

## COMMENTARY

# Know Your Variability: Challenges in Mechanistic Modeling of Inflammatory Response in Inflammatory Bowel Disease (IBD)

Katharine V. Rogers<sup>1</sup>, Indranil Bhattacharya<sup>1</sup>, Steven W. Martin<sup>2</sup> and Satyaprakash Nayak<sup>2,\*</sup>

## INTRODUCTION

Variability in biomarker measurements and their lack of reproducibility is a widely acknowledged concern in clinical and preclinical research. The development of quantitative systems pharmacology (QSP) models for understanding diseases and therapeutics requires integration of data from multiple sources. Since QSP models are heavily dependent on literature-derived biomarker information, uncertainty and variability in biomarker data can have a significant effect on predictive capabilities of the model and its overall applicability to answer research questions.

There is an impetus towards development of QSP models to advance understanding of therapeutics in individual patients, develop tools linking preclinical and clinical studies, and translation between biomarkers and disease outcome.<sup>1</sup> Disease-level models based on biological mechanisms are often constructed using systems of ordinary differential equations and have many unknowns, such as levels of proteins modeled and kinetic rate parameters, that need to be estimated from physiological ranges or from multiple data sources.

The poor reproducibility of preclinical research and the limitations of preclinical tools to translate to the clinical space have been noted in the high failure rate of clinical trials.<sup>2</sup> In a recent commentary in this journal,<sup>3</sup> the author listed inadequate assay condition documentation, improper randomization, and batch effects as the main sources of bias and variability that contribute to irreproducibility of biomarkers and reduce confidence in their applicability to derive robust and meaningful conclusions from clinical studies.

Currently, there is an immense interest in QSP models capturing the modulation of cytokines and cell-types in inflammatory diseases, such as rheumatoid arthritis, psoriasis, and inflammatory bowel disease (IBD). These diseases differ in multiple ways, but all represent a deregulated state of the immune system that can be modeled by assimilating information regarding the levels of cytokines and cells and their interactions. In particular, this commentary is focused on the variability of biomarkers observed in healthy subjects and IBD patients. The two main subtypes of IBD, Crohn's disease

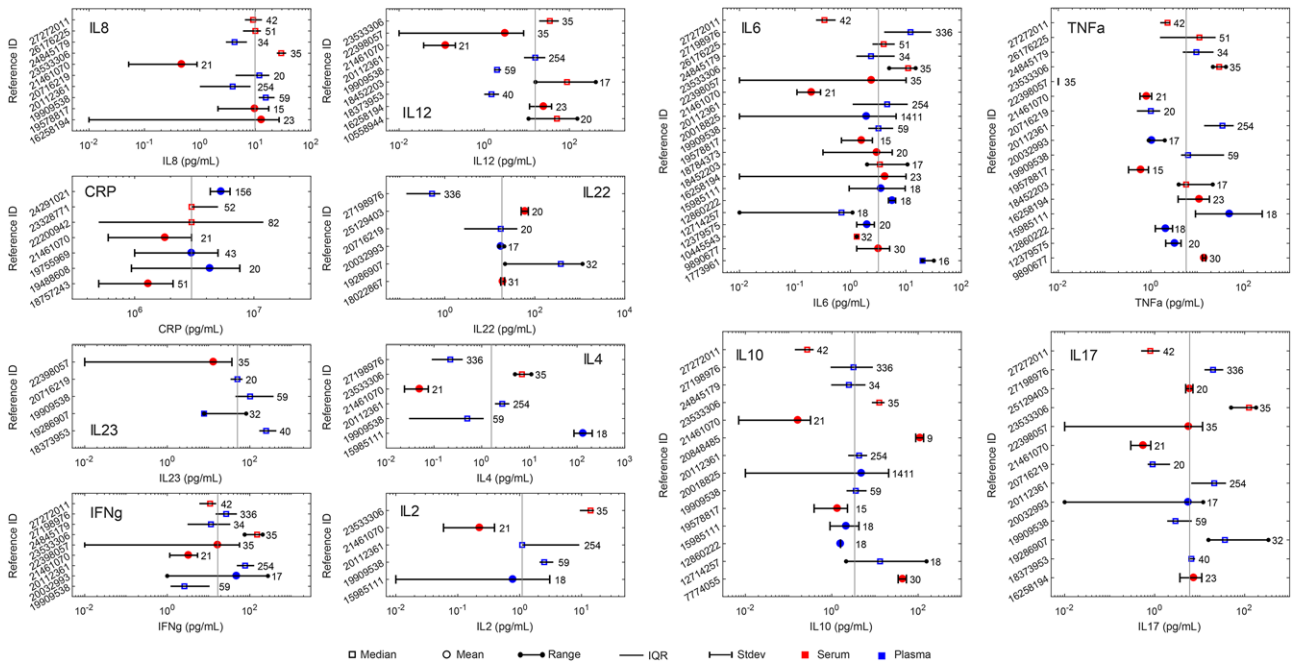
(CD) and ulcerative colitis (UC), differ in the extent and location of inflammation. A thorough review of the literature was conducted to extract data for cytokines and cell numbers in plasma of healthy adults and IBD patients and large variability between studies was seen. Variability in the biomarker data arises from the assay method used, interlab variability, sample storage, sample size, and the analysis method.<sup>4</sup> In this commentary, we present our views on the challenges faced due to variability in biomarkers when constructing and parameterizing a QSP model of immunology.

## VARIABILITY FROM DATA GENERATION METHODS

A major hindrance to the reproducibility of immunology biomarker data originates from the different sample collection methods. Protein biomarkers have been measured using assays that measure either single-analytes or a protein panel. Early (pre-2005) biomarker studies predominately used enzyme-linked immunosorbent assays (ELISAs), providing a limited scope of biomarkers per publication, while multiplex tests now allow for multiple proteins to be measured concurrently. Different methods led to discrepancies in the absolute cytokine concentration (also seen previously<sup>5</sup>), although the ability to measure multiple samples in a single test did not increase the variability of the individual tests (**Supplementary Figure 1**). Breen *et al.* showed the variation of cytokine concentrations measured by multiplex cytokine assay platforms was due to different tests, interlot and interlab variability and recommended using relative values instead of absolute changes.<sup>5</sup>

For flow cytometry measurements, an additional source of variability comes from the definition of markers for the specified cell types. Subsets of immune cells can be defined by cell-surface proteins, intracellular phosphoproteins, or intracellular cytokines. One research article shows how using either cytokines or transcription factors to define T-helper (Th) cell subsets gives significant differences in the percent of cells calculated.<sup>6</sup> The percentage of Th17 cells in the same CD patients differs from ~8% vs. 62% of CD4+ T cells, when defined by an intracellular cytokine, interleukin (IL)-17, or a phosphoprotein, ROR $\gamma$ t+, respectively.<sup>6</sup> Thus,

<sup>1</sup>Clinical Pharmacology, Early Clinical Development, Pfizer Inc., Groton, Connecticut, USA; <sup>2</sup>Pharmacometrics, Global Innovative Pharma Business (GIPB), Pfizer Inc., Groton, Connecticut, USA. \*Correspondence: Satyaprakash Nayak ([Satyaprakash.Nayak@pfizer.com](mailto:Satyaprakash.Nayak@pfizer.com))  
Received 22 June 2017; accepted 16 August 2017; published online on 6 October 2017. doi:10.1111/cts.12503



**Figure 1** Cytokine variability across studies in healthy population. Cytokine levels from healthy adult volunteers vary greatly across multiple studies and multiple cytokines. The way the data are reported is also study-dependent, with either mean (circle) or median (square) of the population. The distribution of biomarker values is reported as interquartile range (blank), minimum and maximum range (dots), or standard deviation (error-bars). Reference ID is the PubMed ID, numbers to the right of line denote the number of subjects in the study, and the red and blue denote measurement in serum and plasma, respectively. The gray vertical line is the median value across all studies.

differences in tags used to define cell types can create confusion when publications are read without understanding the impact of labeling. To use flow cytometry data from multiple sources for mechanistic modeling of cell subtypes, the standardization of flow cytometry methods is required, and others have recommended the adoption of standard reagents and protocols, improved technologies, and the centralization of data.<sup>7</sup>

Another source of variability in biomarker levels comes from specificity of the test used to detect low levels of cytokine concentrations. In some studies most samples fall under the limit of detection (LOD) of the assay, e.g., IL-6 was detected in 18 of 21 CD patients, 2 of 20 UC patients, and 2 of 16 healthy subjects by an ELISA with an LOD of 20 pg/mL.<sup>8</sup> Modern assay techniques seem to have alleviated this problem to an extent, as tests done after year 2000 have reported lower concentrations and a distribution in healthy subjects for IL-6. Even with large-scale “omic” data the LOD needs to be considered in study design, e.g., when using proteomics some important inflammation markers IL-2, IL-4, and tumor necrosis factor alpha (TNF $\alpha$ ) are often under the LOD.<sup>4</sup> Thus, in experiments studying the dynamics of cytokine treatment where most samples are under the LOD, the conclusions derived from such data may not be robust.

### VARIABILITY IN DATA ANALYSIS AND REPORTING

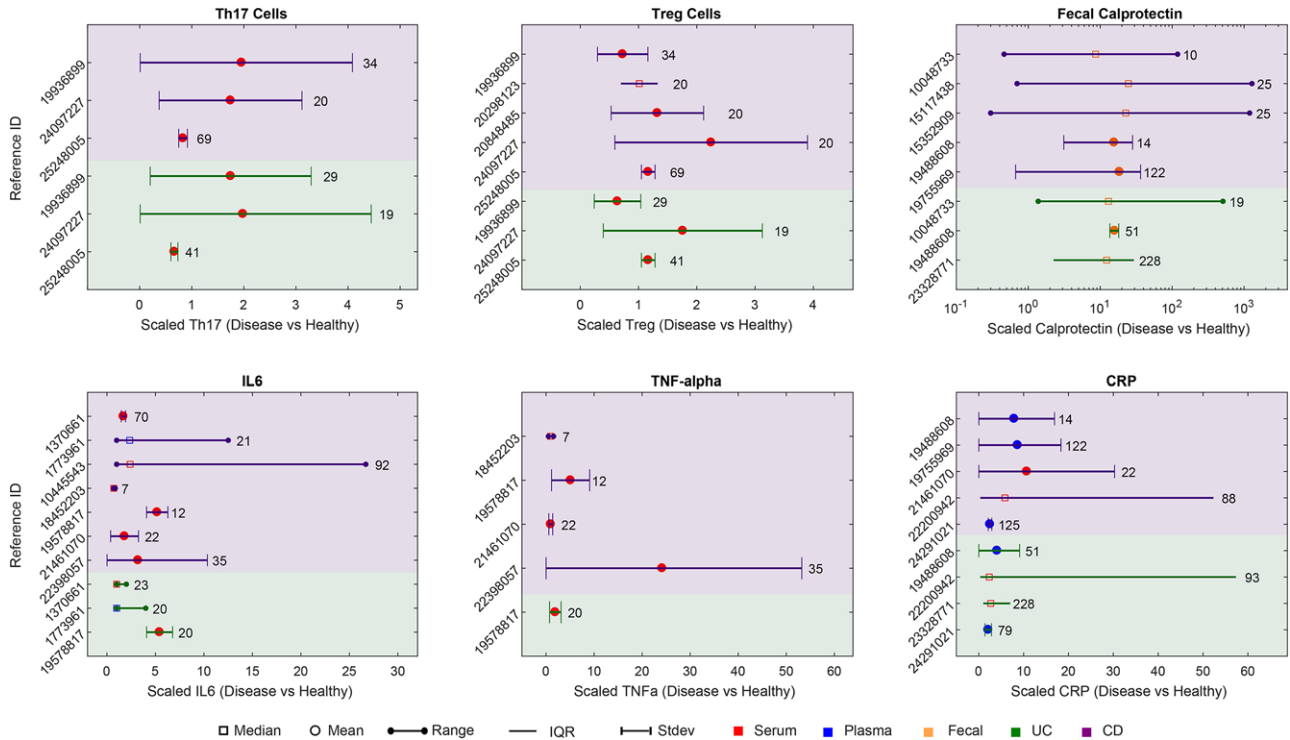
An additional reason for difficulty in comparing biomarker level data across studies is due to the diverse ways in which the results are reported. The cytokine concentrations and cell numbers for subject groups are often reported as mean or

median, with standard deviation, interquartile range, or range (**Figure 1**). Inconsistency in grouping of subjects, i.e., responders vs. nonresponders, active vs. inactive, and CD and UC combined into IBD was also a major confounder in comparing results across studies. Grouping of patients based on response, disease activity, etc., is clearly important for analysis and reporting, but to maximize the use of these data in systems models, one must be able to ungroup the data to study their effect individually. In addition, often patient history such as previous therapies, location of the disease, surgery, and age are published as group summary data, thus limiting the ability of the model to match mean population data without substantial information about covariates at an individual level.

A common inconsistency in data reporting was seen around the consideration of the LOD, with some studies assuming samples under the LOD as 0 pg/mL and others using the LOD of the test, e.g., Ogawa *et al.* assumed values under the LOD to be null even though the sensitivity for the IL-23 assay was 15 pg/mL.<sup>9</sup> Another source of uncertainty could occur from the reporting of biomarker levels in different units, especially in the case of cell-type data reported either in cells/mL or as a percent of a parent cell population.

### SUBJECT VARIABILITY

Even in the absence of variability in data generation, there would still be a level of variability in immunology biomarkers due to natural causes. Naturally occurring variations between individuals occur due to a variety of factors such as age, seasonal variations, gender heritable influences, microbiota,



**Figure 2** Biomarker fold change variation across studies in an IBD population. The figure shows mean (or median) biomarker level scaled with respect to the mean (or median) value in the healthy control group within the study. A fold change value of less than one indicates a decrease in the biomarker level in the patient population and greater than one indicates an increase in the biomarker level. Green lines denote ulcerative colitis patients and purple lines denote Crohn's disease patients.

viruses, and the environment.<sup>10</sup> By contrast, intraindividual variation was observed to be low, as healthy adults have stable immune cell frequencies and serum protein levels, except in the case of an acute immune response.<sup>10</sup> These naturally occurring variations are difficult to control, but the reporting of subject level data with age, sex, and environment may allow for a better understanding of the data.

In growing therapeutic areas, like immunology, biomarkers are typically exploratory end points, often creating a sparsity of data that must be considered while building a model and evaluating model predictions. Exploratory biomarkers are typically investigated in early-stage clinical trials where the total sample sizes rarely go above 100 subjects and are often spread across multiple treatment groups (Figure 1). It is only in phase II trials where subject numbers can reach into the 100's for treatment groups, and these studies often have limited biomarker collection. To incorporate the effect of study size, it would be prudent to report standard error when meta-analyzing studies to account for this factor. Another concern is an inconsistent definition of control populations. For many IBD studies, particularly those with tissue samples, the control subject is defined by not having IBD, but often subjects come in for diagnostic procedures, including screening for polyps or cancer. In some studies irritable bowel syndrome (IBS) patients are a separate disease group, while in others IBS patients are controls, showing inconsistencies in classification of subjects across studies.

## DATA VARIABILITY AND QSP MODELING

A summary of important protein biomarker levels in healthy subjects from the literature showed the wide range of values reported and the differences in assay conditions and reporting standards (Supplementary Table 1). The large variability observed between studies emphasized the challenges in constructing a mechanistic model based on such data (Figure 1). A similar survey of biomarkers in CD and UC reveals that there is also high variability in the levels reported in disease populations. Figure 2 shows the broad inconsistency observed in scaled values of key biomarkers in IBD, which underscores the challenge in defining a disease baseline state. This variability makes QSP modeling especially challenging, as it relies substantially on literature information for model construction and parameter estimation. This is particularly true in the rapidly evolving area of immune systems modeling compared with other therapeutic areas such as diabetes with well-established pathways and better understanding of biomarker variability. However, there are approaches available to mitigate this challenge in modeling, e.g., systematic meta-analysis tools, such as inverse variance weighting, can be used to better estimate a central tendency and spread from different studies. A virtual population simulation can also aid in understanding the variability and accuracy of model predictions. If the differences in biomarker levels are truly representative of diverse populations, they can be incorporated in the model using different parameter sets. Approaches such

as sensitivity analysis should also be used to determine the uncertainty of the model prediction by varying key parameters.

## CONCLUSION

Biomarker data of immune system components reported in the literature were examined, particularly in the context of mathematical modeling in IBD, and a large variation in levels of biomarkers being reported for healthy and diseased populations was observed. Towards achieving the goal of personalized medicine, it is important to understand the translatability of individual biomarker data from multiple sources. It is also essential to comprehend the variability of the data and their impact with regard to application in tools such as QSP models to predict individual response. Although some degree of variability in biomarker levels is expected as a result of natural causes and assay limitations, the standardization of biomarker level reporting protocols will enhance their applicability towards meta-analysis tools and mechanistic disease modeling and allow for deriving robust conclusions.

**Acknowledgment.** We thank Gianluca Nucci, PhD for discussions and critical review of the article.

**Conflict of Interest.** All authors are or were employed by Pfizer Inc. during the conduct of the research.

1. Sorger, P.K. & Allerheiligen, S.R.B. Quantitative and systems pharmacology in the post-genomic era: new approaches to discovering drugs and understanding therapeutic

mechanisms. An NIH White Paper by the QSP Workshop Group 2011; Available from: <http://www.nigms.nih.gov.proxy1.athensams.net/NR/rdonlyres/8ECB1F7C-BE3B-431F-89E6-A43411811AB1/0/SystemsPharmaWPSorger2011.pdf>.

2. Begley, C.G. & Ellis, L.M., Drug development: Raise standards for preclinical cancer research. *Nature* **483**(7391), 531–533 (2012).
3. McShane, L.M., In pursuit of greater reproducibility and credibility of early clinical biomarker research. *Clin. Transl. Sci.* **10**(2), 58–60 (2017).
4. Enroth, S., Hallmans, G., Grankvist, K. & Gyllenstein, U. Effects of long-term storage time and original sampling month on biobank plasma protein concentrations. *EBioMedicine* **12**, 309–314 (2016).
5. Breen, E.C. et al. Multisite comparison of high-sensitivity multiplex cytokine assays. *Clin. Vaccine Immunol.* **18**(8), 1229–1242 (2011).
6. Li, J. et al. Profiles of lamina propria T helper cell subsets discriminate between ulcerative colitis and Crohn's disease. *Inflamm. Bowel Dis.* **22**(8), 1779–1792 (2016).
7. Maecker, H.T., McCoy, J.P. & Nussenblatt, R., Standardizing immunophenotyping for the Human Immunology Project. *Nat. Rev. Immunol.* **12**(3), 191–200 (2012).
8. Mahida, Y.R., Kurlac, L., Gallagher, A. & Hawkey, C.J. High circulating concentrations of interleukin-6 in active Crohn's disease but not ulcerative colitis. *Gut* **32**(12), 1531–1534 (1991).
9. Ogawa, K., Matsumoto, T., Esaki, M., Torisu, T. & Iida, M. Profiles of circulating cytokines in patients with Crohn's disease under maintenance therapy with infliximab. *J. Crohns Colitis* **6**(5), 529–535 (2012).
10. Brodin, P. & Davis, M.M., Human immune system variation. *Nat. Rev. Immunol.* **17**(1), 21–29 (2017).

© 2017 The Authors. Clinical and Translational Science published by Wiley Periodicals, Inc. on behalf of American Society for Clinical Pharmacology and Therapeutics. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Supplementary information accompanies this paper on the *Clinical and Translational Science* website. ([http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1752-8062](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1752-8062))