Carnegie Mellon University

Annual Progress Report: 2009 Formula Grant

Reporting Period

January 1, 2010 – June 30, 2010

Formula Grant Overview

The Carnegie Mellon University received $910,547 in formula funds for the grant award period January 1, 2010 through December 31, 2012. Accomplishments for the reporting period are described below.

Research Project 1: Project Title and Purpose

Mid-Level Feature Representation in Human Visual Cortex - Despite five decades of research, remarkably little is known about the representation of objects in "high-level" human visual cortex. Although both neurophysiology and neuroimaging have informed us about the spatio-temporal properties of early visual cortex (V1->V4), the subsequent encoding of objects within the ventral pathway is largely an unknown. Advances in neuroimaging technologies, including functional Magnetic Resonance Imaging (fMRI), Diffusion Tensor Imaging (DTI), Magnetoencephalography (MEG), and Electroencephalography (EEG) have altered the way vision scientists are able to study object representation. To that end, this project will develop a variety of novel; real-time neuroimaging paradigms and analysis methods designed to study and decode the representation of objects in terms of visual features.

Anticipated Duration of Project

1/1/2010 – 12/31/2012

Project Overview

Human visual object recognition abilities are profoundly better than those of even the most powerful artificial vision systems, yet our understanding of high-level vision in humans remains embarrassingly vague. One of the reasons is a failure to articulate some notion of representation beyond early visual processing. That is, although it is well established that early vision codes for oriented local edges, there is almost no notion of how said edges (and other image features) are combined to represent parts, objects, and scenes. One reason for this lack of theory is that there are, as yet, inadequate computational tools to learn compositional (i.e., hierarchies comprised of reusable parts) structures from real-world images. A second reason is that neurophysiology, including both single-unit and multi-unit recording, is far too narrow in scope to provide a clear picture of how the primate visual system represents visual information. To address these problems in visual coding, we propose a research program that leverages two neuroimaging methods that allow broad coverage of the human brain in action – fMRI and MEG. We will

combine the particular strengths of each of these research tools with state-of-the-art machine learning methods that offer a much more powerful means for exploring and understanding high-dimensional, complex datasets. We will introduce new experimental designs that enable fMRI to function in a real-time system for adaptively assessing neural responses in the visual system and that will enable much better source localization in MEG. For both techniques, coverage of the entire visual system combined with exquisite source localization in fMRI and exquisite temporal resolution in MEG, will allow us to gain a better understanding of the distributed codes used in object and scene representation. There are three facets of this project. First, fMRI and MEG will be used to examine the nature of mid-level features that form the core of object representations. Second, both technologies will be improved with respect to their spatial (MEG) and temporal (fMRI) resolution. Third, a variety of machine learning methods will be applied to more effectively and efficiently identify the optimal stimulus/stimuli for specific brain regions to better identify and understand how the resultant features are combined to support high-level vision.

Principal Investigator

Michael J. Tarr, PhD
Professor and Co-Director, Center for Neural Basis of Cognition
Carnegie Mellon University
115c Mellon Institute, CNBC
4400 Fifth Avenue
Pittsburgh, PA 15213

Other Participating Researchers

None

Expected Research Outcomes and Benefits

Addressing the root causes and concomitant impairments arising from both endogenous (e.g., autism, dyslexia, Alzheimer's, etc.) and exogenous (e.g., TBI, stroke, tumors, etc.) brain disorders requires a detailed account of basic perceptual and cognitive processes from a neural perspective. With respect to the study of high-level vision, the process of object recognition is at the nexus of how we see, interpret, and think about the world around us. Yet there is no comprehensive account of how the human visual system accomplishes this routine and critical action. Thus, a better understanding of how the human brain is able to robustly recognize and interpret objects and scenes in an invariant manner is critical to allowing clinical scientists to more accurately diagnose and more effectively treat a wide range of neural diseases and disorders. For example, several different developmental disorders, including autism, dyslexia, and congenital prosopagnosia manifest, in part, as impairments in visual functioning – advances in understanding the underlying visual mechanisms that are disrupted by such diseases will enable the next generation of behavioral interventions. Likewise, many stroke and TBI patients have impaired visual recognition abilities, the most salient of which is prosopagnosia in which face recognition is rendered non-functional. Elucidating the neural processes whereby face recognition is realized will lead to clearer models as the basis for potential treatments. More generally, the development of improved neuroimaging methods will enhance the study of the

brain across many different cognitive domains: neuroimaging has already revolutionized how we study and ultimately understand the neural bases of cognition, and as such, improvements in these varied techniques have the potential to further enable better discernment of mental processes. It is only through more precise and more coherent models that we will be able to develop better diagnosis and analysis tools for many different brain disorders.

Summary of Research Completed

This research project commenced in May 2010 following project approval and receipt of funds at the end of April 2010. Thus, the reporting period covers only May and June of 2010.

The Carnegie Mellon Scientific Brain Imaging Research Center (SIBR) commenced operations during the first weeks of May 2010. The facilities at this center are integral to the research planned for this project and the initial two months of work has focused on debugging and developing the necessary hardware and software infrastructure necessary to begin both Diffusion Spectrum Imaging (DSI) and fMRI scanning of human subjects at SIBR. Specific issues included porting the DSI sequences developed at the University of Pittsburgh to the somewhat different (and newer) scanner located at SIBR. Having ported the sequences, we conducted a series of validation tests with pilot subjects to ensure that we are obtaining high-quality data for state-of-the-art high-density fiber-tracking. In parallel we have been testing our functional imaging sequences in order to establish optimal imaging parameters with respect to both spatial and temporal resolution, in particular, taking advantage of the parallel imaging using the new 32-channel head coil. Finally, we have been debugging the stimulus presentation component of the SIBR system. The original projector lens provided sub-optimal image quality with respect to running visual psychophysics. Because of the uniquely localized magnetic field lines associated with the Siemens Verio MRI scanner located at SIBR, the back wall where the projector is located is relatively close to the scanner bore (8'). As such standard projector lens cannot project a focused image onto a screen located at the bore opening. To remedy this situation, we worked with a custom optics company to develop a short-throw lens for our high-quality 3 color DLP projector. This lens has been installed and provides superb image quality consistent with our experimental needs (which focus on studying human vision). Almost all elements are in place and we expect to begin collecting both DSI and fMRI data during the latter half of July, 2010.

In our pilot testing, we have been exploring white matter connectivity between posterior and anterior portions of the human ventral visual pathway (typically assumed to be the pathway mediating visual object recognition). As shown in Figure 1, preliminary DSI data reveals a distinct fiber tract connecting the functionally-defined "fusiform face area" ("FFA" – an area of visual cortex known to be critical in face recognition) and anterior portions of the infereo-temporal cortex ("AIT"). AIT is thought to be critical in face identification and may be the brain region in which "complete" representations of individual faces are encoded.

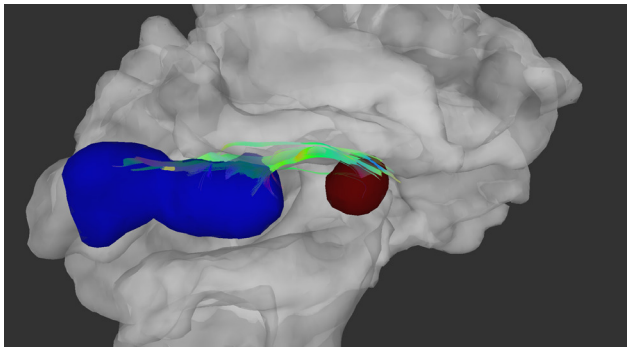

Figure 1. A white matter fiber tract in the human ventral cortex that connects the FFA to OFA (pilot DSI results).

Research Project 2: Project Title and Purpose

Understanding Schizophrenia through Logical and Empirical Analysis - Schizophrenics have persistent experiences of hearing voices (auditory hallucinations) and thoughts that are paradoxically taken to belong to another (inserted thoughts). To explain these symptoms, a common model postulates a mechanism whereby a person's thoughts or actions are "marked" as self-generated. Under this theory, trouble arises when this mechanism fails: a person's thoughts appear "alien" because they are not marked as self-generated. This project suggests an alternative hypothesis: a central problem in patients who hear voices or have alien (inserted) thoughts is not the misattribution of central thoughts, but rather that thinking and imagining are abnormal because they *automatically* generate thoughts and images. This project develops and tests this new hypothesis regarding the origins of these specific symptoms of schizophrenia.

Anticipated Duration of Project

1/1/2010 – 12/31/2012

Project Overview

This project focuses on two symptoms of schizophrenia: auditory verbal hallucinations (AVH) and inserted thoughts. Two of the most common symptoms of schizophrenia are persistent experiences of hearing voices and thoughts that schizophrenics paradoxically take to belong to another individual (i.e., not themselves). Within the psychiatric literature, these symptoms are commonly explained as a dysfunctional "monitoring" mechanism whereby a person's thoughts or actions are usually "marked" as self-generated. Trouble arises for schizophrenic individuals (i.e., they exhibit symptoms associated with schizophrenia) when this putative mechanism fails: a person's thoughts or internal voices appear "alien" because they are not marked as belonging to that person. That is, prevailing theories of AVH invoke notions of controlled processes that involve active monitoring of inner speech. The problem with such models (which dominate the field) is that they fail to answer the right questions about these symptoms. This is due to an inadequate conceptualization of the symptoms themselves; that is, they are taken at face value

rather than considered within the larger framework of cognitive processing. Careful analysis of these symptoms leads to an alternative hypothesis regarding their origin: that a central problem in schizophrenic patients who hear voices or have alien (inserted) thoughts is that *normal* processes of thinking and imagining are abnormal because they *automatically* (and unconsciously) generate thoughts and images (i.e., not that they are mislabeled). This precise and testable hypothesis is reached only by understanding the widely used, but poorly defined, notion of automaticity drawn from cognitive psychology. To investigate the role of automaticity in these two symptoms of schizophrenia this project will involve collaboration with psychologists and psychiatrists at two levels. Conceptually, we shall provide a rigorous, well-specified theoretical account of the notion of automaticity as used in cognitive psychology. Empirically, we will formulate and test more precise alternative models and explanations for some aspects of schizophrenia, specifically investigating mechanisms of AVH, with the aim of exploring both spontaneous cortical oscillatory activity as well as that elicited by external auditory stimuli, with the general hypothesis that increased cortical excitability may mediate the automatic nature of AVH in schizophrenia. Elucidating such mechanisms will further our understanding of one of the core symptoms in schizophrenia and provide the basis for developing novel treatment approaches for the illness.

Principal Investigator

Wayne Wu, PhD
Assistant Professor and Associate Director, CNBC
Carnegie Mellon University
115 Mellon Institute, CNBC
4400 Fifth Avenue
Pittsburgh, PA 15213

Other Participating Researchers

Raymond Cho, MD - employed by University of Pittsburgh Medical Center
Mark Wheeler, PhD - employed by University of Pittsburgh
Lori Holt, PhD - employed by Carnegie Mellon University

Expected Research Outcomes and Benefits

Schizophrenia is a mental disorder affecting approximately 1% of the US population – over 2 million individuals. Although a variety of antipsychotic medications have been developed to treat the disease, these have many side effects and are only partially successful in ameliorating schizophrenic symptoms. Psychosocial treatments are likewise limited in their efficacy. Successful treatment and, ultimately, developing more effective medications and behavioral interventions must rely on more accurate characterizations of the disease and its concomitant symptoms. Thus, a better model of the root causes of two of the most common symptoms that is grounded in cognitive mechanisms will provide a clearer basis for advanced drug and behavioral therapies.

The broader analysis of automaticity and a better model of its role in cognition will also have implications for a wide range of disorders, for example, dyslexia and ADHD, as well as improving our understanding of the symptoms of schizophrenia.

Summary of Research Completed

This research project has not yet commenced because Professor Wu's appointment at Carnegie Mellon has been delayed until August 1, 2010.

Research Project 3: Project Title and Purpose

Computational Immunology for Toleragenic Composite Tissue and Solid-Organ Transplantation
We propose to elucidate the tissue specific biological and informational principles of complex immunological mechanisms and systems that drive rejection and inflammation in solid or composite tissue transplants. Through advanced computational methods, machine learning and other high dimensional analytic techniques, a new understanding of the patterns and governing principles of the immune system will be formulated, and will pave the way for novel clinically relevant toleragenic transplant protocols, therapeutics and long term patient care strategies. The longer term objective is to improve transplant tolerance and reduce side effects from indiscriminant immunosuppression.

Anticipated Duration of Project

1/1/2010 – 6/30/2012

Project Overview

This project seeks to elucidate immunological patterns and mechanisms in pre-clinical composite tissue and solid organ transplants under various conditions of rejection, inflammation and tolerance. To achieve this, machine learning, computational linguistics, agent based modeling, and differential equation based methods of analysis will be applied to proteomic and genetic data. The raw data is being gathered through ongoing collaborative projects with the University of Pittsburgh Departments of Immunology, Plastic and Reconstructive Surgery, The McGowan Institute for Regenerative Medicine, The Pittsburgh Nuclear Magnetic Resonance Center, and the University of Innsbruck School of Medicine.

The objective of the proteomic analysis is a set of models with predictive power for rejection state (and possibly grade), time point, and graft type. Models will be built for specific tissues including skin, muscle, heart, lung, and other tissues for which data may be available including the liver, kidney, pancreas, lymph node, and serum. Protein data is provided in concentration values (pg/ml) as read by a Luminex 100-IS machine at the University of Pittsburgh Department of Immunology.

Genetic analysis will focus on the skin and muscle components, utilizing gene array data provided by the composite tissue transplant program at the University of Pittsburgh Department of Plastic and Reconstructive Surgery as well as the University of Innsbruck School of Medicine.

The objectives of this analysis include development of models with predictive power for rejection state, identification of rejection specific gene expression and enumeration of these targets for verification by collaborating researchers, and corroboration of gene-protein expression under various rejection and inflammation conditions. Time and resources permitting, we will also investigate the feasibility of models with predictive power for time point and physiological process.

Principal Investigator

Jaime Carbonell, PhD
Director, Language Technology Institute
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213

Other Participating Researchers

Ravi Starzl – employed by Carnegie Mellon University

Expected Research Outcomes and Benefits

Both solid organ and composite tissue transplantation has traditionally required extensive life-long immunosuppression therapy to prevent rejection of the graft. It is this long-term high-dose immunosuppression regimen that is the limiting factor for both the quality of organs that can be used for a solid organ transplant and for the widespread adoption of composite tissue transplantation. Further, the present clinical method for detecting rejection in solid organ transplant is fluctuation of organ function (caused by damage from rejection), or skin discoloration in the case of composite tissue rejection.

The recent paradigm of transplant tolerance, or chimerism, has revealed a conceptual framework where it may be possible to detect the onset of rejection before tissue damage occurs, adjust immunosuppression dosage to optimal levels, and potentially wean patients to very low or no maintenance doses of immunosuppression.

This study will generate novel quantitative models, methods and protocols that will allow the toleragenic paradigm to be personalized and adapted to each individual patient. This would allow better allograft management and the treatment of rejection prior to the accumulation of damage that affects organ function. With better proactive management, a larger pool of organs that could be used in transplant would be made available, helping to alleviate the worldwide shortage of kidney, liver, pancreas and other organs. Additionally, this study will enumerate genetic markers of susceptibility to rejection, as well as targets and strategies for the development of novel non-toxic immunoregulatory pharmaceuticals.

Summary of Research Completed

Investigation into the cytokine network profiles in both solid organ and composite tissue

transplant has begun. To elucidate the cytokine dynamics in solid organs, data from rat cardiac (heart and lung) transplant surgeries conducted at the Nuclear Magnetic Resonance (NMR) center of Carnegie Mellon University has been collected. To elucidate cytokine dynamics in composite tissue transplant, data has also been collected from rat hind limb transplant surgeries conducted at the Department of Plastic and Reconstructive Surgery - Composite Tissue Allotransplantation Group, of the University of Pittsburgh. Preliminary results from the analysis of these data sets is promising and affirms our expectations to deliver the project results proposed in this work on schedule.

Materials and Methods

Rat Cardiac Transplantation Surgery and Sampling

Using a Brown-Norway (BN) to Agouti (DA) heterotopic heart and lung transplant technique described in Ho 2001, allogeneic [n=13] and syngeneic [n=4] cardiac transplants were studied, as well as naïve [n=2] rats. The heterotopic transplant technique allows the native heart and lungs to remain functional in the animal, allowing samples to be drawn from both native organs as well as allografts. At every time point native heart, graft heart, native lung, graft lung, lymphnode, spleen, kidney, liver and serum samples were taken.

Rat Hind-Limb Composite Tissue Allotransplantation Surgery and Sampling

Using a Brown-Norway (BN) to Lewis (LEW) rat hind-limb allotransplant technique, syngeneic [n=10], allogeneic transplants without immunosuppression [n=10], allogeneic grafts [n=10] treated with tacrolimus and naïve rats [n=10] were studied. 180 skin and muscle biopsies were taken between postoperative days 0 and 11. Protein levels of 14 cytokines known to be relevant in cellular inflammatory responses were assessed by Luminex™.

All tissue samples are 4mm squares preserved in RNALater and frozen at -25C for 24 hours, then moved to -80C. Blood samples are processed into Serum, then frozen at -25C for 24 hours, then moved to -80C.

Luminex Cytokine Concentration Analysis

Samples were homogenized and protein levels of 14 cytokines known to be relevant in cellular inflammatory responses (GMCSF, Il-1α, Il-1β, Il-5, Il-4, Il-6, Il-8, Il-10, Il-18 TNF-α, IFN-γ, Il-12 p70, GRO/KC and MCP-1) were assessed by Luminex® 100 IS using the Rat Cytokine/Chemokine Luminex® Bead immunoassay Kit, LINCOplex, 14 Plex. Each sample was bisected and run in two separate arrays of wells. The final protein concentration values were calculated as the mean of the values from the two arrays.

Computational Tools

All analysis utilized the Weka machine learning toolkit (v 3.6) and Microsoft Excel (v 2008), Luminex® protein readings were provided in a comma separated value (CSV) format that were then converted to attribute-relation file format (ARFF) for further analysis in Weka Explorer. Normalization of data was performed by manually programmed min-max normalization formula in MS Excel, with the form $X^* = (X - \min(X))/\text{range}(X)$. Linear classification analysis utilized the Naive Bayes classifier tool within Weka, while decision tree analysis utilized the J48 algorithm. Feature extraction was aided by the information gain attribute evaluator (ranker search algorithm), as well as the correlation-based feature subset evaluator (best first algorithm).

Decision trees, entropy values and classification results were generated with Weka visualizer and explorer, while graphs were generated with MS Excel.

Computational Processing, Analysis and Results

Cardiac Transplant:

Allograft/Isograft groups and Native tissue can be clearly distinguished from each other at all time points with substantially different cytokines levels at all time points. Allograft and Isograft groups are also linearly separable from each other as early as POD 1, but with a narrow margin.

Referring to Figure 3, IL-18 is the most reliable differentiator between Native tissue and Isograft/Allograft tissue. Allograft IL-18 levels at POD 1 are higher than Native tissue by a factor of approximately 7, and Isograft IL-18 levels at POD 1 are higher than Native tissue by a factor of approximately 5. The J48 decision tree algorithm (Weka) was able to create a classification rule that distinguished Native heart tissue from Allograft heart tissue with 100% success at all time points in this preliminary data set. The J48 decision tree algorithm was also able to derive a rule that distinguished allograft from native lung tissue 100% accuracy in this preliminary analysis based on elevated IL-1a levels in the allograft.

Composite Tissue Allotransplant:

As a preprocessing step, protein levels for postoperative day 3 are standard normalized (mean 0, standard deviation 1) for each of the cytokines. Feature selection is then applied with the least absolute selection and shrinkage operator (LASSO). This method applies an L1 constraint while minimizing residual sum of squares driving some of the predictors exactly to 0, eliminating them from the model. Picking the following cytokines with the highest weight thus reduced the number of relevant features to three: TNF-a, Il12 p70 and Il-5.

Using normalized protein levels of the chosen three cytokines, multiclass logistic regression is applied to distinguish between Allograft, Isograft and tacrolimus treated samples. Logistic regression is a linear regression technique which applies logit function to compute posterior probabilities of classes given the data. Under this model, probability of a class j with the exception of the last class is for sample X_i is:

$P(C_j \mid X_i) = \exp(\beta_j^T X_i) / \Sigma_k \exp(\beta_k^T X_i)$ k = number of classes

where parameters β_i are estimated using an efficient algorithm called iterative reweighted least squares. We applied leave one out cross validation (LOOCV) to test the classifier.

When multiclass logistic regression is applied without feature selection, 4 out of 14 samples are misclassified leading to 71.4% accuracy, indicating that multiple cellular or physiological processes generate the cytokine concentrations observed. The final concentration value of the cytokines/chemokines likely represents overlapping distributions from these multiple processes. However, when we reduce the feature space to be more tractable through LASSO feature selection (yielding TNF-a, Il12 p70 and Il-5), and then apply a logistic regression classifier (utilizing 10-fold cross validation), all samples were successfully classified to their correct groups at post operative day 3, obtaining %100 accuracy. Lasso reduced noise from the model, revealed cytokines that are discriminative for rejection, inflammation and increased the learning performance.

Discussion - Cardiac Transplant
While initial results are very promising, additional analysis is needed to establish how robust to noise differentiation between Allograft and Isograft is, particularly at the earliest time points. The difference between Allograft and Isograft groups becomes quite clear and progressively more pronounced at POD 3, 5, 7 and 9. Promising techniques developed during the analysis of CTA data will be adapted and applied to the cardiac transplant model.

Discussion – Composite Tissue Allotransplant
Initial results of CTA data are also very promising and already suggest some methods for building an in-silico model of composite tissue allografts during rejection. While diagnostic accuracy is very high in these preliminary results, testing for robustness against noise is again very important.

Conclusion
Initial results are promising and suggest an important link exists between cytokine network dynamics and the state (as well as progression) of rejection in solid and composite tissue transplantation. The cytokine dynamics indicative of rejection appear to be linearly separable at early time points and appear to be predictive of acute rejection. Further, the cytokine network profiles observed suggest that there is significant attribute interaction in the underlying immunological processes that drive rejection, and that interrupting these processes, either at a cytokine or cellular level, may be likely to alter the course of rejection or rejection pathology.

In both the composite tissue allotransplantation and cardiac transplantation models making additional connections to underlying mechanisms through corroborative findings in the literature, or direct observation, will further strengthen these findings, elucidate the mechanisms of rejection, and help direct the clinical application of these findings.

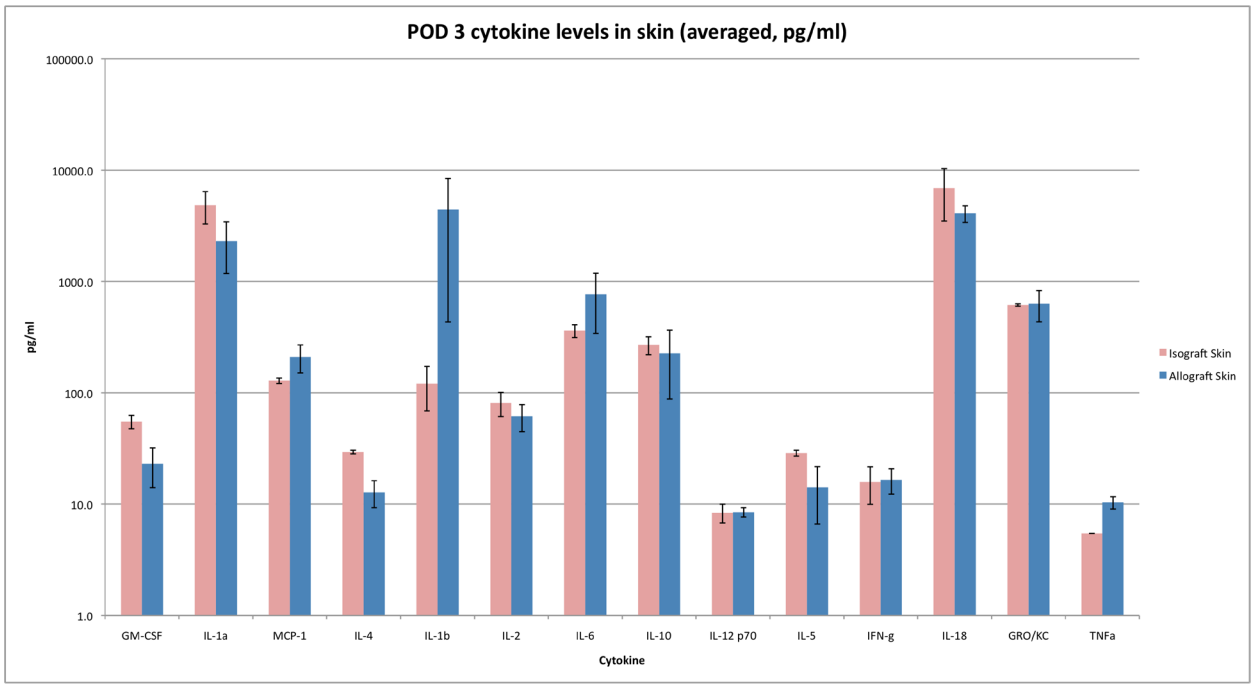

Figure 1: Cytokine levels in skin of composite tissue allografts on postoperative day 3.

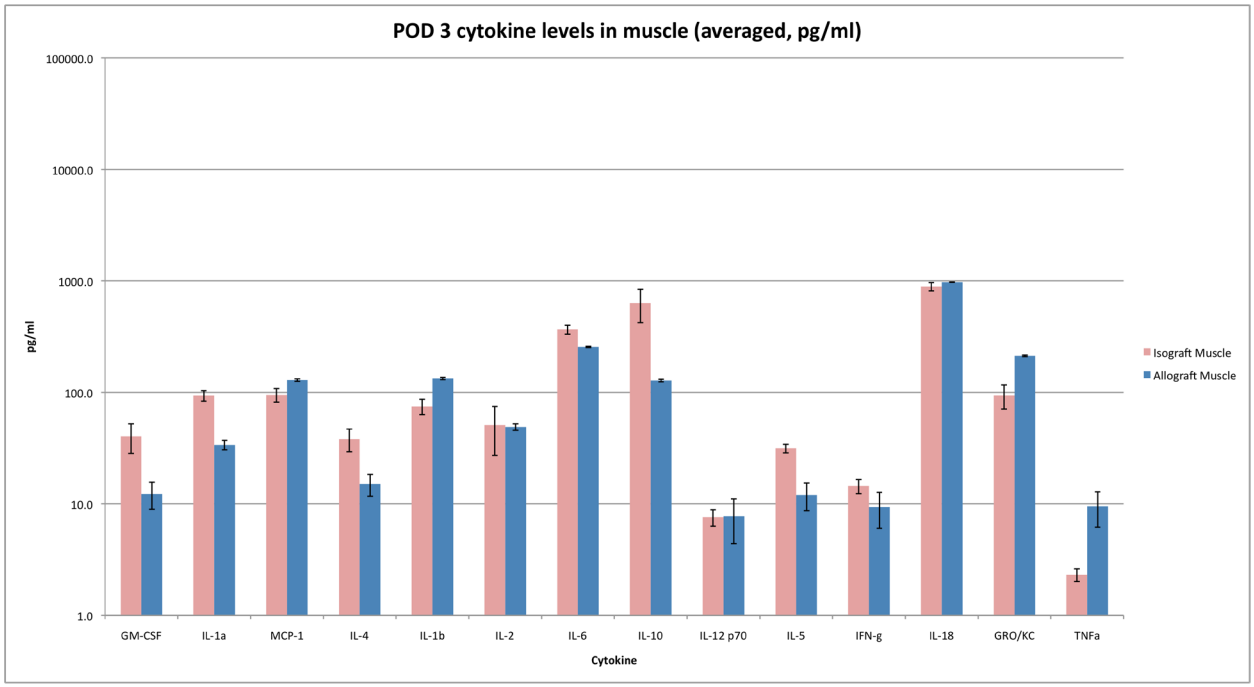

Figure 2: Cytokine levels in muscle of composite tissue allografts on postoperative day 3.

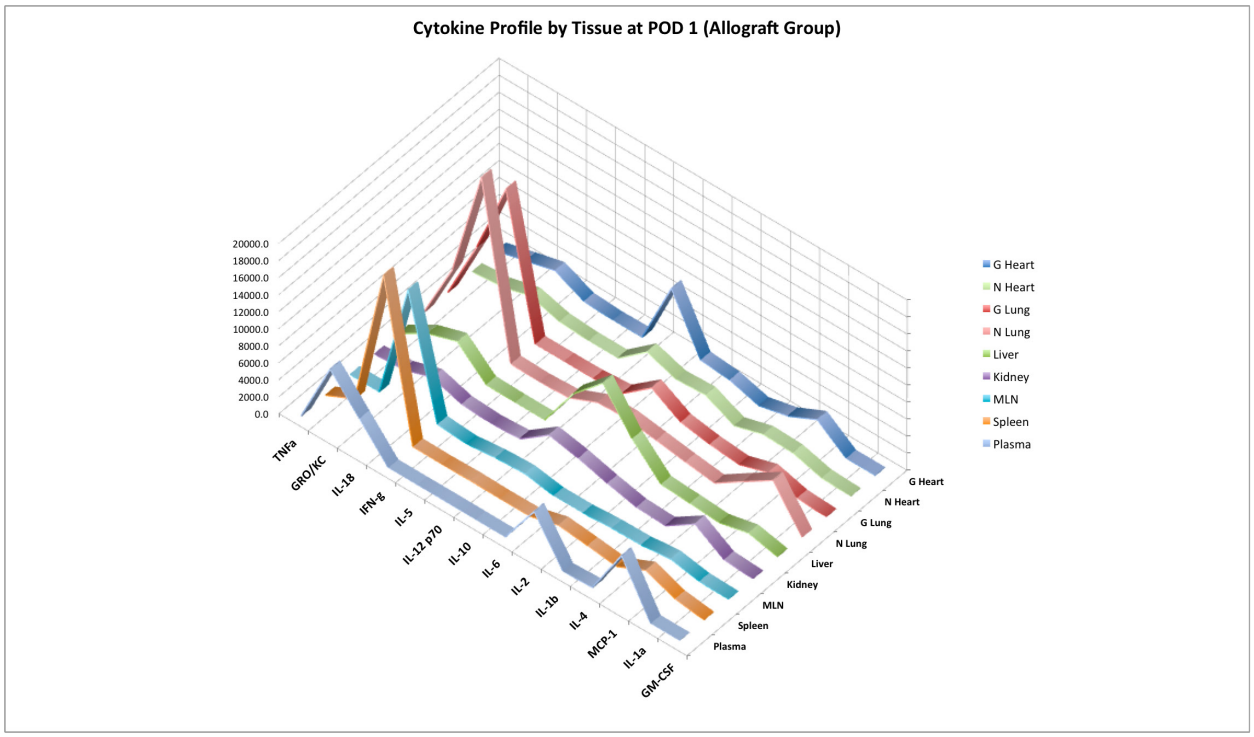

Figure 3: Cytokine levels in all sampled tissues of allograft cardiac transplant on postoperative day 1.