Statistical strategies to analyze multiparametric assessments of innate immune response modulating impurities (IIRMI)

Logan Kelley-Baker, Cheng Her, Mohanraj Manangeeswaran, Derek Ireland, Daniela Verthelyi, Seth Thacker,
Laboratory of Immunology, OBP, CDER, FDA

Abstract

Our lab and others have shown that trace levels of process related impurities derived from host cells, adventitious agents, or raw materials, when co-delivered with a product, can act as adjuvants modifying the milieu where a product is delivered, inducing local inflammation and recruitment of antigen presenting cells and enhancing the formation of germinal centers in the draining lymph nodes. We termed these innate immune response modulating impurities (IIRMI) and developed a series of in vitro assays to monitor them. The readout for these assays consists of mRNA or protein levels for one or more cytokines linked to innate immune responses (e.g. TNF and IL-6), complex panels of RNA for innate immune related genes, or activation markers on dendritic cells. In this study, we explore different statistical strategies to process the resulting data in order to identify generic products as inducing a different innate immune response pattern relative to the reference licensed product tested in clinical trials. We assess the results using strategies that compare the number of markers induced by the products in peripheral blood cells from each donor, and/or the magnitude of each response in the context of the gene and/or donor. Further, we then try to quantify that difference, to ascertain if is relevant or negligible. Our studies will provide reviewers with multiple strategies to probe the results for assays monitoring the presence of innate immune response modulating impurities in peptides and proteins.

Introduction

Protein and peptide therapeutics are commonly used to treat multiple serious conditions. The development of immune responses to these products, even when these have human sequences, are frequent and can significantly affect the safety and efficacy of the product. Innate immune response modulating impurities (IIRMI) such as product aggregates, residual solvents, host cell proteins, remnants of adventitious agents or leachates from the manufacturing train are risk factors of immunogenicity as they can act as adjuvants facilitating the induction of an immune response to the product.

Improvements in analytical chemistry and product characterization are enabling new regulatory paths for peptides and proteins that could require fewer (351(k), 505(b)(2)) or no clinical studies (505(j)), however residual uncertainty over the potential of product and process related impurities to impact on the immunogenicity risk hinder these regulatory decisions. Our lab and others have established assays to monitor the IIRMI presence using the differential expression of a variety of genes in the innate immune pathways as biomarkers.[1] Strategies to integrate the genes expression data to make regulatory decisions are lacking.

Critical parameters in these assays include: (1) The range of genes interrogated. Too narrow a range of genes could miss critical innate immune activation markers, where as too large a screen may miss a pointed response. (2) The number of samples tested, given the high variability of in vitro tests with primary cells.

Here we used a larger gene panel (30- genes) and 22 PBMC to explore differences in gene expression as well as in donor activation between a generic and its reference licensed product. We find that the use of larger panel of genes can improve the reliability of the tests being performed and propose novel testing strategies to combine expression outputs per donor, consider gene expression covariance and synthesize the effect size using a modified principal component analysis.

References

1) Haile LA, Puig M, Polumuri SK, Ascher J, and Verthelyi D. InVivo Effect of Innate Immune Response Modulating Impurities on the Skin Milieu Using a Macaque Model: Impact on Product Immunogenicity. Journal of pharmaceutical sciences. 2017;106(3):751-60.
2) R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
3) Cinar O, Viechtbauer W (2022). "The poolr Package for Combining Independent and Dependent \$p\$ Values." _Journal of Statistical Software_, *101*(1), 1-42. doi:10.18637/jss.v101.i01 <https://doi.org/10.18637/jss.v101.i01>.
4) Strube, M. J. (1985). Combining and comparing significance levels from nonindependent hypothesis tests. *Psychological Bulletin*, *97*(2), 334–341.

Experimental Design

- PBMCs were isolated from healthy donors and plated at 5x10⁶ cells/ml. Cells were incubated with media, the RLD, or corresponding follow on product for 24 hours after which RNA was isolated using the Trizol method per manufactures recommendations. 500ng of total RNA was used to transcribe cDNA. Gene expression was quantified by qPCR on a TLDA plate.
- Statistical analysis was performed in R 4.2.0.[2]
- Sample integrity was determined by ensuring that each housekeeping gene was above the detection threshold, and in consistent ratios between samples.
- The expression of housekeeping genes was used to transform CT values into dCT. Calculated log2 FC by subtracting the untreated or the RLD dCT from the test product dCT. Very Extreme values (>6 IQR) and missing values were converted to a logFC of zero.
- Single gene analysis and multiple gene analysis were performed, as specified in their respective sections.

Single Gene Analysis

- Single gene analysis consisted of performing one-sided Wilcoxon signed rank test was performed with $\mu = 0$ on the logFC data, p-values were left unadjusted to simulate single gene analysis.

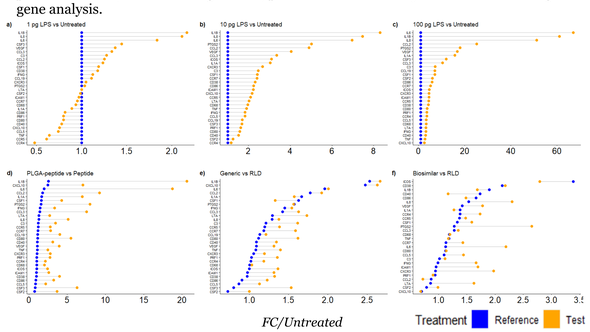

Figure 1. The inflammatory response of PBMC as seen in the FC relative to untreated samples of a test/potentially impure product [a-c) LPS, d) Generic Peptide vs RP e) Peptide vs PLGA encapsulated peptide f) Biosimilar vs RLD]. Genes were organized from highest expression in the reference product to lowest.

Gene	1 pg LPS	10 pg LPS	100 pg LPS	PLGA Peptide vs peptide	Generic vs RLD	Biosimilar vs RLD
IL1A	1.84E-01	**2.80E-03**	**1.14E-05**	**2.38E-05**	3.51E-01	5.92E-01
IL1B	**9.12E-03**	**7.63E-06**	**3.81E-06**	**4.77E-06**	2.61E-01	7.52E-01
IL2	6.65E-01	*9.07E-02*	**9.12E-03**	1.01E-01	7.43E-01	4.17E-01
IL6	**1.52E-02**	**3.81E-05**	**3.81E-06**	**9.54E-07**	3.37E-01	**7.24E-03**
IL8	**1.92E-02**	**7.63E-06**	**3.81E-06**	**1.31E-04**	3.37E-01	1.52E-01
IL10	4.44E-01	**1.43E-03**	**2.30E-04**	**5.34E-03**	8.53E-01	7.15E-01
IL17	5.00E-01	**1.05E-02**	**1.05E-02**	**1.55E-02**	3.33E-01	1.47E-01
CCL3	2.21E-01	**9.54E-05**	**7.63E-06**	**1.34E-05**	2.61E-01	5.67E-01
CCL19	3.67E-01	**4.07E-02**	**7.63E-06**	**2.05E-03**	4.08E-01	**9.53E-03**
CCL2	1.73E-01	**3.36E-04**	**7.63E-06**	**2.38E-05**	1.83E-01	6.95E-01
CCL5	7.52E-01	**2.69E-02**	**1.64E-04**	**5.08E-04**	**4.48E-02**	*5.42E-02*
CCR4	9.83E-01	2.09E-01	**7.25E-05**	**1.81E-05**	1.56E-01	6.49E-01
CCR5	8.58E-01	*5.42E-02*	**3.29E-03**	**1.16E-03**	1.31E-01	5.19E-01
CCR7	6.17E-01	**1.68E-03**	**2.67E-05**	**1.97E-04**	**2.66E-02**	*7.29E-02*
CXCR3	4.66E-01	**7.25E-05**	**1.26E-04**	**3.54E-04**	1.74E-01	**6.94E-03**
CXCL10	7.31E-01	**4.42E-02**	1.28E-01	**5.25E-05**	2.73E-01	5.10E-01
CSF1	2.48E-01	**6.45E-04**	**7.63E-06**	**1.55E-03**	7.84E-01	1.72E-01
CSF2	3.18E-01	**3.24E-02**	**1.61E-04**	**7.58E-04**	1.28E-01	6.08E-01
CSF3	5.34E-01	1.01E-01	**2.74E-04**	**8.10E-03**	2.57E-01	1.47E-01
CD38	1.62E-01	**6.02E-03**	**2.10E-04**	**1.34E-05**	**2.00E-02**	1.97E-01
CD40	8.77E-01	*9.82E-02*	**6.45E-04**	**1.34E-05**	1.84E-01	8.31E-01
CD68	7.39E-01	**1.92E-02**	**6.45E-04**	**3.54E-04**	4.93E-01	3.83E-01
CD80	9.41E-01	**3.68E-02**	**1.91E-05**	**1.91E-06**	**1.64E-02**	5.34E-01
CD86	7.25E-01	*6.49E-02*	**4.20E-04**	**3.15E-05**	*7.68E-02*	5.38E-01
ICOS	1.96E-01	**6.45E-04**	**3.81E-05**	**2.58E-03**	**3.07E-02**	6.99E-01
ICAM1	8.17E-01	**1.40E-03**	**3.81E-06**	**8.39E-05**	*8.84E-02*	**2.16E-02**
PTGS2	5.34E-01	**3.81E-06**	**3.81E-06**	**8.45E-04**	*5.17E-02*	*5.17E-02*
IFNG	5.34E-01	*7.08E-02*	**2.80E-03**	**4.01E-03**	3.18E-01	*9.87E-02*
PRF1	6.65E-01	*7.08E-02*	**3.81E-05**	**4.77E-06**	**2.20E-02**	2.27E-01
C3	4.49E-01	**9.65E-04**	**7.63E-06**	**1.36E-03**	1.56E-01	4.66E-01
TNF	8.68E-01	**3.00E-02**	**7.90E-04**	**3.65E-03**	2.05E-01	5.17E-01
LTA	5.17E-01	*5.42E-02*	**1.18E-02**	**1.03E-02**	**3.07E-02**	**4.64E-02**
VEGF	1.96E-01	**3.81E-06**	**7.63E-06**	**2.38E-05**	*8.84E-02*	1.84E-01

Table 1. Single gene analysis, on the 30 genes that appear in at least one of the chosen sets. Each dataset is one of the columns, and the p-values are from a one-sided Wilcoxon-test of the ratio of the data against 0. Text colored yellow and italicized had a p-values of less than 0.1, while text that was colored red and bolded has p-values of 0.05 or less. Notably, p-values are left unadjusted to better emulate what a researcher might see if they looked at 1-3 genes.

Multiple Gene Analysis

- Selected sets of genes that line up with canonical gene pathways
- Donor-based meta analysis: Treated each donor as an individual test, and performed a Wilcoxon test of the logFC for the genes within the set for a donor against the housekeeping genes. Then combined the p-values using Stouffer's method (as implemented in the R library poolr)[3].
- Gene-based Meta analysis: Performed Wilcoxon test for each of the genes within a set, then combined the p-values using Stouffer's method with adjustment for covariance (Generalized adjustment (Strube's) as implemented in poolr).[3, 4]
- MM (Modified Mahalanobis) Distance: Defined the number of canonical pathways within a set that were expected to be activated in at least 2 genes (k). Performed PCA, calculated the standardized distance on the first k dimensions and took the sum of squares.

Figure 2. Gene expression pathways and their component genes. The different sets of genes that we are using, and the genes that compromise them. Gene sets Inflam, APC, Chemo are defined by pathways that are seen in the innate immune response, and they are short for Inflammatory response, APC activation and maturation, and Chemokines. Custom is a set of 20 genes based on historical and scientific knowledge that our lab uses to represent a wider array of responses that could potentially occur due to different impurities. The All set is the collection of genes that were represented in any of the other four pathways.

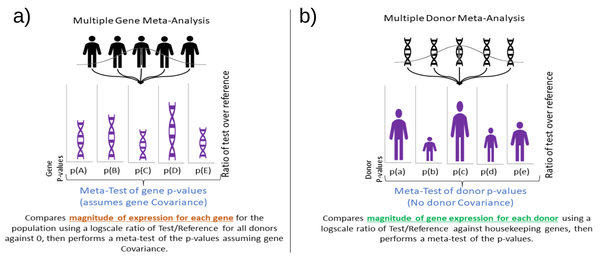

Figure 3. The proposed tests to be used for multiple gene detections of impurity within generic products. The first method (a) treats each gene like an individual test, and combines with Stouffer's test with an adjustment for covariance. The second method (b) treats each donor as an individual test, and combines using Stouffer's test without an adjustment for covariance.

Set	1 pg LPS	10 pg LPS	100 pg LPS	PLGA-Peptide vs peptide	Generic vs RLD	Biosimilar vs RLD
Inflam	1.83E-01	**2.52E-05**	**1.06E-03**	**2.02E-02**	1.90E-01	2.39E-01
	5.21E-02	**1.80E-13**	**2.74E-24**	**2.11E-16**	**4.94E-02**	3.73E-01
APC	6.66E-01	*6.20E-02*	**1.77E-02**	**4.35E-06**	**2.47E-02**	2.33E-01
	9.86E-01	**4.49E-04**	**7.58E-14**	**2.64E-16**	**6.60E-03**	1.97E-01
Chemo	5.09E-01	**1.09E-02**	**4.58E-03**	**1.35E-04**	*6.38E-02*	1.97E-01
	6.35E-01	**6.50E-09**	**1.30E-18**	**2.20E-16**	**1.14E-02**	1.08E-01
Custom	4.24E-01	**4.51E-03**	**4.40E-03**	**5.03E-05**	**3.34E-02**	1.61E-01
	4.04E-01	**4.06E-11**	**4.24E-24**	**1.40E-24**	**1.80E-03**	1.38E-01
All	4.77E-01	**4.42E-03**	**2.55E-03**	**1.58E-04**	**4.59E-02**	1.67E-01
	5.32E-01	**4.00E-11**	**1.33E-25**	**1.62E-23**	**2.48E-03**	1.05E-01

Table 2. Significance values for the two meta-tests performed on each dataset using the five defined sets of genes. The top value of each cell is the p-value obtained from the gene based meta test, and the bottom value is the donor based meta test p-value. Text colored yellow and italicized had a p-value of less than 0.1, while text that was colored red and bolded text had p-values of 0.05 or less.

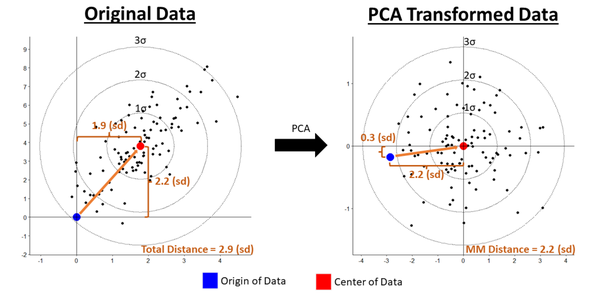

Figure 4. Using simulated data, the figure on the left shows that if we just aggregate the Cohen's d values into one distance function, it fails to account for covariance of the data. To create a covariance adjusted distance, we can convert the data into PCA units, and find the value of Cohen's d between the original origin point and the center of the data in each of the principal components. Mahalanobis distance is typically the aggregate of all the distances in PCA space – and is the standard for multivariate effect size, however it overestimates differences when sample sizes were small. The MM-Distance method uses the first k component, where k is the number of expected activated pathways represented in the chosen set of genes.

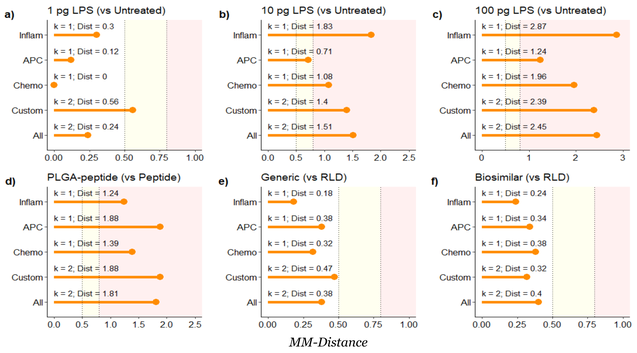

Figure 5. The MM-Distances as calculated for each of the 5 sets of genes for each of the 6 datasets [a-c) LPS, d) Generic Peptide vs RP e) Peptide vs PLGA encapsulated peptide f) Biosimilar vs RLD]. MM-Distance was calculated as explained in figure 2. The MM-Distance and number of expected total sets of canonical pathways activated is displayed above each point. A medium distance thresholds of 0.5-0.8 was visualized with yellow – which we would mean further exploration is necessary. A large distance threshold of >0.8 was visualized with red – which we would take to mean that the inflammatory response elicited by the two products cannot be assumed to be the same.

Conclusion

- Innate and inflammatory responses are complex and cannot be assessed comprehensively by looking at 1-3 genes.
- Meta tests that use multiple genes can reduce biases and assumptions as well as stochastic significant values.
- Genes based on canonical pathways is a logical approach to setting up pools of genes that represent known innate immune paths
- The proposed combination of meta-tests consistently detected gene expression differences.
- The MM-Distance provides a convenient threshold for assessing differences between products.