



Published in final edited form as:

Cytokine. 2011 November ; 56(2): 430–434. doi:10.1016/j.cyto.2011.06.011.

Interlaboratory Reproducibility of Female Genital Tract Cytokine Measurements by Luminex: Implications for Microbicide Safety Studies

Mark E. Scott^{a,*}, Sarah S. Wilson^{b,1}, Lisa A. Cosentino^c, Barbra A. Richardson^d, Anna-Barbara Moscicki^a, Sharon L. Hillier^{c,e}, and Betsy C. Herold^b

^aDivision of Adolescent Medicine, Department of Pediatrics, School of Medicine, University of California, San Francisco, 513 Parnassus Avenue, San Francisco, CA 94143, USA. scottm@peds.ucsf.edu (MES) and moscickia@peds.ucsf.edu (ABM)

^bDepartments of Pediatrics and Microbiology-Immunology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA. swilson@einstein.yu.edu (SW) and betsy.herold@einstein.yu.edu (BCH)

^cDepartment of Obstetrics, Gynecology, and Infectious Disease Research, Magee-Womens Research Institute, 204 Craft Avenue, Pittsburgh, PA 15213, USA. lcosentino@mwri.magee.edu (LAC) and hillsll@mwri.magee.edu (SLL)

^dDepartment of Biostatistics, University of Washington, Box 359909, 325 Ninth Avenue, Seattle, WA 98104, USA. barbrar@u.washington.edu

^eDepartment of Obstetrics, Gynecology, and Reproductive Services, University of Pittsburgh, 300 Halket Street, Pittsburgh, PA 15213, USA

Abstract

The interlaboratory reproducibility of cytokine measurements from cervicovaginal samples by Luminex has not been reported. Using cervicovaginal lavage specimens collected on three study days from 12 women participating in a Phase I microbicide study, we measured a panel of eight cytokines in three independent laboratories. Four (IFN- γ , IL-10, IL-17 and TNF) were below the limit of detection in the majority (85%) of samples in either two or all three laboratories, an observation that may guide analyte selection for future studies. Good interlaboratory agreement (intraclass correlation coefficient, $r > 0.7$) in absolute levels was observed for IL-1 β , IL-6, and IL-8, while poor agreement was seen for IFN- $\alpha 2$ ($r = 0.47$). When considering within-subject change from baseline (pre-product, at study-day 0) to either post-product visit (study-days 7 and 14), IL-1 β and IL-6 exhibited good interlaboratory agreement ($r > 0.7$), while IFN- $\alpha 2$ and IL-8 did not. Future studies addressing the clinical utility of specific biomarkers of inflammation for microbicide trials should consider reproducibility in the context of defining biologically meaningful thresholds of change for candidate biomarkers, ensuring that such change can be reliably distinguished from background variability.

© 2011 Elsevier Ltd. All rights reserved.

*Corresponding author. Address: Division of Adolescent Medicine, Department of Pediatrics, Box 1374, University of California, San Francisco, 513 Parnassus Avenue, San Francisco, CA 94143-1374 USA. Phone: +1 415-476-3260 Fax: +1 415-502-1222. scottm@peds.ucsf.edu.

¹Present address: Department of Microbiology, University of Washington, Seattle, WA 98195, USA. sarahsw@uw.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

cytokine measurement; reproducibility; multiplex methods; cervicovaginal secretions; female reproductive tract; microbicide studies

1. Introduction

Vaginal microbicides hold great promise as a female-controlled strategy for the prevention of human immunodeficiency virus (HIV) and other sexually transmitted infections. The recent CAPRISA 004 trial, in which a 39% reduction in HIV and an unanticipated 51% reduction in HSV-2 acquisition were observed in women who applied 1% tenofovir gel before and after sex, illustrates the exciting potential of this strategy [1]. These encouraging results contrast with those obtained in earlier microbicide trials with surfactants and polyanionic entry inhibitors. Not only did the earlier products fail to protect against HIV, but several (Nonoxynol-9, C31G [Savvy], and cellulose sulfate) were associated with at least a trend towards higher rates of HIV infection [2–4]. Subsequent work indicated that those products may have facilitated HIV acquisition by inducing local inflammation and disrupting the epithelial barrier [5–9].

The central role inflammation plays in promoting HIV infection is further supported by studies with other sexually transmitted infections, suggesting that the inflammatory response recruits T cells into the genital tract to increase the risk of HIV infection. Inflammatory cytokines may also promote HIV replication by activating the long terminal repeat. Together these observations suggest that inflammatory mediators in genital tract secretions may serve as biomarkers of HIV risk and their measurement could prove predictive of the safety of vaginal microbicides, mucosal vaccines, or other interventions. A critical prerequisite in the development of biomarkers is the validation and standardization of assays. While Luminex multiplex technology is commonly used to measure cytokines and chemokines in various specimen types including female genital tract specimens, there are little or no data about the reproducibility of results across different laboratories. This has important implications for future clinical studies, as it will determine whether assays need to be performed at a single centralized laboratory and the extent to which results obtained from different studies can be compared.

Therefore to address this gap, convenience cervical samples that had been collected as part of a Phase I microbicide safety trial were evaluated in a blinded fashion in three independent laboratories to determine the interlaboratory variability in cytokine and chemokine measurements using the Luminex-100 multiplex system. We selected a panel of mediators that included those that had been previously shown to be associated with HIV infection risk in vitro (IL-1 β , IL-6, IL-8, and TNF), antiviral cytokines (IFN- α 2 and IFN- γ), the anti-inflammatory cytokine IL-10, and IL-17, which plays a role in neutrophil recruitment, promotes the production of β -defensins [10], and has been shown to play a role in the immune response to *N. gonorrhoeae* [11].

2. Materials and Methods

2.1 Patient population and specimen collection

Cervical vaginal lavage (CVL) specimens collected from women participating in a Phase I clinical safety trial of a candidate vaginal microbicide were used for this study [12]. The trial is registered at www.ClinicalTrials.gov (NCT00331032) and all participants provided informed consent. De-identified specimens from days 0, 7, and 14 were evaluated, where day 0 represents baseline (pre-product) and days 7 and 14 represent the respective number of

days on either product or placebo (in the same gel carrier). At each study visit, the cervix was lavaged with 5 ml normal saline, following which the specimen was aspirated and frozen at -80°C without centrifugation.

2.2 Sample handling and testing

Thirty-six CVL specimens from 12 subjects were tested in a blinded (for subject and study day) fashion by three independent research laboratories for eight cytokines (IFN- α 2, IFN- γ , IL-1 β , IL-6, IL-8, IL-10, IL-17, and TNF). As it has previously been reported that assays from different vendors have poor intervendor agreement [13–15], a single vendor, Millipore (Billerica, MA), was used for this study. Specimens were thawed, vortexed to ensure homogeneity, aliquoted, refrozen, and distributed to each laboratory. Following identical sample handling and testing protocols at each of the three laboratories, the specimens were thawed at room temperature, centrifuged (2000 G at 4°C for 10 min) to remove mucus and cellular debris, and tested in duplicate using MilliPlex MAP human cytokine/chemokine immunoassay kits (Millipore), per manufacturer's instructions. Kits from a single manufacturing lot were used by all three laboratories. Briefly, samples were incubated with antibody-conjugated microspheres, overnight, in 96-well filter-membrane assay plates with agitation. Plates were then washed with wash buffer provided in the assay kits and vacuum filtration, following which analyte-bound beads were incubated with a biotinylated detection antibody cocktail and finally with streptavidin-phycoerythrin. Following additional wash steps and resuspension of beads in instrument sheath fluid, plates were run on Luminex 100 instruments (Luminex, Austin, TX). Regression curves (5-parameter logistic) were fit, and unknown concentrations in pg/ml determined by interpolation, by each laboratory using their local software (Laboratories A and B: STarStation [Applied Cytometry Systems, Sacramento, CA]; Laboratory C: MiraiBio MasterPlex QT version 2.5 [Hitachi Software, South San Francisco, CA]). Concentrations from duplicate wells were averaged.

2.3 Statistical analysis

Sample values below the lowest standard (3.2 pg/ml) were set at the midpoint between zero and this value; values above the highest standard (10,000 pg/ml) were set at 10,000 pg/ml. Agreement of the cytokine measurements among the three laboratories was assessed using the intraclass correlation coefficient (r), which provides an index of the intersubject variability relative to the total variability [16], for all measurements and stratified by study day. Within-subject changes from baseline (day 0) at day 7 (“ $\Delta 7$ ”) and day 14 (“ $\Delta 14$ ”) were calculated by subtracting log-transformed baseline levels for each subject from log-transformed day 7 and (separately) day 14 levels (equivalent to ratios of the non-log-transformed values). Interlaboratory agreement in detecting within-subject change at the two post-product visits was then examined using the intraclass correlation coefficient. SPSS version 19 (SPSS, Inc.) was used for all statistical analyses.

3. Results

3.1 Interlaboratory agreement of absolute levels

The median and range of baseline (day 0) values obtained for each cytokine and from each lab are shown in Table 1. Four of the cytokines tested (IFN- γ , IL-10, IL-17 and TNF) exhibited expression levels too low in CVLs to measure reliably, with 85% or more of samples (from all study days) falling below detection limits in either two or all three laboratories (Table 2). These were therefore excluded from further analyses. Of the other four cytokines, good interlaboratory agreement ($r > 0.7$) was seen for IL-1 β and IL-6, both overall and stratified by study day, and for IL-8 overall and on study-days 7 and 14 (Table 3). IFN- α 2, in contrast, showed poor interlaboratory agreement except for the day-7 samples.

3.2 Interlaboratory agreement of within-subject cytokine-level change

Because of the variable degree of interlaboratory agreement in absolute cytokine measurements, and because normal ranges for these markers in CVL specimens have not been established, it was important to assess whether the ability to detect within-subject change from baseline was consistent across laboratories. Interlaboratory agreement in within-subject change was good at both post-product study days for IL-1 β and IL-6 (Table 4). IFN- α 2 and IL-8, in contrast, showed poor interlaboratory agreement.

In the case of IL-8 it was noted that, despite the overall interlaboratory agreement in absolute levels shown in Table 3, three of 36 specimens showed marked interlaboratory discordance (not shown) and that these were all clustered in the baseline specimen group. Because the discordance among baseline samples could affect both the day-7 and day-14 within-subject change assessments, which might not have been the case had these been distributed across study days, we repeated the analysis with these three subjects omitted. In this analysis, the intraclass correlation coefficient rose to 0.46 (95% CI: 0.05, 0.82) at day 7 and 0.81 (95% CI: 0.53, 0.95) at day 14.

4. Discussion

Using the intraclass correlation coefficient, which, in repeatability studies, provides an index of the natural variability between samples relative to the total variability [16], we found that three (IL-1 β , IL-6, and IL-8) of the four cytokines with measurable levels had good interlaboratory agreement. This was the case both in absolute level measurements and (for IL-1 β and IL-6) when examining within-subject change when baseline and post-product samples are tested within the same laboratory. IFN- α 2, in contrast, showed poor reproducibility. We note, also, that several important cytokines of biologic interest (IFN- γ , IL-10, IL-17, and TNF) demonstrated levels too low to be reliably measured using the Luminex-based assay kits, an observation which may guide analyte selection for future studies of genital-tract immune markers.

While our results suggest reasonably good interlaboratory agreement, there was variability in absolute measurements among the laboratories, which was more marked for some samples than others. This impacted the IFN- α 2 reproducibility most obviously, but was observed to a lesser extent for the other cytokines tested as well. Among known causes of variability of Luminex measurements [17–19], it has been demonstrated that different instruments can give significantly different readings even when calibrated to the same standard, presumably due to differences in their opto-electrical response curves [20]. Also of note in our study is that the instruments used by the participating laboratories were outfitted with different software packages for acquisition and analysis. This raises the possibility of variability being introduced through different underlying curve-fitting algorithms in the respective software packages, even with all three laboratories using a 5-parameter logistic curve fit. While the 5-parameter model most closely fits the ligand-binding kinetics of immunoassays, nearly eliminating the lack-of-fit error of the 4-parameter logistic model while avoiding the pitfalls of overparameterized models [21], it can be much more difficult to fit via software algorithms. In determining best fit by minimizing the weighted sum of squared errors, most algorithms are unable to reliably distinguish a local minimum in an ill-conditioned regression from the global minimum of the correct result [21]. Unfortunately, employment of a proprietary data-file format by STarStation [22] precludes our exchanging raw fluorescence data for reanalysis on the opposite platform to further examine the role of software in our findings.

Low levels may also have contributed to the poor interlaboratory agreement observed for IFN- α 2. Wong et al. reported coefficients of variation in cytokine measurements between

replicate serum samples as high as 44%, which was in sharp contrast to previous studies that had assessed reliability of Luminex measurements in the linear portion of the standard curves [23]. Because these authors studied physiologic levels, many cytokines fell into the lowest portion of the sigmoidal standard curve, where a leveling off of the curve increases the imprecision in unknown interpolation. In spite of that, they concluded that — in cases where the intersubject variability results in high intraclass correlation coefficients irrespective of assay coefficients of variation — the method has potential utility in epidemiologic studies. Thus, in considering the two cytokines with low, measurable levels in our own study, caution is urged in measurement of IFN- α 2 from CVL specimens, whereas the high intraclass correlation coefficients we report for IL-1 β are reassuring for measurement of that cytokine.

To our knowledge, this is the first study of the interlaboratory reproducibility of cytokine measurements by Luminex using clinical, cervicovaginal specimens. In a multicenter study of cytokine immunoassay performance, Fichorova et al. examined the contributions of interlaboratory variability, matrix effect, and assay method on recovery, using recombinant reference standards for IL-1 β and IL-6 spiked into different matrices [24]. The authors concluded that, in the commercial Luminex kit studied, interlaboratory reproducibility is good for IL-1 β (able to detect a 1.84-fold difference between measurements performed in different laboratories), but less so for IL-6 (able to detect only a 6.5-fold or higher difference). The relative contributions to that variability of manufacturing lot, software package, and curve-fit model were not addressed. They reported that recoveries are better for both cytokines when prepared in saline (as was used for CVL collection in the present study) than in phosphate-buffered saline, highlighting an important matrix effect of specimen collection medium. Another important conclusion from that study was that biologically active reference standards or endogenous cytokines should be used to validate assay performance and reproducibility, rather than the calibrators included with assay kits. Our results, using clinical study specimens, confirm their findings for IL-1 β but differ for IL-6.

Although the intraclass correlation coefficients reported here suggest good agreement for several cytokines, and thus potential utility for the Luminex platform in microbicide safety studies, they do not provide a context for evaluating whether that level of agreement is sufficient. For a biomarker assay to be useful, it must have a level of reproducibility (interlaboratory and other) that allows one to distinguish biologically meaningful changes in expression from background variability. Adopting concepts and terminology from Lee et al. [25], assay validation is best regarded as an iterative process and is intertwined with biomarker "qualification" (i.e., identification of specific biomarkers that can serve as acceptable surrogates for an endpoint of interest). Part of the biomarker qualification process for microbicide safety studies will entail defining meaningful thresholds of change and should include revisiting the question of reproducibility. Thus, as candidate biomarkers of microbicide safety are identified and characterized, assay acceptance criteria must include demonstration that the fold differences that can be reliably measured within and between laboratories allow clinically meaningful changes to be detected. Whether a centralized laboratory is needed will depend on interlaboratory variability as evaluated within that context. Irrespective of whether multiple laboratories are used, there appears to be broad support in the literature for selection of a single assay kit vendor for use throughout a study [13–15] as well as support for using either biologically active reference standards or actual clinical specimens for validation [24]. Our results also support testing pre- and post-product specimens from a given subject in the same laboratory. Lastly, we recommend that either the same software package be used for curve fitting, or that software packages be employed that allow exchange of raw fluorescence data for cross-laboratory reanalysis and validation. If

such cross-validation were to indicate interoperator variability, curve fitting and unknown interpolation of data from different laboratories could then be centralized.

Highlights

- > Three laboratories measure cytokines in cervicovaginal lavage samples by Luminex
- > IFN- γ , IL-10, IL-17 and TNF are below detection in a majority of CVL samples
- > IL-1 β , IL-6 and IL-8, but not IFN- α 2, show good agreement in absolute measurements
- > IL-1 β and IL-6 show good agreement in within-subject change after microbicide gel use
- > Cytokine measurement by Luminex has potential utility in microbicide safety studies

Abbreviations

CVL	cervicovaginal lavage
HIV	human immunodeficiency virus

Acknowledgments

The authors thank Dr. Craig Cohen, protocol chair of the VivaGel Phase I trial, for making specimens available for this study and for facilitating pre-submission review of the manuscript by the STI Clinical Trials Group Executive Committee. This study was supported by the National Institutes of Health (NIH) grants AI065309 from the National Institute of Allergy and Infectious Diseases (NIAID) and R37 CA051323 from the National Cancer Institute, and by the STI Clinical Trials Group (NIAID Division of Microbiology and Infectious Diseases HHSN266200400074C). In addition, this research was supported by NIH/NCRR UCSF-CTSI grant number UL1 RR024131 from the National Center for Research Resources (NCRR). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. Information on NCRR is available at <http://www.ncrr.nih.gov>. Information on Re-engineering the Clinical Research Enterprise can be obtained from <http://nihroadmap.nih.gov/clinicalresearch/overview-translational.asp>.

References

1. Abdool Karim Q, Abdool Karim SS, Frohlich JA, Grobler AC, Baxter C, Mansoor LE, et al. Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. *Science*. 2010; 329:1168–1174. [PubMed: 20643915]
2. Feldblum PJ, Adeiga A, Bakare R, Wevill S, Lendvay A, Obadaki F, et al. SAVVY vaginal gel (C31G) for prevention of HIV infection: a randomized controlled trial in Nigeria. *PLoS ONE*. 2008; 3:e1474. [PubMed: 18213382]
3. Van Damme L, Ramjee G, Alary M, Vuylsteke B, Chandeying V, Rees H, et al. Effectiveness of COL-1492, a nonoxynol-9 vaginal gel, on HIV-1 transmission in female sex workers: a randomised controlled trial. *Lancet*. 2002; 360:971–977. [PubMed: 12383665]
4. Van Damme L, Govinden R, Mirembe FM, Guédou F, Solomon S, Becker ML, et al. Lack of effectiveness of cellulose sulfate gel for the prevention of vaginal HIV transmission. *N Engl J Med*. 2008; 359:463–472. [PubMed: 18669425]
5. Fichorova RN, Tucker LD, Anderson DJ. The molecular basis of nonoxynol-9-induced vaginal inflammation and its possible relevance to human immunodeficiency virus type 1 transmission. *J Infect Dis*. 2001; 184:418–428. [PubMed: 11471099]

6. Galen BT, Martin AP, Hazrati E, Garin A, Guzman E, Wilson SS, et al. A comprehensive murine model to evaluate topical vaginal microbicides: mucosal inflammation and susceptibility to genital herpes as surrogate markers of safety. *J Infect Dis.* 2007; 195:1332–1339. [PubMed: 17397004]
7. Cheshenko N, Keller MJ, MasCasullo V, Jarvis GA, Cheng H, John M, et al. Candidate topical microbicides bind herpes simplex virus glycoprotein B and prevent viral entry and cell-to-cell spread. *Antimicrob Agents Chemother.* 2004; 48:2025–2036. [PubMed: 15155195]
8. Mesquita PM, Cheshenko N, Wilson SS, Mhatre M, Guzman E, Fakioglu E, et al. Disruption of tight junctions by cellulose sulfate facilitates HIV infection: model of microbicide safety. *J Infect Dis.* 2009; 200:599–608. [PubMed: 19586414]
9. Wilson SS, Cheshenko N, Fakioglu E, Mesquita PM, Keller MJ, Herold BC. Susceptibility to genital herpes as a biomarker predictive of increased HIV risk: expansion of a murine model of microbicide safety. *Antivir Ther.* 2009; 14:1113–1124. [PubMed: 20032541]
10. Liang SC, Tan XY, Luxenberg DP, Karim R, Dunussi-Joannopoulos K, Collins M, et al. Interleukin (IL)-22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides. *J Exp Med.* 2006; 203:2271–2279. [PubMed: 16982811]
11. Feinen B, Jerse AE, Gaffen SL, Russell MW. Critical role of Th17 responses in a murine model of *Neisseria gonorrhoeae* genital infection. *Mucosal Immunol.* 2010; 3:312–321. [PubMed: 20107432]
12. Cohen CR, Moscicki AB, Scott ME, Ma Y, Shiboski S, Bukusi E, et al. Increased levels of immune activation in the genital tract of healthy young women from sub-Saharan Africa. *AIDS.* 2010; 24:2069–2074. [PubMed: 20588163]
13. Djoba Siawaya JF, Roberts T, Babb C, Black G, Golakai HJ, Stanley K, et al. An evaluation of commercial fluorescent bead-based luminex cytokine assays. *PLoS ONE.* 2008; 3:e2535. [PubMed: 18596971]
14. Khan SS, Smith MS, Reda D, Suffredini AF, McCoy JP. Multiplex bead array assays for detection of soluble cytokines: comparisons of sensitivity and quantitative values among kits from multiple manufacturers. *Cytometry B Clin Cytom.* 2004; 61:35–39. [PubMed: 15351980]
15. Liu MY, Xydakis AM, Hoogeveen RC, Jones PH, Smith EO, Nelson KW, et al. Multiplexed analysis of biomarkers related to obesity and the metabolic syndrome in human plasma, using the Luminex-100 system. *Clin Chem.* 2005; 51:1102–1109. [PubMed: 15976097]
16. Bland JM, Altman DG. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. *Comput Biol Med.* 1990; 20:337–340. [PubMed: 2257734]
17. Hanley B. Variance in multiplex suspension array assays: carryover of microspheres between sample wells. *J Negat Results Biomed.* 2007; 6:6. [PubMed: 17459164]
18. Hanley BP, Xing L, Cheng RH. Variance in multiplex suspension array assays: microsphere size variation impact. *Theor Biol Med Model.* 2007; 4:31. [PubMed: 17716372]
19. Hanley BP. Variance in multiplex suspension array assays: a distribution generation machine for multiplex counts. *Theor Biol Med Model.* 2008; 5:3. [PubMed: 18226252]
20. Hanley B. Variance in multiplex suspension array assays: intraplex method improves reliability. *Theor Biol Med Model.* 2007; 4:32. [PubMed: 17727693]
21. Gottschalk PG, Dunn JR. The five-parameter logistic: a characterization and comparison with the four-parameter logistic. *Anal Biochem.* 2005; 343:54–65. [PubMed: 15953581]
22. Applied Cytometry Systems. [Date accessed: February 25, 2011] Technote #24: Can I import data from the Luminex software into SStarStation for data analysis?. 2007. http://www.appliedcytometry.com/Technotes/Data_Conversion/24.html
23. Wong HL, Pfeiffer RM, Fears TR, Vermeulen R, Ji S, Rabkin CS. Reproducibility and correlations of multiplex cytokine levels in asymptomatic persons. *Cancer Epidemiol Biomarkers Prev.* 2008; 17:3450–3456. [PubMed: 19064561]
24. Fichorova RN, Richardson-Harman N, Alfano M, Belec L, Carbonneil C, Chen S, et al. Biological and technical variables affecting immunoassay recovery of cytokines from human serum and simulated vaginal fluid: a multicenter study. *Anal Chem.* 2008; 80:4741–4751. [PubMed: 18484740]

25. Lee JW, Devanarayan V, Barrett YC, Weiner R, Allinson J, Fountain S, et al. Fit-for-purpose method development and validation for successful biomarker measurement. *Pharm Res.* 2006; 23:312–328. [PubMed: 16397743]

Table 1

Cytokine levels in CVL specimens at baseline.

	Laboratory A		Laboratory B		Laboratory C	
	BDL ^a	Median (IQR) ^b	BDL ^a	Median (IQR) ^b	BDL ^a	Median (IQR) ^b
IL-1β	4 (33%)	11.0 (1.50, 73.6)	2 (17%)	9.19 (3.72, 52.1)	3 (25%)	8.85 (3.02, 112)
IL-6	2 (17%)	44.6 (18.4, 50.4)	0 (0%)	23.6 (10.0, 34.1)	0 (0%)	33.3 (23.9, 48.5)
IL-8	0 (0%)	590 (282, 913)	0 (0%)	1,270 (240, 1,163)	0 (0%)	779 (142, 1,163)
IL-10	11 (92%)	ND	10 (83%)	ND	6 (50%)	ND
IL-17	12 (100%)	ND	8 (67%)	ND	10 (83%)	ND
IFN-α2	7 (58%)	1.50 (1.50, 6.30)	0 (0%)	9.00 (5.69, 10.5)	0 (0%)	17.3 (9.53, 25.5)
IFN-γ	8 (67%)	ND	10 (83%)	ND	9 (75%)	ND
TNF	11 (92%)	ND	11 (92%)	ND	11 (92%)	ND

^aNumber (percent) of 12 baseline (day-0) samples below detection limit (BDL) of 3.2 pg/ml.

^bMedian (IQR) cytokine levels at baseline, in pg/ml, of averaged duplicate determinations, with below-detection values recoded as 1.5 pg/ml; ND, not determined.

Table 2

Cytokine levels below detection limits in CVL specimens.

	1 of 3 laboratories^a	2 of 3 laboratories^a	3 of 3 laboratories^a	2 or 3 laboratories^a
IL-1 β	5 (14%)	4 (11%)	4 (11%)	8 (22%)
IL-6	7 (19%)	0 (0%)	3 (8%)	3 (8%)
IL-8	3 (8%)	0 (0%)	1 (3%)	1 (3%)
IL-10	3 (8%)	7 (19%)	24 (67%)	31 (86%)
IL-17	3 (8%)	3 (8%)	30 (83%)	33 (92%)
IFN- α 2	15 (42%)	8 (22%)	1 (3%)	9 (25%)
IFN- γ	3 (8%)	7 (19%)	24 (67%)	31 (86%)
TNF	0 (0%)	0 (0%)	33 (92%)	33 (92%)

^aNumber (percent) of 36 samples in which cytokine measurement was below the detection limit of 3.2 pg/ml in exactly the indicated number of laboratories.

Table 3Interlaboratory agreement in absolute cytokine measurements.^{a,b}

	Overall	By Study Day		
		Day 0	Day 7	Day 14
IL-1 β	0.92 (0.87, 0.96)**	0.88 (0.71, 0.96)**	0.90 (0.75, 0.97)**	0.97 (0.91, 0.99)**
IL-6	0.90 (0.83, 0.94)**	0.83 (0.61, 0.94)**	0.93 (0.84, 0.98)**	0.90 (0.77, 0.97)**
IL-8	0.72 (0.57, 0.83)**	0.32 (-0.05, 0.71)*	0.74 (0.46, 0.91)**	0.86 (0.68, 0.95)**
IFN- α 2	0.47 (0.27, 0.66)**	0.19 (-0.13, 0.60)	0.72 (0.43, 0.90)**	0.49 (0.13, 0.79)*

^aIntraclass correlation coefficient, r (95% CI).^bP-value from hypothesis test that r is different from zero (F-test):* $p \leq 0.05$,** $p \leq 0.001$.

Table 4Interlaboratory agreement in within-subject change in cytokine measurements.^{a,b}

	$\Delta 7^c$	$\Delta 14^c$
IL-1 β	0.71 (0.42, 0.90)**	0.86 (0.67, 0.95)**
IL-6	0.79 (0.55, 0.93)**	0.86 (0.67, 0.95)**
IL-8	0.10 (0.20, 0.53)	0.51 (0.46, 0.80)*
IFN- $\alpha 2$	0.40 (0.04, 0.74)*	0.36 (0.00, 0.72)*

^aIntraclass correlation coefficient, r (95% CI).^bP-value from hypothesis test that r is different from zero (F-test):* $p \leq 0.05$,** $p \leq 0.001$.^cChange from baseline (day 0) at days 7 and 14, calculated by subtracting log-transformed baseline levels for each subject from log-transformed day-7 and day-14 levels.