	<0.64	1.3	<0.64	1.9	<0.64	3,432	87.8	2.2	<0.64	<0.64
10	2	M	1788	5.9	8.1	18.7	<0.64	<0.64	<0.64	2.9	2.6	<0.64	<0.64	4,863	56.8	2.9	1.4	1.3
11	3	M	1697	5.7	3.8	26.2	1.1	<0.64	<0.64	8.8	1.9	<0.64	<0.64	5,334	53.4	3.9	3.5	0.8
12	7	M	2133	5.9	6.0	34.3	7.1	<0.64	28.3	6.5	<0.64	<0.64	6.6	2,075	124.9	1.7	5.7	2.0
13	2	M	1655	4.7	8.8	29.5	3.5	<0.64	<0.64	6.6	<0.64	<0.64	<0.64	1,746	106.9	1.8	3.2	0.7
14	1	M	2119	5.8	<0.64	<0.64	1.4	<0.64	<0.64	5.9	<0.64	<0.64	<0.64	647	107.8	<0.64	6.1	<0.64
15	1	M	2019	3.7	LBC	11.2	1.0	<0.64	<0.64	5.8	<0.64	1.2	11.9	5,063	69.6	2.7	2.6	1.1
16	1	M	1752	5.3	<0.64	23.6	10.8	<0.64	<0.64	3.7	<0.64	<0.64	6.5	701	104.8	<0.64	4.3	5.0
17	1	M	1659	5.6	2.9	122.3	12.3	<0.64	0.7	7.1	<0.64	<0.64	1.2	1,991	151.7	6.4	7.9	<0.64
18	2	M	1373	3.9	5.0	44.7	3.3	<0.64	<0.64	6.8	<0.64	2.4	<0.64	6,191	80.3	2.9	5.6	1.1
19	1	M	1876	3.6	<0.64	0.7	1.0	<0.64	<0.64	4.8	<0.64	0.7	<0.64	606	45.3	1.5	2.2	<0.64
20	1	M	1954	7.3	<0.64	<0.64	<0.64	<0.64	<0.64	9.5	<0.64	<0.64	<0.64	1,263	27.9	<0.64	1.1	<0.64
21	1	M	2041	4.5	4.1	<0.64	4.7	<0.64	<0.64	13.7	<0.64	<0.64	<0.64	442	83.9	<0.64	6.2	<0.64
22	1	M	2193	7.9	LBC	<0.64	<0.64	<0.64	<0.64	27.0	<0.64	<0.64	<0.64	3,840	73.6	2.6	3.1	<0.64
23	1	M	2214	6.0	1.5	2.1	47.9	3.5	16.6	32.8	<0.64	<0.64	4.2	442	584.4	1.9	18.4	<0.64
24	1	M	1976	6.4	<0.64	<0.64	0.8	<0.64	<0.64	3.0	<0.64	4.7	<0.64	271	115.3	<0.64	0.9	<0.64
25	3	M	1782	4.4	5.3	4.7	1.4	<0.64	<0.64	10.1	<0.64	1.4	7.6	3,300	122.2	4.7	2.8	0.9

Table 18 Continued

Animal ID	Avg of how many samples	Sex	Age (Days)	Weight (Kg)	MCP-1	MIP-1β	MIP-1α	sCD40L	TGFα	TNFα	VEGF	IL-18
1	1	M	2194	6.6	211.8	<0.64	<0.64	>10000	3.1	4.1	4.7	16.2
2	2	F	2331	6.4	359.2	7.3	8.1	5,125	12.8	2.0	42.8	24.4
3	1	M	2018	7.1	462.2	0.7	<0.64	>10000	5.6	8.9	<0.64	16.9
4	1	M	1700	4.5	439.9	0.7	10.0	7,631	14.7	<0.64	11.9	78.7
5	1	M	2296	5.6	321.2	0.6	256.3	188	8.6	0.8	1798.1	106.3
6	1	M	1954	6.7	447.1	4.3	5.4	9,633	4.3	2.1	4.3	25.2
7	1	M	1968	5.1	294.5	<0.64	<0.64	>10000	5.0	2.6	4.7	53.8
8	2	F	2345	5.3	285.9	11.7	7.1	3,923	15.0	82.2	318.7	74.9
9	1	M	2176	5.9	275.4	1.1	3.9	513	2.6	7.7	<0.64	30.0
10	2	M	1788	5.9	246.1	1.1	1.6	5,279	6.3	<0.64	73.9	102.9
11	3	M	1697	5.7	488.7	2.7	4.5	5,758	13.2	1.1	10.9	145.4
12	7	M	2133	5.9	631.5	4.0	9.6	3,512	13.7	9.3	7.9	46.0
13	2	M	1655	4.7	419.1	2.6	13.9	5,185	8.3	1.4	9.5	30.7
14	1	M	2119	5.8	398.5	1.0	<0.64	3,995	4.0	<0.64	<0.64	<0.64
15	1	M	2019	3.7	356.2	2.8	7.7	10,978	3.5	3.7	228.5	67.3
16	1	M	1752	5.3	350.2	<0.64	54.2	3,138	22.8	2.8	9.8	8.0
17	1	M	1659	5.6	281.2	3.8	7.5	4,750	7.2	14.8	35.3	109.3
18	2	M	1373	3.9	575.9	1.7	14.6	6,357	16.7	3.6	5.4	94.1
19	1	M	1876	3.6	170.1	1.5	4.3	475	6.6	<0.64	<0.64	3.1
20	1	M	1954	7.3	200.0	<0.64	4.6	1,693	14.5	<0.64	<0.64	<0.64
21	1	M	2041	4.5	102.4	<0.64	5.9	5,169	4.5	0.7	7.3	<0.64
22	1	M	2193	7.9	336.3	<0.64	2.6	>10000	3.8	<0.64	12.2	36.0
23	1	M	2214	6.0	293.1	12.4	17.8	6,762	10.4	38.6	201.8	81.4
24	1	M	1976	6.4	338.2	<0.64	0.8	10,829	6.1	14.9	<0.64	<0.64
25	3	M	1782	4.4	462.3	7.6	12.2	7,600	7.8	1.2	28.6	158.3

Table 18 Continued

Animal ID	Avg of how many samples	Sex	Age (Days)	Weight (Kg)	G-CSF	GM-CSF	IFNγ	IL-1β	IL-1ra	IL-2	IL-4	IL-5	IL-6	IL-8	IL-12/23 (p40)	IL-13	IL-15	IL-17
26	2	M	1794	5.1	0.9	19.7	6.4	<0.64	<0.64	7.6	<0.64	4.7	1.5	1,599	74.9	2.9	1.5	<0.64
27	1	M	1657	4.7	3.6	7.1	3.3	<0.64	<0.64	4.8	LBC	<0.64	3.0	2,408	24.4	2.3	2.9	1.0
28	1	M	1768	5.1	3.0	43.5	<0.64	<0.64	<0.64	77.8	<0.64	<0.64	<0.64	8,973	136.6	<0.64	1.7	0.9
29	2	M	1918	5.1	6.8	50.5	2.7	<0.64	<0.64	10.0	<0.64	<0.64	<0.64	1,593	117.9	2.7	1.6	0.7
30	1	M	1865	4.0	3.9	101.1	16.4	2.2	<0.64	11.8	<0.64	15.4	5.5	943	146.0	3.0	13.8	<0.64
31	1	M	1812	4.2	<0.64	13.8	1.1	<0.64	<0.64	6.6	<0.64	<0.64	1.2	1,281	74.0	1.7	5.4	<0.64
32	1	M	1248	4.0	8.7	141.4	4.1	<0.64	<0.64	12.5	2.8	<0.64	<0.64	13,345	46.1	11.1	1.0	2.4
33	1	M	1881	3.9	<0.64	19.2	3.8	<0.64	<0.64	1.8	<0.64	<0.64	1.2	556	60.0	1.4	<0.64	0.7
34	2	M	2179	4.2	<0.64	50.4	<0.64	<0.64	35.9	9.1	<0.64	<0.64	7.4	1,460	76.1	6.4	3.3	6.3
35	1	M	1854	5.1	6.8	68.9	97.6	<0.64	6.0	17.6	<0.64	<0.64	<0.64	454	222.3	5.1	13.2	1.1
36	1	M	1922	4.8	13.2	6.6	1.2	<0.64	<0.64	6.1	<0.64	<0.64	<0.64	2,150	57.8	6.3	5.7	4.7

Animal ID	Avg of how many samples	Sex	Age (Days)	Weight (Kg)	MCP-1	MIP-1β	MIP-1α	sCD40L	TGFα	TNFα	VEGF	IL-18
26	2	M	1794	5.1	335.4	2.4	20.7	4,448	3.3	6.7	17.8	93.4
27	1	M	1657	4.7	304.8	2.4	11.3	4,054	12.0	<0.64	5.2	38.2
28	1	M	1768	5.1	602.2	20.3	18.2	4,038	8.5	5.5	32.4	102.3
29	2	M	1918	5.1	505.6	2.7	5.1	3,123	7.0	3.6	11.2	81.6
30	1	M	1865	4.0	429.8	6.3	10.6	2,606	22.3	17.2	4.7	10.4
31	1	M	1812	4.2	542.4	1.2	14.8	4,675	11.3	<0.64	5.2	22.0
32	1	M	1248	4.0	474.2	7.5	11.4	4,787	20.2	4.5	53.6	334.7
33	1	M	1881	3.9	238.7	1.5	<0.64	1,673	2.4	2.8	<0.64	7.1
34	2	M	2179	4.2	380.3	7.4	37.3	3,502	22.0	1.1	344.7	39.7
35	1	M	1854	5.1	142.7	4.8	32.7	8,861	36.1	193.2	38.6	127.5
36	1	M	1922	4.8	242.8	8.3	15.7	8,120	47.4	<0.64	16.3	116.6

In this cohort of 36 animals, 11 animals had multiple samples tested and of those 11, eight animals had only two samples tested. In order to provide an analysis of intra-animal variability, the fold change of each cytokine was calculated in each animal with multiple observations. The fold change is calculated by dividing the highest concentration by the lowest concentration. A result of 1.0 indicates no change. If more than two observations were included in the fold-change-analysis, the highest and lowest observations were used in calculations and if censored data were included, it was changed to half the detection limit (0.32 pg/mL rather than <0.62 pg/mL). To evaluate the overall changes rather than the direction of change (increased or decreased concentration) from initial to final sample, the highest concentration was always divided by the lowest concentration so that all results are reported as fold changes of ≥ 1.0. These data are presented in Table 19 for the accepted and marginal cytokines, while Table 20 shows these data for the rejected cytokines. Animal E is represented twice in Tables 19 and 20, as E1 and E2. This animal had seven total samples tested (represented as E1), three of which were collected on the same day then tested in a single assay (represented as E2). The average variability for the 15 accepted and marginal cytokines within a day for this animal was a 2.4 fold change, compared to the average fold change for all animals of 2.6 in these same 15 cytokines. This supports the idea that the observed intra-animal variability may not be due to changes in health status but perhaps are part of normal, daily secretion patterns. Additional studies are needed to define this variability.

Table 19: Fold change data for each of the 11 animals with more than 1 sample tested over time in the accepted and marginal cytokines.

Animal	A (n=2)	B (n=2)	C (n=3)	D (n=3)	E1 (n=7)	E2 (n=3)	F (n=2)	G (n=2)	H (n=3)	I (n=2)	J (n=2)	K (n=2)	Average
G-CSF	1.3	1.0	1.3	1.1	5.0	1.7	2.2	1.2	2.5	1.8	1.7	1.0	1.8
IFNγ	2.2	1.3	1.0	1.3	3.6	1.4	1.9	1.3	1.7	1.5	1.1	3.5	1.8
IL-4	3.1	1.0	16.0	1.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	3.3	2.6
IL-2	3.4	1.2	1.1	2.3	2.6	2.1	1.2	2.0	1.2	1.5	1.1	2.1	1.8
IL-5	4.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	2.1	29.1	1.0	1.0	3.7
IL-6	1.0	2.3	1.0	1.0	1.0	1.0	1.0	1.0	5.9	4.0	1.0	46.2	5.5
IL-8	2.1	1.1	1.3	1.6	7.7	3.0	1.6	1.0	1.8	1.6	1.2	1.1	2.1
IL-12/23 (p40)	2.0	1.9	1.2	1.7	2.9	1.2	1.8	1.3	1.4	1.2	1.1	1.1	1.6
IL-13	1.2	1.0	1.8	2.5	2.2	1.3	1.1	1.3	1.5	1.5	1.0	20.0	3.0
IL-15	1.4	1.0	2.5	1.6	3.0	2.1	3.3	2.9	2.3	2.4	1.4	10.4	2.9
MCP-1	1.6	1.2	1.2	2.0	2.9	2.9	1.2	2.0	1.7	2.0	1.1	1.2	1.8
MIP-1α	1.3	1.4	9.7	2.0	3.7	1.2	1.2	1.8	1.7	1.6	1.4	7.0	2.7
sCD40L	1.1	1.0	1.1	1.3	10.5	10.5	1.5	1.3	1.4	2.2	2.5	2.1	3.0
TGFα	1.1	1.8	1.4	2.3	4.9	2.1	2.1	1.1	6.3	1.6	2.4	4.7	2.2
VEGF	1.7	1.3	2.1	2.2	5.4	3.5	1.1	1.5	2.9	1.2	1.0	8.6	2.5
Average	1.9	1.3	2.9	1.7	3.8	2.4	1.6	1.4	2.4	3.6	1.3	7.6	2.6

Table 20: Fold change data for each of the 11 animals with more than 1 sample tested over time in the rejected cytokines.

Animal	A (n=2)	B (n=2)	C (n=3)	D (n=3)	E1 (n=7)	E2 (n=3)	F (n=2)	G (n=2)	H (n=3)	I (n=2)	J (n=2)	K (n=2)	Average
GM-CSF	4.7	1.1	3.5	8.2	14.6	1.5	2.0	4.8	1.3	2.0	1.1	56.3	8.4
IL-1ra	3.8	1.4	1.0	1.0	30.6	8.1	1.0	1.0	4.5	1.0	1.0	224.4	23.2
IL-1β	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
IL-17A	1.3	1.6	1.8	1.7	4.7	1.4	1.5	2.7	1.1	1.0	1.5	11.0	2.6
MIP-1β	1.2	1.4	6.7	2.5	5.6	1.3	1.4	2.7	2.8	2.2	2.5	1.9	2.7
TNFα	1.3	1.8	1.0	1.3	2.9	1.2	8.8	1.0	3.1	9.2	1.1	7.0	3.3
IL-18	2.6	1.1	2.2	1.6	58.5	3.1	1.4	1.2	1.0	1.2	1.1	6.1	6.7
Average	2.2	1.3	2.5	2.5	16.8	2.5	2.4	2.0	2.1	2.5	1.3	44.0	6.9