

Systems Immunology Dynamics Linking Biomedical Data Science to Immune Disease

Sadia Tasneem

May 25, 2025

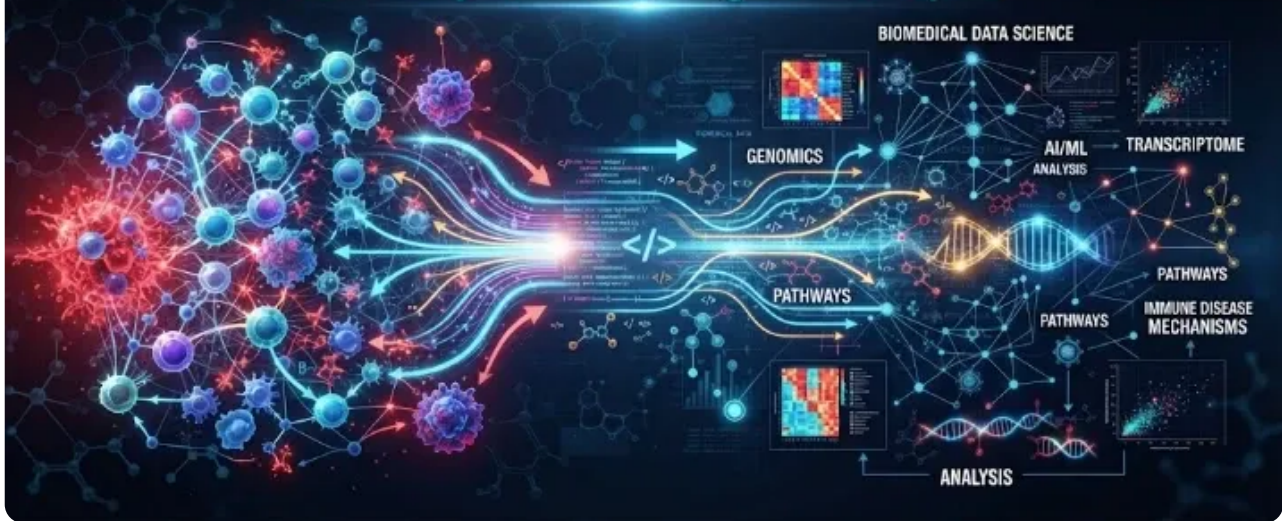
Abstract

The article examines immune disease as changing cellular states that move through flare, treatment response, remission, relapse, and infection. Longitudinal single-cell studies show that epithelial, myeloid, and T-cell compartments can follow different trajectories under therapy, making static pretreatment measurements insufficient. Systems-immunology models can identify coordinated modules that rise or fall together, while receptor sequencing can track persistent or replaced lymphocyte clones. Sampling frequency must match biological timescales, from rapid cytokine signaling to slower clonal replacement and tissue remodeling. State-space, differential-equation, and machine-learning models can describe these transitions, but mechanistic predictions require independent experimental validation. Age, sex, genetics, microbiome, disease duration, and treatment history should be modeled without averaging away meaningful subgroups. Biomedical data science is most useful when it predicts the direction of an immune trajectory and reduces the space of plausible mechanisms or interventions rather than merely describing high-dimensional change.

SYSTEMS IMMUNOLOGY DYNAMICS

Linking Biomedical Data Science to Immune Disease

From Complex Data to Targeted Therapies



Although clinical labels are the same, immune disease is dynamic. For someone with the same diagnosis, the process of reorganizing the underlying immune system can include flare, partial response, remission, relapse, infection, and medication change. Those transitions can be made measurable through the combination of single cell profiles, soluble proteins, receptor repertoires, and clinical data across time, in the context of Biomedical data science. The big challenge is not finding the differences but determining which changes are evidence of immune dynamics rather than technical variability or downstream effects.

This approach is seen to be useful in longitudinal single cell studies. Repeated sampling of intestinal tissue during anti inflammatory treatment has demonstrated differential trajectory of remission and nonremission of epithelial, myeloid and T cell compartments in IBD. It is impossible to obtain this kind of result completely from a single pretreatment biopsy because therapeutic outcome relies on the response of cellular networks following therapy.

From the perspective of systems immunology, those trajectories are considered as “coordinated state changes.” A decrease in one inflammatory population can have an effect on the availability of cytokines, antigen presentation and recruitment of other cells. Network analysis is able to detect modules that go up and down together. A dynamic module might be more stable (stable between different patients) than one specific cell subtype, since different patients may be in similar inflammatory states by similar cellular

pathways.

What can be learned depends on the time resolution. Only sampling before treatment and months later will not capture early signaling events that may help determine response. Early and late samples can provide tissue remodeling and pharmacodynamic changes, respectively, if testing around treatment initiation. If testing around treatment initiation, early and late sampling can measure pharmacodynamic changes and tissue remodeling, respectively. The best schedule is determined by the biological process as cytokine signaling can alter in minutes, and clonal replacement can take weeks.

The various timescales can be represented in computational models. Transitions among cell populations or immune states can be estimated using ordinary differential equation models, state space models, and machine learning approaches. Models can be constrained more effectively by data that is observed in the cells over multiple time points (longitudinal data) as opposed to data that is only observed at a single time point (static measurements). The resulting parameters can then lead to making mechanistic hypotheses regarding proliferation, death, migration or differentiation.

Those hypotheses should be directly tested with validation. Independent assays of neutrophil turnover or tissue abundance could assess the prediction made by a model if it suggests that neutrophils are a key player in the onset of remission. A computational explanation is useful if it suggests an experiment that is realizable but might fail. If not, the modelling will start to be merely a description of the data.

Biological memory is added to the analysis using clonal tracking. The T cell receptor and B cell receptor sequences recognize similar cells at different times. Experimentally, it is possible to determine if remission is due to elimination of pathogenic clones, functional reprogramming of persisting clones, or replacement by new populations. Two patients with no more than identical cell counts can thus exhibit very different clonal dynamics.

Spatial profiling can determine if these clones use pathogenic niches. Blood sampling does not reflect tissue aggregates, epithelial interfaces, granulomas or inflamed vessels. Association of spatial cell states with circulating markers can be used to determine which blood markers are systemic responses rather than association with tissue processes.

The same can be achieved with biomedical data science, which can detect early warning signs. Before clinical relapse occurs there can be a gradual drift of a cell state or increased variance in that state, or both. These markers should be tested for in the future as retrospective information can create a more homogenous appearance than is present in

pre-relapse patterns. To assess the ability to predict, future samples are to be used with rules that are set prior to the occurrence of results.

Longitudinal studies are especially prone to batch effects. Technical drift can simulate biological change if all early samples are processed in a first laboratory versus later samples in a second laboratory. Randomization of samples, repetition of standards, reference controls and Batch correction are essential. Not all data collected in a confounded design can be corrected using computations.

The act of treatment presents an additional interpretation challenge. A drug may directly affect the number of cells and indirectly the inflammation of the tissue. Therefore, a decreasing immune signature may not represent disease clearance, but rather, target engagement. Combination of pharmacology, clinical response and mechanistic assays can yield a distinction between the expected drug effect and a biomarker for durable remission.

Don't average the population out of the data. Immune states are influenced by age, sex, genetics, microbiome, previous infection, and disease duration. Hierarchical Models can track common patterns and individual paths. The key to precision medicine is to know what is common to all patients and what is predictable in subgroups.

At the end of the day, a clinically useful dynamic signature must be simple. Discovery can be thousands of features, but routine care is not likely to support a large multiomic panel. A complex model can be reduced, or a gene expression assay and/or flow cytometry can be developed, then tested for preservation. Translation, when it works, is the production of a measurement that is useful outside the research laboratory, not an analysis.

The dynamics of state transitions can also be modeled as trajectories in low dimensional immune space. The dimensionality reduction can be used to co-ordinate the thousands of features, and pseudo-time can be used to sequence the cell states along developmental pathways. In longitudinal clinical data, time should always be the base, as pseudotime may imply an order that is different from the real temporal order.

Relapse biology may be different than primary disease. One pathogenic population can be removed by a treatment and another one established later. This shift can be detected by comparing relapse to scores obtained from the sample of healthy controls and also with scores obtained prior to the relapse. Dynamic models, therefore, should have the potential of inducing new states to occur rather than being constrained to assigning the later samples to any of the defined states at enrollment.

The use of cross tissue integration is very informative. Repeat sampling is easy for blood, but may be difficult for tissue. A model can find a set of blood characteristics that can track a tissue program, and confirm this set of feature against matched biopsies. This provides a meaningful link from deep tissue biology to less invasive monitoring.

Experimental design can be aided by mechanistic simulations. Simulated perturbation can approximate which time point in a model has the greatest effect if one population of cells is driving the other by means of a cytokine circuit. Thereafter, laboratory experiments can be taken around that interval. Computation is most useful when it cuts down on the space of plausible experiments instead of replacing them.

Clinical events can be used as interventions, not background noise in the trajectory model. Sudden changes in immune systems can result from infection, steroid bursts, surgery or vaccination. These events are annotated so the model can differentiate between disease driven change and external perturbations.

A useful relapse model ought to also offer lead time. The value of prediction one day in advance is different from prediction several weeks in advance. Lead time distributions and false alert rates can give an indicator of the feasibility of a dynamic immune signature for practical monitoring.

Uncertainty in cell annotation can also be included in dynamic immune models. Not all states are easily assignable to a single category, and putting each cell into a specific category can result in time jumps. Biological gradients can be preserved using a probabilistic annotation or continuous state scores. This can be helpful if therapy is a gradual exchange of one discrete population with another.

A relapse trajectory should also be correlated with simpler clinical markers. Any complexity added to the measurement is warranted only to add value to lead time, accuracy, or actionability that goes beyond routine measurement.

The immune disease research landscape is centered, and this is what Nima Aghaeepour's biomedical research interests are about, biomedical data science and systems immunology. Longitudinal measurements can show the changes in immune networks when treated and computational models can predict transitions associated with response or relapse. The best models are still based on cell biology, grow independently and do not result in more complex validated assays. The information gained from systems immunology is most powerful when it helps to describe the immune state that exists, its dynamics, and the direction it is likely to go.