

City Research Online

City St George's, University of London

Citation: Bukkuri, A. (2026). The mask of sanity: Cancer and autoimmunity as complementary failures of semiotic classification. Mathematical Biosciences and Engineering, 23(10), pp. 3169-3183. doi: 10.3934/mbe.2026109

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/38563/

Link to published version: https://doi.org/10.3934/mbe.2026109

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 23(10): 3169–3183.
DOI: 10.3934/mbe.2026109
Received: 23 June 2026
Revised: 02 October 2026
Accepted: 15 October 2026
Published: 27 October 2026

Letter

The mask of sanity: Cancer and autoimmunity as complementary failures of semiotic classification

Anuraag Bukkuri*

Department of Mathematics, City St George's, University of London, London, UK

* **Correspondence:** Email: anuraag.bukkuri@city.ac.uk.

Abstract: The immune system encounters innumerable cells every day and must decide whether these cells are healthy and thus must be tolerated, or are malignant and thus must be eliminated. These decisions are made from imperfect molecular signals in a complex, noisy environment. Errors in classification may lead to diseases such as cancer and autoimmunity. In this paper, we use ideas from cancer biosemiotics and signal detection theory to formalize this notion and show that cancer and autoimmunity are two sides of the same semiotic corruption problem, rather than separate pathologies. We show how biological factors such as noise, aging, deceptive signaling, and chronic inflammation all converge on reducing the discriminability of the immune system. Using this framework, we make two clinical predictions that are consistent with results from randomized clinical trials: 1) checkpoint inhibitors increase the aggressiveness of the immune classifier, leading to improved tumor control at the cost of higher rates of immune-related adverse events, and 2) the adverse side effects associated with checkpoint inhibitors will be more systemic than those associated with antigen-directed therapies. Then, we allow the immune system's reference for normalcy to update over time, and show that the lag between the true tissue state and the immune system's perception of it leads to a transient shift in the aggressiveness/tolerance of the classifier. In accordance with the discontinuity theory of immunity, we show that abrupt changes are registered as abnormal, whereas gradual changes are absorbed into the immune system's notion of normalcy. Although we frame this discussion in the context of cancer, it applies more broadly to any biological classification problem in which the state of the sender is unknown, and signals are imperfect and corruptible.

Keywords: biosemiotics; semiotic corruption; cancer; autoimmunity; signal detection theory; discontinuity theory

1. Introduction

The cells in our immune system are continually patrolling our bodies, on the lookout for "the other". During this process, they encounter a number of cells, some malignant and some healthy. The challenge for the immune system is determining which is which, based on imperfect signals in a complex, noisy environment. They cannot directly observe the state of the cells they encounter, but must detect and interpret cellular malignancy or normalcy via signals such as neoantigens presented on MHC-1 molecules and stress responses (e.g., NKG2D ligands). Their interpretation impacts their response, either towards tolerance or attack, thereby influencing downstream physiological consequences.

When signaling integrity is intact and normal and malignant cells occupy distinct regions of a high-dimensional signal space, the classification problem is straightforward, and the immune system can effectively prevent or combat disease. However, when the semiotic landscape is corrupted, the immune system may interpret normal cells as dangerous (type 1 error leading to autoimmunity) or malignant cells as normal (type 2 error leading to cancer). For example, this may occur when the signal space dimensionality collapses due to aging, the signal overlap increases due to deceptive signaling or chronic stress, or there is a high degree of noise in the environment. In this paper, we frame cancer-autoimmune classification as a biosemiotic problem and draw on signal detection theory to show that fundamental constraints on immune classification force trade-offs between minimizing type 1 and type 2 errors. This trade-off between autoimmunity and cancer tolerance is observed in clinical settings. For instance, in patients with melanoma who were treated with immunotherapy, the induction of vitiligo (an autoimmune response in which the immune system attacks melanocytes) was associated with improved survival [1]. Conversely, patients who received solid organ transplants, and consequently experienced immunosuppression, were shown to have significantly elevated cancer risks across 32 distinct malignancies [2]. We demonstrate how the biological factors described above influence immune classification error rates and suggest that immune checkpoint inhibitor therapies may induce autoimmunity as a side effect of intentionally biasing the immune system towards an aggressive, type 2 error-minimizing regime [3, 4]. Then, we investigate how immune phenomenology dynamically changes, showing how a classifier with an evolving notion of "normalcy" captures Pradeu's discontinuity theory [5, 6]: slow, gradual changes can be absorbed into the immune system's conception of "normal", whereas abrupt changes are classified as "abnormal". Specifically, we show that the lag between the tissue's true state and the immune system's perception of it shifts the classifier towards a more aggressive (tolerant) strategy when stress rises (falls) faster than the immune system habituates to it.

2. Biosemiotic grounding

Before presenting our mathematical formalization, we ground the cancer-autoimmune classification problem in biosemiotics. There is no better place to start than Jakob von Uexküll's conception of the *Umwelt* [7]. Building on Kant's distinction between the noumena and the phenomena, von Uexküll stressed the importance of recognizing that individuals do not perceive the world as it is, but through a perceptual experience that is unique to them. They do not perceive objects directly, but via an indirect representation of them that they interpret, derive meaning from, and act upon.

This is particularly striking in the context of immune cells whose job, broadly, is to identify and eliminate "the other". However, immune cells do not directly observe the "true state" of cells that they interact with (e.g., normal or malignant). Rather, they encounter cell surface receptors, signaling molecules, and metabolic outputs within a noisy, complex environment and must extract actionable meaning within this milieu. Under physiological conditions, this classification task is straightforward: the immune system is able to effectively detect, interpret, and eliminate malignant cells while tolerating normal cells. However, under diseased conditions, this classification may become more difficult*.

To understand why this is the case, we turn to Peirce's triadic model of signs, which formalizes the communication process described above [9]†. Under this model, meaning is manifested through a dynamic three-part relationship that involves an object (the referent: normal or cancer cell), a representamen (the sign's form: cell surface markers), and an interpretant (the sense made of the sign: immune cell interpretation machinery). Peirce, and later Eco, additionally distinguished between the dynamical object, the real extra-semiotic entity, and the immediate object, the representation of that entity within the sign relation [9, 12]. In this paper, this distinction is critical: immune cells do not have access to the dynamical object, and can only act upon the immediate object–it is this phenomenology, both how it arises and its implications for pathology, that we investigate here.

There are two aspects of this triad that pathology corrupts, thereby driving a divergence between the dynamical and immediate object: the representamen and the interpretant. Specifically: 1) cancer cells can deceptively signal their state to appear like normal cells to the immune system, and 2) immunosenescence, signal noise, and chronic disease can interfere with the way immune cells interpret the signals they receive. Figure 1 visually illustrates Peirce's model of semiotics, adapted for our cancer-autoimmune classification problem. The mathematical framework presented in the next section will help operationalize this conceptual model and allow us to explore how corruption of the semiotic landscape may lead to cancer or autoimmunity.

3. Mathematical framework

Based on this biosemiotic grounding, we introduce the mathematical model we will use to explore the cancer-autoimmune classification problem. This framework is grounded in signal detection theory, a field that develops techniques to measure differences between patterns in the presence of noise. For simplicity, we assume that there are two cell types that the immune system interacts with: θ_N and θ_C, representing normal and cancer cells, respectively. The immune system cannot directly observe these types, but must instead infer the cell type from an observable signal $m \in \mathbb{R}^d$ drawn from the following Gaussian distributions:

$$m|\theta_N \sim \mathcal{N}(\mu_N, \sigma^2 I_d),$$
$$m|\theta_C \sim \mathcal{N}(\mu_C, \sigma^2 I_d).$$

*Our use of the term "classification" draws on a long intellectual history that reaches back to medieval scholastic philosophy, a tradition which, in part, aimed at classifying natural phenomena from their forms and essences. This logic shaped early inquiries into biological systematics [8] and inspired Peirce's triadic model of the signs. Here, we adopt this approach to understand cellular phenomenology.

†The cybernetics school of semiotics approaches this problem in a complementary manner, by analyzing types of information transmission via code to understand how information transforms into meaning across physical, biological, and social systems [10]. Sharov, in particular, has attempted to link this tradition with biosemiotics via his notion of functional information [11].

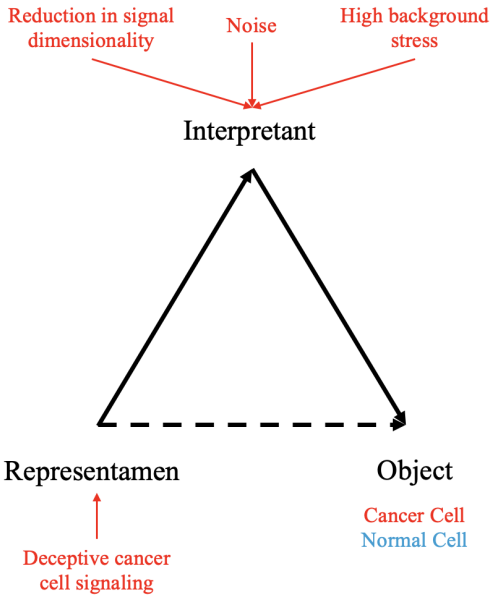


Figure 1. Peirce's triadic model adapted to cancer-autoimmune classification. Reduction in signal dimensionality, noise, and high background stress corrupt the immune interpretant, thus driving a divergence between the dynamical and immediate object. Deceptive cancer cell signaling corrupts the representamen.

In the language of Peircean semiotics, θ_i and m represent the object and the representamen, respectively. Without loss of generality, we set $\mu_N = \mathbf{0}$ and $\mu_C = \delta\mathbf{1}$. Under this formulation, the number of dimensions d can be understood as the number of distinct signals that the immune system uses to classify cells as normal vs. cancer. Similarly, δ can be seen as the typical fold-change in signal markers between normal and cancer cells, and σ is the perceived level of noise and stochastic variation in the signals.

In a given local environment, let $P(\theta_C)$ be the prior cancer probability. Then, via Bayes' theorem, the posterior probability is as follows:

$$P(\theta_C|m) = \frac{P(m|\theta_C)P(\theta_C)}{P(m|\theta_C)P(\theta_C) + P(m|\theta_N)(1 - P(\theta_C))}. \tag{3.1}$$

Given this probability, the immune system must determine whether to attack or tolerate the cell. To incorporate the costs of misclassification, we introduce parameters λ_1 and λ_2, which capture the cost of type 1 (attacking normal cells) and type 2 (tolerating cancer cells) errors, respectively. Then, basic algebra gives us a threshold $\tau^* = \frac{\lambda_1}{\lambda_1+\lambda_2}$ such that if $P(\theta_C|m) > \tau^*$, then the immune system will attack the cell and will tolerate it otherwise. Note that if autoimmunity is more costly in the local environment ($\lambda_1 > \lambda_2$), then this threshold will increase (such that stronger evidence for malignancy is required for attack), corresponding to a tolerant immune state. Conversely, if cancer is more costly ($\lambda_1 < \lambda_2$), then the threshold will decrease (such that weaker evidence for malignancy is required for attack), corresponding to an aggressive immune state. In Peircean semiotics, this internal processing represents the interpretant (one may make a finer distinction here: the dynamic interpretant is the classification into normal vs. cancer, and the energetic interpretant is the decision to tolerate vs. attack).

Next, we introduce the discriminability index from signal detection theory, which captures the

separation between the means of two distributions. For the multivariate Gaussians with isotropic covariance $\Sigma = \sigma^2 I_d$ that our model involves, the Mahalanobis distance between means is as follows:

$$D_M = \sqrt{(\mu_C - \mu_N)^T \Sigma^{-1} (\mu_C - \mu_N)} = \frac{\|\mu_C - \mu_N\|}{\sigma}.$$

With $\mu_N = \mathbf{0}$ and $\mu_C = \delta \mathbf{1}$, we have $\|\mu_C - \mu_N\| = \delta \sqrt{\mathbf{1}^T \mathbf{1}} = \delta \sqrt{d}$. Thus, the discriminability index is given by the following:

$$\mathcal{D}_0 = \frac{\delta \sqrt{d}}{\sigma}. \tag{3.2}$$

As one would expect, the greater the difference in signal markers between normal and cancer cells, the higher the dimensionality of the signal space, and the lower the signal noise, the greater the discriminability index will be. We can analytically compute the log-likelihood ratio as follows:

$$\Lambda(m) = \ln \frac{P(m|\theta_C)}{P(m|\theta_N)} = \frac{(\mu_C - \mu_N)^T \left(m - \frac{1}{2}(\mu_C + \mu_N)\right)}{\sigma^2}, \tag{3.3}$$

and write the decision threshold, $P(\theta_C|m) > \tau^*$, equivalently as $\Lambda(m) > c$, where

$$c = \ln \frac{\lambda_1 (1 - P(\theta_C))}{\lambda_2 P(\theta_C)}. \tag{3.4}$$

Note that the error costs and prior only enter the decision rule through c, meaning that a tissue with a higher cost ratio is indistinguishable from one with a low prior cancer probability. Thus, $\lambda_1, \lambda_2, P(\theta_C)$ are unidentifiable from just the behavior of the immune classifier. We can additionally compute the type 1 (α) and type 2 (β) error rates as follows:

$$\alpha = \Phi\left(-\frac{c}{\mathcal{D}} - \frac{\mathcal{D}}{2}\right) \text{ and } \beta = \Phi\left(\frac{c}{\mathcal{D}} - \frac{\mathcal{D}}{2}\right), \tag{3.5}$$

where Φ is the standard normal cumulative distribution function. Table 1 lists key model parameters, their biological interpretation, and the values used in the following simulations.

Table 1. Model parameters.

Parameter	Interpretation	Values
d	Signal space dimensionality	Varies
δ	Signal separation	0.6
σ	Noise level	0.3
s	Stress intensity	Varies
$P(\theta_C)$	Cancer prevalence	0.5
λ_1/λ_2	Cost ratio	1

4. Results

4.1. Static immune classifier

To build an intuition for our framework, consider a classification problem for $d = 2$. Although the immune system operates in a higher-dimensional signal space, this simplified setting captures the

key relevant features. First, we define the normal and cancer distributions as $\mathcal{N}((0,0), 0.3^2 I_2)$ and $\mathcal{N}((0.6,0.6), 0.3^2 I_2)$, respectively. In Figure 2, we sample 2000 signals from each, with green points representing normal cell signals and red points representing cancer cell signals. The outlined circle and square represent the mean signal for the normal and cancer distributions, respectively. The dashed line represents the Bayes-optimal boundary for classifying cells as normal vs. cancer for $\lambda_1/\lambda_2 = 1$, given as the perpendicular bisector of the line connecting signal distribution means. Unequal costs for type 1 and type 2 errors shift this boundary (dotted lines), towards the cancer mean when autoimmunity is more costly and towards the normal mean when cancer is more costly. Although the normal and cancer cells clearly occupy distinct regions of signal space, there is an overlapping region, an inevitability due to the probabilistic nature of signal expression. Thus, the signals expressed by some cancer cells will be interpreted as normal (type 2 error) and vice versa (type 1 error). When signaling integrity is intact, the overlap in signal distributions will be minimal, as will misclassification events.

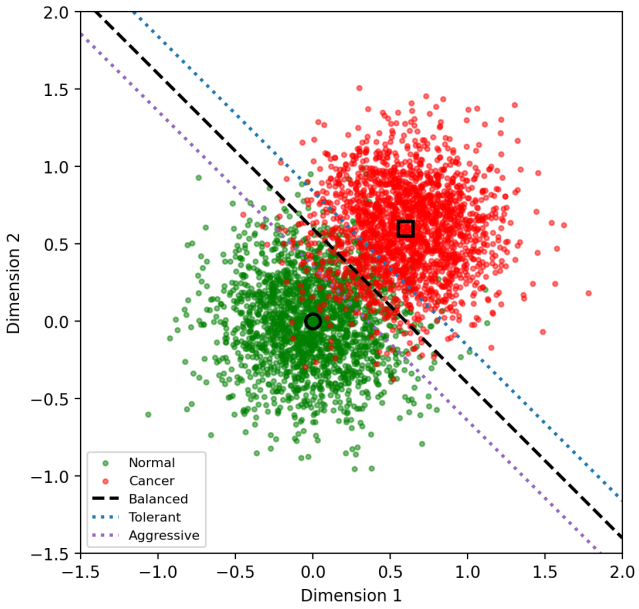


Figure 2. Visualization of overlapping distributions in two-dimensional signal space. 2D signal space showing normal (green) and cancer (red) distributions with three immune classifiers. The dashed line corresponds to the balanced total error minimizer ($\lambda_1/\lambda_2 = 1$), and the dotted lines correspond to the Bayes-optimal boundaries for aggressive ($\lambda_1/\lambda_2 = 1/5$) and tolerant classifiers ($\lambda_1/\lambda_2 = 5$). Overlapping distributions mean that misclassification will occur, regardless of where the classification boundary is set.

However, semiotic corruption can significantly increase classification errors and lead to autoimmunity or cancer. To see how this may occur, consider how the discriminability index can degrade. We propose four main biological mechanisms: degradation of signal dimensionality, noise, chronic stress, and deceptive cancer cell signaling. First, consider the erosion of signal dimensionality, as occurs during aging (e.g., via thymic involution). Early in life, individuals have a highly diverse T cell repertoire that allows them to recognize a wide range of antigens. However, the thymus, which is responsible for T cell repertoire formation [13] via V(D)J recombination, progressively involutes with age, dramatically contracting the T cell repertoire [14–16]. This reduces

the number of molecular features the immune system can assess when encountering a cell in the body, which we formalize as a decrease in d. This is purely meant to be a qualitative relationship rather than a quantitative mapping–in reality, T cells are highly cross-reactive [17], so distinct clonotypes are far from independent detectors and d is more accurately interpreted as the number of functionally independent signals that the immune cells can process than any specific repertoire count. Note from Eq 3.2 that a reduction in d also decreases the baseline discriminability index $\mathcal{D}_0$, hampering the immune system classification accuracy.

Other contributors to discriminability degradation include high levels of background stress and deceptive cancer cell signaling. Concerning the former, an inflammatory microenvironment that accompanies chronic disease, persistent infections, or repeated injuries can activate stress response programs in the tissue such as the upregulation of heat shock proteins to combat protein misfolding from oxidative damage [18]. However, these molecular signals overlap with those expressed by cancer cells: indeed, heat shock proteins are overexpressed across a range of cancers [19]. This leads to a high degree of overlap in signal space, which is formalized as a stress-induced shift in the mean of the normal cell distribution: $\mu_N(s) = (1-s)\mu_N + s\mu_C = s\delta\mathbf{1}$, where $s \in [0,1]$ quantifies the level of stress. Thus, signal discriminability is eroded by a factor of $1-s$: $\mathcal{D}(s) = \delta(1-s)\sqrt{d}/\sigma$. Cancer mimicry has an analogous impact on classification. Namely, cancer cells often mask their identity to appear normal by, e.g., reducing MHC-1 expression [20] or shedding NKG2D ligands from their cell surface [21], thereby similarly eroding the discriminability of the immune system. Finally, another mechanism that can adversely impact the immune system's classification ability is the presence of high levels of noise in signaling space, which can result from, e.g., stochastic gene expression. Mathematically, this increases the variance, σ, in our model and degrades discriminability.

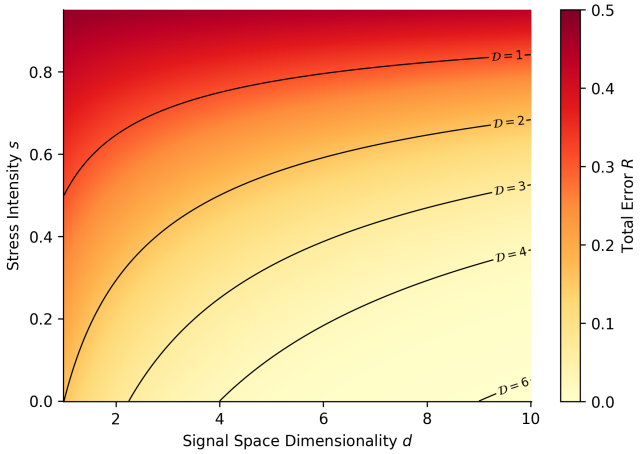


Figure 3. Total error rate and discriminability landscapes. Total error rate increases with stress and decreases with dimensionality. Discriminability decreases with stress and increases with dimensionality.

In this paper, we focus on the role of signal space dimensionality, d, and signal overlap as induced by chronic stress, s, on classification. To visualize this, we plot the total classification error $R = P(\theta_C)\beta + (1-P(\theta_C))\alpha$ as functions of d and s for a balanced classifier ($\lambda_1 = \lambda_2$) in Figure 3. From Eq 3.5, we have $R = \Phi(-\mathcal{D}(s)/2)$. As we can see, discriminability decreases with stress and increases with dimensionality. Due to this, the total error rate shows the opposite trend, increasing with stress

and decreasing with dimensionality. Note that d, s, and σ need not act in isolation. For example, chronic stress may simultaneously increase noise and reduce signal dimensionality, say with $\sigma(s) = \sigma_0(1 + as)$ and $d(s) = d_0(1 - bs)$. Then, we have a discriminability that is degraded even further: $\mathcal{D}(s) = \delta(1 - s)\sqrt{d_0(1 - bs)}/\sigma_0(1 + as)$. Thus, treating each route in isolation provides a conservative estimate on the error rates; the qualitative results hold for when these mechanisms act together.

These findings have potential clinical implications, particularly for cancer immunotherapy. Here, we distinguish between two different kinds of immunotherapy. The first are antigen-directed therapies such as CAR-T therapy and bispecific T cell engagers that aim to help immune cells recognize cancer cells based on their surface antigens. In other words, they aim to improve discriminability. The second class of immunotherapies are checkpoint blockade therapies that activate the immune system by inhibiting checkpoint proteins such as PD-1, its ligand PD-L1, and CTLA-4 and blocking these "off switches". In other words, they shift the immune system towards a more aggressive strategy by increasing λ_2. This decreases c, which moves the immune system along a fixed receiver operating characteristic (ROC) curve (red arrow) with a trade-off between cancer cells tolerated and normal cells attacked. Conversely, an increase in $\mathcal{D}$ corresponds to a shift of the ROC curve entirely (blue arrow), allowing both type 1 and type 2 error rates to simultaneously decrease (Figure 4).

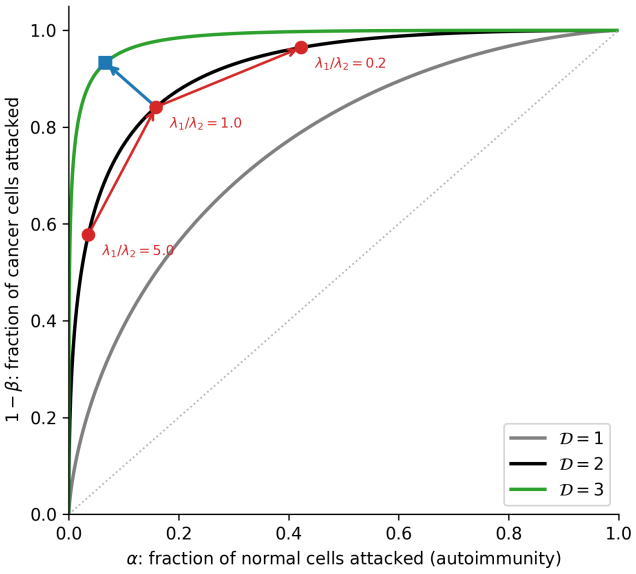


Figure 4. Threshold Shift vs. Discriminability Improvement. ROC curves are plotted for $\mathcal{D} = 1, 2, 3$. Increasing λ_2 (checkpoint inhibitors) moves the immune system along a fixed ROC curve, increasing cancer cell killing at a cost of increased autoimmunity. Increasing $\mathcal{D}$ (antigen-directed therapies) moves the ROC curve, allowing both type 1 and type 2 errors to decrease.

Based on the framework we developed here, we expect these two classes of therapies to have different side effects. Specifically, we make two predictions. First, a therapy that only changes c moves the immune system along a fixed ROC curve; thus, tumor control must come with a cost of autoimmunity. Second, checkpoint inhibitors and antigen-directed therapies should differ in the localization of their type 1 errors: checkpoint inhibitors "release the brakes on the immune system" broadly, whereas antigen-directed therapies are highly specific; thus, immune-related adverse events

(irAEs) should be much more systemic for the former compared to the latter.

Results from randomized, phase III clinical trials of checkpoint inhibitors in advanced melanoma support our first prediction. Patients who received 10 mg/kg of ipilimumab had improved survival compared to those who received 3 mg/kg (median overall survival of 15.7 vs. 11.5 months), but also experienced treatment-related adverse events (TRAE) at higher rates (37% vs. 18%) [22]. Similarly, in the Checkmate 067 trial, a combination nivolumab and ipilimumab treatment led to improved patient survival: median progression-free survival of 11.5 months vs. 2.9 months (ipilimumab alone) or 6.9 months (nivolumab alone), at the cost of higher TRAE: 55% vs. 27.3 % (ipilimumab alone) or 16.3 % (nivolumab alone) [23]. A phase III clinical trial used Tebentafusp, a bispecific T cell engager for gp100, to treat metastatic uveal melanoma, which allows us to test our second prediction [24]. Patients who received this drug showed improved survival outcomes compared to the control, standard-of-care therapy (overall survival at one year of 73% vs. 59%) and experienced higher levels of TRAE (44% vs. 17%). However, key side effects were skin-related events, such as rash and pruritus, due to the targeting of gp100-positive melanocytes. These tissue-specific side effects are in accordance with our second prediction, as contrasted with the systemic irAEs often observed with checkpoint inhibition [25, 26].

4.2. Dynamic calibration of normalcy

Thus far, we have fixed a static immune classifier, assuming that the immune system is always calibrated to the tissue it inhabits, and analyzed the implications of different tissue conditions and thresholds on type 1 and type 2 error rates. However, what remains unexplored is how the immune system updates its classifier in response to a dynamically changing environment. A key conceptual tension naturally arises when we consider this question: how does the immune system perceive "normality"? The traditional account is the self/non-self theory, which holds that the immune system is shaped through ontogeny to tolerate self and responds to entities that fall outside this repertoire of tolerated stimuli. An alternative is Pradeu's discontinuity theory, which argues that immune responses are driven not primarily by foreignness, but by abrupt changes in the molecular interactions to which the immune system is ordinarily exposed to [5, 6]. Here, we show that adopting a phenomenological perspective in which the immune system updates its internal reference for normalcy over time naturally captures the discontinuity theory: rapid changes are flagged as abnormal, whereas slow changes are absorbed into the system's notion of normalcy.

In our static approach, we assumed that the immune system classifies signals relative to the true normal mean, μ_N. However, this assumes that the immediate object coincides with the dynamical object. Here, we take a phenomenological perspective and relax this assumption, allowing the immune system to classify signals instead based on an evolving internal reference $\hat{\mu}_N$, which dynamically updates based on the signals it encounters as follows:

$$\frac{d\hat{\mu}_N}{dt} = \eta(\mu_N - \hat{\mu}_N), \tag{4.1}$$

where $\eta \geq 0$ is the rate of habituation. The decision rule remains the same as before, $\Lambda_t(m) > c$, with the true normal mean replaced by the immune system's internal model of the normal mean. Although we do not explicitly deal with the inherent stochasticity of this system, it is worth briefly noting that if the signal drifts as a random walk with volatility v and observational noise n, then the optimal estimate of the tissue state, given by the Kalman-Bucy filter, is Eq 4.1 with $\eta = v/n$ [27]. Intuitively, this

suggests that the immune memory is longer when signals are noisy and shorter when the tissue state rapidly changes.

There is empirical evidence that $\hat{\mu}_N$ adapts in response to the environment. Lymphocytes have been shown to dynamically tune their responsiveness to recurrent stimuli [28, 29], sustained NKG2D ligand expression in normal epithelia causes (reversible) systemic downregulation of NKG2D expression and increased tumor susceptibility [30], and mature NK cells adapt their responsiveness when placed into an altered MHC environment [31].

For the following analysis, we allow stress to enter the tissue via $\mu_N(t) = s(t)\delta\mathbf{1}$ (and thus $\hat{\mu}_N(t) = \hat{s}(t)\delta\mathbf{1}$) with $s \in [0, 1)$, and we assume that the immune system starts calibrated to the tissue: $\mu_N(0) = \hat{\mu}_N(0)$. For simplicity, we assume that stress does not alter μ_C. Then, we can rewrite Eq 4.1 as follows:

$$\frac{d\hat{s}}{dt} = \eta(s - \hat{s}), \tag{4.2}$$

where the lag $l(t) = s(t) - \hat{s}(t)$ quantifies the divergence between the stress in the tissue (dynamical object) and the immune system's perception of it (immediate object): it is positive (negative) when stress has risen (fallen) faster than the immune system has habituated to it. Letting $\hat{\mathcal{D}} = (1 - \hat{s})\mathcal{D}_0$ be the perceived discriminability of the immune system, we have the following proposition that links the lag to a shift in classifier threshold.

Proposition 4.1. *Let $\alpha(t)$ and $\beta(t)$ be the type 1 and type 2 error rates, respectively, of the decision rule $\Lambda_t(m) > c$. Then,*

$$\alpha(t) = \Phi\left(-\frac{c}{\hat{\mathcal{D}}} - \frac{\hat{\mathcal{D}}}{2} + \mathcal{D}_0\, l\right), \qquad \beta(t) = \Phi\left(\frac{c}{\hat{\mathcal{D}}} - \frac{\hat{\mathcal{D}}}{2}\right), \tag{4.3}$$

and the point $(\alpha(t), 1 - \beta(t))$ lies on the ROC curve of the true discriminability $\mathcal{D}$.

Proof. Define $z = \mathbf{1}^T m/\sqrt{d}$ as the projection of m onto the axis connecting the means. Under the normal state θ_N, we have $z \sim \mathcal{N}(s\delta\sqrt{d}, \sigma^2)$ and under the cancer state θ_C, we have $z \sim \mathcal{N}(\delta\sqrt{d}, \sigma^2)$. Then, the decision rule $\Lambda_t > c$ is equivalent to the condition $z > c\sigma/\hat{\mathcal{D}} + \delta\sqrt{d}(1 + \hat{s})/2$. From this, we can compute the error rates as $\alpha = \Phi\left(-\frac{c}{\hat{\mathcal{D}}} - \frac{\hat{\mathcal{D}}}{2} + \mathcal{D}_0\, l\right)$ and $\beta = \Phi\left(\frac{c}{\hat{\mathcal{D}}} - \frac{\hat{\mathcal{D}}}{2}\right)$, as desired. Then, we can compute $\Phi^{-1}(1 - \beta) - \Phi^{-1}(\alpha) = \hat{\mathcal{D}} - \mathcal{D}_0 l = \mathcal{D}(s)$. Thus, $(\alpha(t), 1 - \beta(t))$ lies on the ROC curve of the true discriminability. □

Note that for a balanced classifier ($c = 0$), Eq 4.3 reduces to the following:

$$\alpha(t) = \Phi\left(-\frac{\mathcal{D}}{2} + \frac{\mathcal{D}_0 l}{2}\right), \qquad \beta(t) = \Phi\left(-\frac{\mathcal{D}}{2} - \frac{\mathcal{D}_0 l}{2}\right). \tag{4.4}$$

Thus, when stress rises faster than the immune system has habituated ($l > 0$), normal cells are perceived as more cancer-like by the immune classifier, leading to an increase in type 1 errors relative to a classifier that is instantaneously calibrated to the tissue state and a move along the true ROC curve towards a more aggressive strategy. Conversely, when stress falls faster than the immune system has habituated ($l < 0$), cancer cells are perceived as more normal-like by the immune classifier, leading to an increase in type 2 errors and a move along the true ROC curve towards a more tolerant strategy. Eventually, assuming $\eta > 0$, as the immune classifier calibrates to the new tissue state, $\hat{s} \to s$, $l \to 0$, and $\alpha, \beta \to \Phi(-\mathcal{D}/2)$.

The properties of this transient mismatch depend on the relative speeds of tissue change and immune habituation. To formalize this, we consider a stress that moves linearly from s_0 to s_1 from time 0 to time T, and stays at s_1 for $t > T$: $s(t) = s_0 + (s_1 - s_0)\min(t/T, 1)$. The smaller T is, the more abrupt the stress is. The case $s_1 > s_0$ captures the development of stress (e.g., the initiation of chain smoking), whereas $s_0 > s_1$ captures recovery (e.g., cessation of smoking). Then, we have the following proposition.

Proposition 4.2. *For a tissue with a balanced immune classifier subject to linearly moving stress from s_0 to s_1 over the interval $[0, T]$, assuming $\hat{s}(0) = s_0$, $s_0 \neq s_1$, and $T, \eta > 0$, the type 1 error rate attains its unique global extremum (maximal if $s_1 > s_0$ and minimal if $s_0 > s_1$) at $t = T$, where*

$$\alpha_{\text{ext}} = \Phi\left(-\frac{\mathcal{D}(s_1)}{2} + \frac{\mathcal{D}_0\,(s_1 - s_0)\,f(\eta T)}{2}\right), \qquad f(x) = \frac{1 - e^{-x}}{x}. \tag{4.5}$$

Proof. Differentiating $l = s - \hat{s}$ gives us $dl/dt = ds/dt - \eta l$, with $l(0) = 0$. For $t \leq T$, we can solve this to obtain $l(t) = \frac{s_1 - s_0}{\eta T}(1 - e^{-\eta t})$, and for $t > T$, we have $l(t) = l(T)e^{-\eta(t-T)}$. Thus, we have $l(T) = (s_1 - s_0)f(\eta T)$. Then, by substituting in the appropriate values into Equation 4.4, we can write $\alpha(T) = \Phi\left(-\frac{\mathcal{D}(s_1)}{2} + \frac{\mathcal{D}_0\,(s_1 - s_0)\,f(\eta T)}{2}\right)$. To show that this is the unique global extremum, let $g := s + l - 1$ such that $\alpha(t) = \Phi(\mathcal{D}_0 g(t)/2)$ and note that $\operatorname{sgn}(\alpha') = \operatorname{sgn}(g')$. Since $dg/dt = ds/dt + dl/dt$, for $t \leq T$, this simplifies to $\frac{s_1 - s_0}{T}(1 + e^{-\eta t})$ and $\operatorname{sgn}(g') = \operatorname{sgn}(s_1 - s_0)$. For $t > T$, $dg/dt = -\eta l(T)e^{-\eta(t-T)}$ and $\operatorname{sgn}(g') = -\operatorname{sgn}(s_1 - s_0)$. Since g is continuous, it is strictly monotone on $[0, T]$ and is strictly monotone in the opposite direction on $[T, \infty)$, so $t = T$ is the unique global extremum. □

$f(\eta T)$ represents the fraction of the stress change that remains a discrepancy between the dynamical and immediate object at time T, and entirely depends on the duration of the tissue change relative to the immune system's memory. When stress abruptly rises ($\eta T << 1$), almost the full change remains unabsorbed ($f \approx 1$), so the tissue is attacked beyond what the true state warrants and is treated as abnormal. When stress gradually rises ($\eta T >> 1$), changes are absorbed ($f \approx 1/(\eta T)$) and the new tissue state is classified as normal. Abrupt resolutions of stress produce analogous responses, leading to transient hyporesponsiveness rather than a heightened response. These findings are closely related to the broader use of delay differential equations in modeling tumor-immune dynamics [32–34]. Namely, we can integrate Eq 4.1, assuming $s(t) = s_0$ for $t < 0$, to get $\hat{s}(t) = \int_0^\infty \eta e^{-\eta u} s(t - u)du$. Thus, the immune system's perception is a distributed delay over the tissue's past states, with an exponential kernel and a mean delay of $1/\eta$[‡].

Not only do these results naturally capture the discontinuity theory of immunity through a phenomenological lens, but they also have a natural interpretation in the context of cancer. A tissue that absorbs increasingly stressed signals into its notion of normalcy, as often occurs during chronic disease, tolerates cancer cells that display similar signals at higher rates. As observed in [30], sustained expression of NKG2D ligands in a healthy epithelium trained the immune system to interpret these signals as normal. As a result, systemic NKG2D expression downregulation was noticed, which increased tumor susceptibility. Thus, our results provide a semiotic interpretation of the long-recognized connection between chronic disease and cancer.

[‡]As an aside, our findings do not rely on the exponential nature of this kernel. For any delay distribution U such that $\mathbb{E}(U) = 1/\eta$, the unabsorbed fraction is $\mathbb{E}[\min(U, T)]/T$. For $U \sim \text{Exp}(\eta)$, this simplifies to $f(\eta T)$, but for any distribution, this approaches 1 for abrupt changes and $1/(\eta T)$ for gradual changes.

5. Conclusions

In this paper, we used a Bayesian semiotic modeling approach to argue that cancer and autoimmunity may simply be two sides of the same semiotic corruption problem, rather than separate biomedical pathologies[§]. We focused on two aspects of the Peircean triad that, under pathological conditions, are corrupted: the representamen and the interpretant. The former is caused by deceptive cancer cell signaling and the latter is caused by high levels of noise in signal space, significant background stress in the microenvironment, or a degradation of signal dimensionality as caused by immunosenescence. This leads to the inaccurate registration of malignant cells as "safe" objects to be tolerated (type 2 error: cancer) or normal cells as "dangerous" objects to be attacked (type 1 error: autoimmunity).

We provided a mathematical formalization of these ideas using concepts from signal detection theory and illustrated how immune discriminability and classification error rate change as a function of signal dimensionality, stress, deceptive cancer cell signaling, and noise. We predicted that checkpoint inhibitors that shift the immune classifier to a more aggressive state on the ROC curve would lead to an improved tumor response at the cost of autoimmune toxicity, a prediction supported by recent randomized clinical trials in melanoma [22, 23]. We also predicted that the immune-related adverse effects would be more systemic in response to checkpoint inhibitors and more localized to tissues that share the target antigen in response to antigen-directed therapies, consistent with the melanocyte-toxicities observed in response to Tebentafusp [24].

Additionally, we extended our static framework to allow for an immune system that dynamically adapts its notion of normalcy over time. We showed that a lag between the real state of the tissue (dynamical object) and the immune system's perception of it (immediate object) transiently shifts the immune classifier along the true ROC curve: towards aggression if stress rises faster than immune habituation, and towards tolerance if it falls faster. The fraction of change that registers as a discrepancy only depends on the duration of the change relative to the immune system's memory; thus, abrupt changes are classified as abnormal, whereas gradual changes are absorbed into the system's evolving notion of normalcy. Although we framed this discussion in the context of cancer, our results are broadly applicable to any biosocial system in which classification decisions must be made based on imperfect and corruptible signals without direct knowledge of the underlying state of the sender.

Use of AI tools declaration

The author declares the use of Claude Opus 4.8 for copyediting purposes and to help in writing the code for numerical simulations, which the author ran to produce Figures 2–4. All results have been verified for accuracy by the author, who takes full responsibility for the code and the content of this article.

Conflict of interest

The author declares there is no conflict of interest.

[§]This approach joins recent work that models semiosis as inference under uncertainty via Bayesian inference or economy-driven heuristics in the context of representations between human and SARS-CoV-2 and translation [35, 36]. Our work here places this within a cellular phenomenology context and further formalizes the Bayesian semiotic machinery.

References

1. H. E. Teulings, J. Limpens, S. N. Jansen, A. H. Zwinderman, J. B. Reitsma, P. I. Spuls, et al., Vitiligo-like depigmentation in patients with stage III-IV melanoma receiving immunotherapy and its association with survival: A systematic review and meta-analysis, *J. Clin. Oncol.*, **33** (2015), 773–781. https://doi.org/10.1200/JCO.2014.57.4756
2. E. A. Engels, R. M. Pfeiffer, J. F. Fraumeni, B. L. Kasiske, A. K. Israni, J. J. Snyder, et al., Spectrum of cancer risk among US solid organ transplant recipients, *JAMA*, **306** (2011), 1891. https://doi.org/10.1001/jama.2011.1592
3. S. M. Amos, C. P. M. Duong, J. A. Westwood, D. S. Ritchie, R. P. Junghans, P. K. Darcy, et al., Autoimmunity associated with immunotherapy of cancer, *Blood*, **118** (2011), 499–509. https://doi.org/10.1182/blood-2011-01-325266
4. M. A. Postow, R. Sidlow, M. D. Hellmann, Immune-related adverse events associated with immune checkpoint blockade, *N. Engl. J. Med.*, **378** (2018), 158–168. https://doi.org/10.1056/NEJMra1703481
5. T. Pradeu, S. Jaeger, E. Vivier, The speed of change: towards a discontinuity theory of immunity?, *Nat. Rev. Immunol.*, **13** (2013), 764–769. https://doi.org/10.1038/nri3521
6. T. Pradeu, E. Vivier, The discontinuity theory of immunity, *Sci. Immunol.*, **1** (2016). https://doi.org/10.1126/sciimmunol.aag0479
7. J. J. von Uexkül, *A Foray Into the Worlds of Animals and Humans: With A Theory of Meaning*, University of Minnesota Press, Minneapolis, 2010.
8. I. Pavlinov, *Biological Systematics: History and Theory*, CRC Press, 2022.
9. C. S. Peirce, *Collected Papers of Charles Sanders Peirce*, Harvard University Press, 1934.
10. S. Brier, *Cybersemiotics: Why Information is Not Enough*, University Of Toronto Press, 2013.
11. A. A. Sharov, Functional information: Towards synthesis of biosemiotics and cybernetics, *Entropy*, **12** (2010), 1050–1070. https://doi.org/10.3390/e12051050
12. U. Eco, *A Theory of Semiotics*, Indiana University Press, 1978.
13. J. E. Park, R. A. Botting, C. Domínguez Conde, D. M. Popescu, M. Lavaert, D. J. Kunz, et al., A cell atlas of human thymic development defines T cell repertoire formation, *Science*, **367** (2020). https://doi.org/10.1126/science.aay3224
14. J. J. Goronzy, C. M. Weyand, Understanding immunosenescence to improve responses to vaccines, *Nat. Immunol.*, **14** (2013), 428–436. https://doi.org/10.1038/ni.2588
15. N. Song, M. A. Elbahnasawy, N. P. Weng, General and individualized changes in T cell immunity during aging, *J. Immunol.*, **214** (2025), 872–879. https://doi.org/10.1093/jimmun/vkae033
16. N. P. Weng, Numbers and odds: TCR repertoire size and its age changes impacting on T cell functions, *Semin. Immunol.*, **69** (2023), 101810. https://doi.org/10.1016/j.smim.2023.101810
17. A. K. Sewell, Why must T cells be cross-reactive?, *Nat. Rev. Immunol.*, **12** (2012), 669–677. https://doi.org/10.1038/nri3279

18. P. S. Kumar, T. Kodi, A. Gopinathan, B. H. Byregowda, K. Nandakumar, A. Kishore, New insights on heat shock proteins as regulators of reactive oxygen species across various stressors in diseases, *Cell Biochem. Funct.*, **44** (2026). https://doi.org/10.1002/cbf.70173

19. D. R. Ciocca, S. K. Calderwood, Heat shock proteins in cancer: diagnostic, prognostic, predictive, and treatment implications, *Cell Stress Chaperones*, **10** (2005), 86.

20. K. Dhatchinamoorthy, J. D. Colbert, K. L. Rock, Cancer immune evasion through loss of MHC class I antigen presentation, *Front. Immunol.*, **12** (2021). https://doi.org/10.3389/fimmu.2021.636568

21. G. Tan, K. M. Spillane, J. Maher, The role and regulation of the NKG2D/NKG2D ligand system in cancer, *Biology*, **12** (2023), 1079. https://doi.org/10.3390/biology12081079

22. P. A. Ascierto, M. Del Vecchio, C. Robert, A. Mackiewicz, V. Chiarion-Sileni, A. Arance, et al., Ipilimumab 10 mg/kg versus ipilimumab 3 mg/kg in patients with unresectable or metastatic melanoma: a randomised, double-blind, multicentre, phase 3 trial, *Lancet Oncol.*, **18** (2017), 611–622. https://doi.org/10.1016/S1470-2045(17)30231-0

23. J. Larkin, V. Chiarion-Sileni, R. Gonzalez, J. J. Grob, C. L. Cowey, C. D. Lao, et al., Combined nivolumab and ipilimumab or monotherapy in untreated melanoma, *N. Engl. J. Med.*, **373** (2015), 23–34. https://doi.org/10.1056/NEJMoa1504030

24. P. Nathan, J. C. Hassel, P. Rutkowski, J. F. Baurain, M. O. Butler, M. Schlaak, et al., Overall survival benefit with tebentafusp in metastatic uveal melanoma, *N. Engl. J. Med.*, **385** (2021), 1196–1206. https://doi.org/10.1056/NEJMoa2103485

25. J. Chen, Y. Alduais, B. Chen, Therapeutic and systemic adverse events of immune checkpoint inhibitors targeting the PD-1/PD-L1 axis for clinical management of NSCLC, *Cell Transplant.*, **30** (2021). https://doi.org/10.1177/09636897211041587

26. Q. Yin, L. Wu, L. Han, X. Zheng, R. Tong, L. Li, et al., Immune-related adverse events of immune checkpoint inhibitors: a review, *Front. Immunol.*, **14** (2023). https://doi.org/10.3389/fimmu.2023.1167975

27. R. E. Kalman, R. S. Bucy, New results in linear filtering and prediction theory, *J. Basic Eng.*, **83** (1961), 95–108. https://doi.org/10.1115/1.3658902

28. Z. Grossman, W. E. Paul, Adaptive cellular interactions in the immune system: the tunable activation threshold and the significance of subthreshold responses, *PNAS*, **89** (1992), 10365–10369. https://doi.org/10.1073/pnas.89.21.10365

29. Z. Grossman, W. E. Paul, Dynamic tuning of lymphocytes: Physiological basis, mechanisms, and function, *Annu. Rev. Immunol.*, **33** (2015), 677–713. https://doi.org/10.1146/annurev-immunol-032712-100027

30. D. E. Oppenheim, S. J. Roberts, S. L. Clarke, R. Filler, J. M. Lewis, R. E. Tigelaar, et al., Sustained localized expression of ligand for the activating NKG2D receptor impairs natural cytotoxicity in vivo and reduces tumor immunosurveillance, *Nat. Immunol.*, **6** (2005), 928–937. https://doi.org/10.1038/ni1239

31. N. T. Joncker, N. Shifrin, F. Delebecque, D. H. Raulet, Mature natural killer cells reset their responsiveness when exposed to an altered MHC environment, *J. Exp. Med.*, **207** (2010), 2065–2072. https://doi.org/10.1084/jem.20100570

32. H. J. Alsakaji, F. A. Rihan, K. Udhayakumar, F. El Ktaibi, Stochastic tumor-immune interaction model with external treatments and time delays: An optimal control problem, *Math. Biosci. Eng.*, **20** (2023), 19270–19299. https://doi.org/10.3934/mbe.2023852

33. F. A. Rihan, D. H. Abdelrahman, F. Al-Maskari, F. Ibrahim, M. A. Abdeen, Delay differential model for tumour-immune response with chemoimmunotherapy and optimal control, *Comput. Math. Methods Med.*, **2014** (2014), 1–15. https://doi.org/10.1155/2014/982978

34. F. A. Rihan, *Delay Differential Equations and Applications to Biology*, Springer Singapore, Singapore, 2021. https://doi.org/10.1007/978-981-16-0626-7

35. S. Torres-Martínez, An integrated Bayesian-heuristic semiotic model for understanding human and SARS-CoV-2 representational structures, *Biosemiotics*, **16** (2023), 415–439. https://doi.org/10.1007/s12304-023-09546-7

36. S. Torres-Martínez, Semiosic translation: a Bayesian-heuristic theory of translation and translating, *Lang. Semiotic Stud.*, **10** (2024), 167–202. https://doi.org/10.1515/lass-2023-0042



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0)