

Mechanistically Informed Prioritization of RNA–LNP Therapeutics: a Safety-by- Design Framework Integrating Virtual Patients, Synthetic Multi-Omics and Biological Consistency Assessment

Valentina Di Salvatore

University of Catania

Giulia Russo

University of Catania

Francesco Pappalardo

`francesco.pappalardo@unict.it`

University of Catania

Research Article

Keywords: RNA therapeutics, lipid nanoparticles, safety-by-design, mechanistic modeling, virtual patients, synthetic multi-omics, explainable artificial intelligence, systems immunology

Posted Date: August 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10513597/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Bioinformatics on September 7th, 2026. See the published version at <https://doi.org/10.1186/s12859-026-06630-w>.

Mechanistically Informed Prioritization of RNA–LNP Therapeutics: a Safety-by-Design Framework Integrating Virtual Patients, Synthetic Multi-Omics and Biological Consistency Assessment

Valentina Di Salvatore
*Department of Health and
Drug Sciences
University of Catania
Catania, Italy
valentinadisalvatore@unict.it*

Giulia Russo
*Department of Health and
Drug Sciences
University of Catania
Catania, Italy
giulia.russo@unict.it*

Francesco Pappalardo*
*Department of Health and
Drug Sciences
University of Catania
Catania, Italy
francesco.pappalardo@unict.it*

*Corresponding author

Abstract

Background: RNA therapeutics are rapidly transforming precision medicine, yet their development requires simultaneous consideration of delivery efficiency, biological efficacy, and immune safety. Existing computational approaches frequently address these objectives separately and rarely connect mechanistic knowledge, artificial intelligence, formulation prioritization, and independent biological consistency assessment within a unified Safety-by-Design framework.

Methods: We developed an integrated computational framework combining virtual patient cohorts, an RNA–lipid nanoparticle (RNA-LNP) formulation design space, mechanistic immune descriptors derived from previously published Universal Immune System Simulator (UISS) simulations, synthetic multi-omics generation, explainable artificial intelligence, and Pareto-based multi-objective prioritization. The mechanistic descriptors were condensed into five Mechanistically Informed Immune Signatures (MIIS), which served as biological constraints for generating coherent synthetic molecular profiles. A Random Forest surrogate model estimated a composite Safety-by-Design score, while biological consistency was retrospectively assessed against the independent human single-cell RNA-sequencing dataset GSE171964.

Results: Using formulation-grouped five-fold cross-validation, the Random Forest surrogate model predicted the composite Safety-by-Design score with $R^2 = 0.887$ and $RMSE = 0.055$, while permutation-based importance identified zeta potential, nanoparticle size, targeting-ligand status, and PEGylation density as the principal determinants of model predictions. Retrospective comparison with GSE171964 yielded BPCI values ranging from 0.57 to 0.88 (mean = 0.76), with

the highest concordance observed at Days 2, 22, and 28, supporting preservation of systems-level immune-program organization.

Conclusions: The proposed framework provides a reproducible proof-of-concept strategy for the early computational prioritization of RNA-LNP formulations. MIIS provide an interpretable interface between mechanistic immune information and synthetic molecular representations, while the BPCI enables systems-level assessment of biological consistency. Validation across additional independent cohorts and prospective experimental studies will be required to establish predictive performance, generalizability, and potential translational utility.

Keywords: RNA therapeutics; lipid nanoparticles; safety-by-design; mechanistic modeling; virtual patients; synthetic multi-omics; explainable artificial intelligence; systems immunology.

1. Background

1.1 RNA therapeutics: opportunities and challenges

RNA-based therapeutics have emerged as one of the most transformative innovations in modern medicine, offering unprecedented opportunities for the prevention and treatment of infectious diseases, cancer, genetic disorders, and inflammatory conditions (1). The clinical success of messenger RNA (mRNA) vaccines during the COVID-19 pandemic (2)(3) demonstrated the remarkable potential of nucleic acid-based medicines, accelerating investments and research efforts across the broader field of RNA therapeutics(4). Beyond conventional mRNA vaccines, the therapeutic landscape now includes self-amplifying RNA (saRNA), small interfering RNA (siRNA), antisense oligonucleotides (ASOs), RNA aptamers, RNA editing technologies, and other programmable nucleic acid platforms designed to modulate gene expression or induce targeted immune responses (5).

The appeal of RNA therapeutics stems from their intrinsic versatility. Unlike traditional small molecules or biologics, RNA molecules can be rapidly designed and manufactured once a target sequence is identified, enabling unprecedented flexibility in responding to emerging pathogens, personalized medicine applications, and rare genetic diseases. Moreover, advances in synthetic biology, computational design, and high-throughput screening have significantly expanded the range of diseases potentially addressable through RNA-based interventions(6).

Despite these advances, several challenges continue to limit the widespread adoption of RNA therapeutics. RNA molecules are inherently unstable, susceptible to enzymatic degradation, and unable to efficiently cross biological membranes without specialized delivery systems. Furthermore, both the therapeutic cargo and the delivery vehicle may trigger unintended immune activation, inflammatory responses, or off-target biological effects. These challenges highlight the need for integrated strategies capable of simultaneously optimizing efficacy, biodistribution, and safety throughout the therapeutic development process (7).

1.2 Lipid nanoparticles as delivery platforms

The success of RNA therapeutics is closely linked to the development of effective delivery technologies. Among the available non-viral delivery systems, lipid nanoparticles (LNPs) have emerged as the most clinically advanced and widely adopted platform. LNPs protect RNA molecules from extracellular degradation, facilitate cellular uptake, promote endosomal escape (8), and contribute significantly to tissue-specific biodistribution(9), (10). Their central role in the development of approved mRNA vaccines has established LNPs as a cornerstone of modern RNA medicine. However, the biological behavior of LNPs is highly sensitive to their physicochemical properties (11). Some ionizable or cationic LNP compositions can induce substantial innate inflammatory responses (12), (13), whereas particle size and surface charge influence lymph-node transitivity and distribution (14). PEGylation modulates colloidal stability and circulation and may also affect cellular uptake and immunogenicity.

The multidimensional nature of LNP design creates a complex optimization problem involving multiple competing objectives(15,16). An ideal formulation should maximize delivery efficiency and therapeutic activity while minimizing systemic toxicity, off-target biodistribution, and unwanted immune activation. Exploring this vast design space exclusively through experimental approaches remains challenging, costly, and time-consuming, motivating the development of complementary computational strategies capable of guiding formulation design prior to laboratory testing. These challenges naturally motivate computational Safety-by-Design frameworks capable of exploring large formulation spaces before experimental testing.

1.3 Computational approaches for RNA therapeutic design

Recent advances in computational biology, artificial intelligence (AI), systems pharmacology, and mechanistic modeling have created new opportunities for the rational design of RNA therapeutics (15). Machine learning algorithms are increasingly being applied to predict RNA stability, optimize sequence design, estimate delivery performance, and identify critical formulation parameters. At the same time, mechanistic simulation platforms have demonstrated their value in reproducing complex biological processes, including immune responses, pharmacokinetics, biodistribution, and tissue-level interactions. The growing regulatory interest in New Approach Methodologies (NAMs) further emphasizes the importance of computational frameworks capable of reducing reliance on animal experimentation while supporting evidence generation during preclinical development. Despite remarkable advances, current computational approaches remain fragmented, typically addressing isolated aspects of RNA therapeutic development rather than supporting integrated decision-making across efficacy, safety, delivery, and biological validation. In this context, the integration of mechanistic simulations with data-driven AI approaches represents a promising paradigm for model-informed drug development (17)(18).

Nevertheless, significant limitations remain. Computational approaches to RNA therapeutic development are still commonly organized around specific tasks or modeling scales, including molecular design, formulation optimization, delivery modeling, and transcriptomic analysis. Integrated multiscale frameworks connecting delivery-system design, mechanistic immune simulation, safety-related endpoints, and biological validation remain comparatively underdeveloped(19). Moreover, optimization efforts have frequently emphasized delivery efficiency and therapeutic performance, underscoring the need to incorporate safety and immunogenicity considerations earlier in candidate selection(9).

As RNA therapeutics continue to expand beyond vaccines toward broader clinical applications, there is an increasing need for computational platforms capable of supporting a safety-by-design strategy in which efficacy and safety are evaluated simultaneously throughout the development process.

1.4 Previous work and rationale

In our previous work, we presented a proof-of-concept computational strategy for the rational design of RNA–lipid nanoparticle (RNA-LNP) formulations by integrating mechanistic immune modeling, synthetic transcriptomic data generation, machine learning, and computational optimization. That study demonstrated that mechanistically informed synthetic datasets could support the identification of candidate formulations with favorable delivery characteristics while reducing predicted inflammatory activation, highlighting the potential of computational approaches to accelerate early-stage formulation design (20).

Although encouraging, the previous framework remained primarily focused on formulation optimization and represented an initial step toward a more comprehensive computational pipeline for RNA therapeutics. It employed a single synthetic transcriptomic layer and represented biological variability through stochastic perturbations rather than explicit virtual patient cohorts. Retrospective validation against the independent GSE171110 dataset was performed at the gene level using differential-expression overlap, directionality concordance, and fold-change correlation. However, the framework did not include a systems-level strategy for determining whether the simulated responses preserved the relative organization and hierarchy of coordinated immune programs.

These limitations motivated the development of the present study, whose objective is not merely to improve formulation optimization, but to establish a modular safety-by-design framework that integrates mechanistic immune modeling, synthetic multi-omics generation, explainable artificial intelligence, virtual patient modeling, and biological consistency assessment within a unified computational workflow. The conceptual advance introduced here is not simply the addition of new computational components, but the establishment of a unified framework in which

mechanistic simulations provide biologically interpretable constraints for downstream AI-driven prioritization and validation.

In the present study, these previously generated mechanistic simulations are not repeated but are repurposed as a biologically grounded source of immune dynamics supporting the generation of synthetic multi-omics representations within a substantially extended safety-by-design framework.

1.5 Objectives of the present study

The objective of this study was to develop an integrated safety-by-design computational framework for the early prioritization of RNA therapeutics by combining mechanistic immune modeling, synthetic multi-omics generation, explainable artificial intelligence, and retrospective biological validation.

The proposed workflow combines heterogeneous virtual patient cohorts with mechanistic immune descriptors derived from previously published UISS simulations. These descriptors are condensed into Mechanistically Informed Immune Signatures (MIIS) and used as biological constraints for generating coherent synthetic multi-omics representations encompassing transcriptomic, cytokine, immune-cellular, and pathway-level responses. These representations are used to characterize the synthetic immune landscape and provide the functional space for retrospective biological consistency assessment. In a parallel modeling branch, a Random Forest surrogate uses seven formulation parameters and four latent immune-state variables to estimate the composite Safety-by-Design score. Candidate RNA-LNP formulations are subsequently prioritized through Pareto-based analysis of five efficacy- and safety-related endpoint objectives, with the composite score used for post hoc ranking and visualization.

To assess the biological plausibility of the framework, predicted immune programs were retrospectively compared with an independent human single-cell transcriptomic dataset obtained following vaccination with BNT162b2 (Pfizer–BioNTech), a lipid nanoparticle (LNP)-formulated mRNA vaccine encoding the prefusion-stabilized SARS-CoV-2 spike protein and representing one of the best-characterized human mRNA vaccine systems currently available (21). Rather than evaluating agreement at the level of individual genes, validation was performed by assessing preservation of the functional hierarchy of major immune programs. To support this analysis, we introduce the Biological Program Concordance Index (BPCI), a composite metric that quantifies biological agreement by integrating immune-program ranking, activation-direction consistency, and top-program overlap.

Collectively, the proposed framework provides a reproducible and interpretable computational strategy for the early-stage development of RNA therapeutics while supporting emerging New Approach Methodologies (NAMs) and model-informed drug development. More broadly, it establishes a modular computational paradigm for mechanistically informed Safety-by-Design that

variables are used by a Random Forest surrogate to estimate the composite Safety-by-Design score. Pareto analysis is performed on five efficacy- and safety-related endpoints, and the composite score is used for post hoc ranking and visualization of non-dominated candidates. Framework-level biological consistency is assessed against independent human transcriptomic data using the Biological Program Concordance Index (BPCI).

Within this architecture, the Universal Immune System Simulator (UISS) serves as a pre-existing mechanistic knowledge layer. Rather than conducting a new simulation campaign, the present study reused previously published temporal profiles of immune-cell activation, antibody production, and response dynamics. These profiles were summarized into quantitative descriptors and aggregated into five Mechanistically Informed Immune Signatures (MIIS): early innate activation, cellular adaptive immunity, humoral response, immune memory, and global immune persistence. The resulting MIIS were then used to constrain coherent synthetic transcriptomic, cytokine, immune-cellular, and pathway-level representations (Tables 1 and 2). Importantly, UISS-derived descriptors, MIIS, and synthetic multi-omics variables belonged to the biological-consistency branch and were not used as predictors of the Random Forest surrogate or as Pareto-dominance objectives.

UISS output	Biological interpretation	Mechanistic descriptor	Synthetic multi-omics layer influenced
EP (Epithelial/target cells)	Early interaction between RNA-LNP and target cells; initiation of innate sensing	Early innate activation	Interferon response; inflammatory response
T cells	Helper T-cell activation	Cellular adaptive immunity	Adaptive immunity; antigen presentation
TC (Cytotoxic T cells)	Cytotoxic immune response	Cytotoxic adaptive activation	Adaptive immunity; antigen presentation
B cells	Activation of humoral immunity	Humoral activation	Adaptive immunity
Tmem (Memory T cells)	Development of long-term immune memory	Immune memory potential	Adaptive immunity
Ig (Immunoglobulins)	Antibody production	Humoral response intensity	Adaptive immunity
Temporal activation curves	Dynamic evolution of immune response	Peak activation, activation persistence, time-to-peak	All synthetic layers

Table 1. Mechanistic integration of UISS within the computational framework

Compartment-specific ΔAUC values were retrieved from the previously published UISS simulation outputs. By integrating baseline-adjusted immune-response trajectories over time, ΔAUC provides a cumulative descriptor that reflects both response intensity and persistence, although it does not separately quantify individual temporal features such as time to peak or response duration. From each previously generated compartment-specific UISS trajectory, five summary descriptors were computed to characterize response magnitude, cumulative activation, kinetics, and residual activity. Let $Ac(t)$ denote the simulated activation level of immune compartment c at time t . The extracted descriptors and their biological interpretation are summarized in Table 2.

Descriptor	Formula/definition	Biological interpretation
Peak activation	$\max_t(Ac(t_j))$	Maximum activation reached during the simulation
Mean activation	$\frac{1}{n} \sum_{j=1}^n (Ac(t_j))$	Average activation intensity over the simulated interval
Cumulative activation	$\int_0^{t_f} (Ac(t)) dt$, estimated using the trapezoidal rule	Overall activation burden integrating intensity and duration
Time-to-peak	$\arg(\max_t(Ac(t)))$	Timing of the maximum response
Final activation	$Ac(t_f)$	Residual activity at the final simulated time point

Table 2. Mechanistic descriptors extracted from UISS simulations

These mechanistic descriptors were aggregated into Mechanistically Informed Immune Signatures (MIIS) and used to constrain the generation of synthetic multi-omics representations. In the current implementation, these representations were used to characterize the simulated immune landscape and to support retrospective biological consistency assessment; they were not included as predictors of the Random Forest surrogate. The surrogate model instead estimated the precomputed composite Safety-by-Design score from seven formulation parameters and four latent immune-state variables. The overall objective of the framework is not to replace experimental validation but to provide a transparent, reproducible, and biologically interpretable computational platform capable of prioritizing candidate RNA-LNP formulations during the early stages of development, thereby reducing experimental burden while supporting model-informed drug development and New Approach Methodologies (NAMs).

The framework comprises a shared computational foundation and two complementary analytical branches.

Shared design and mechanistic-knowledge foundation:

1. generation of virtual patient cohorts;
2. construction of the RNA-LNP formulation design space;

3. integration of mechanistically derived immune descriptors from previously published UISS simulations;
4. construction of Mechanistically Informed Immune Signatures (MIIS).

Candidate-prioritization branch:

5. derivation of five efficacy- and safety-related response variables and computation of the composite Safety-by-Design score;
6. training and evaluation of an explainability-oriented Random Forest surrogate using seven formulation parameters and four latent immune-state variables;
7. Pareto-based multi-objective prioritization using the five individual response variables, followed by post hoc ranking of the non-dominated formulations according to the predicted Safety-by-Design score.

Biological-consistency branch:

8. generation of MIIS-constrained synthetic multi-omics representations and downstream immune-program scores;
9. retrospective biological consistency assessment against one independent human transcriptomic dataset;
10. computation of the Biological Program Concordance Index (BPCI).

The resulting architecture is partially branched rather than strictly sequential. Synthetic multi-omics variables are used to characterize the simulated immune landscape and support retrospective biological consistency assessment but are not included as predictors of the Random Forest surrogate. The BPCI assesses preservation of immune-program organization at the framework level and should not be interpreted as experimental validation of individual prioritized RNA-LNP formulations. The computational architecture comprises two complementary analytical branches connected by a shared mechanistic knowledge layer. In the candidate-prioritization branch, seven RNA-LNP formulation parameters and four latent immune-state variables are used to derive efficacy- and safety-related response variables and to train a Random Forest surrogate for prediction of the composite Safety-by-Design score. Pareto analysis is performed separately using five endpoint objectives, while the predicted composite score is retained for post hoc ranking of the non-dominated formulations. In the biological-consistency branch, UISS-derived Mechanistically Informed Immune Signatures (MIIS) constrain the generation of synthetic multi-omics representations and downstream immune-program scores, which are compared with independent human transcriptomic observations using the Biological Program Concordance Index. Synthetic multi-omics variables are therefore not used as predictors of the Random Forest model, and the BPCI evaluates framework-level biological consistency rather than the experimental validity of individual prioritized formulations.

2.2 Generation of virtual patient cohorts

Following established virtual-population modeling principles (25)(26) a synthetic cohort of 200 virtual patients was generated using a stochastic sampling procedure implemented in R. The cohort size was selected to provide computationally tractable coverage of heterogeneous immune-response scenarios and was not based on a formal statistical power calculation. Rather than representing identifiable individuals or digital replicas of a defined clinical population, each virtual patient was characterized by a combination of demographic and immunological attributes sampled from prespecified probability distributions selected to represent structured inter-individual heterogeneity (Table 3).

The demographic profile included age and biological sex. Age was sampled from a uniform distribution between 18 and 85 years and rounded to the nearest integer, whereas sex was assigned according to fixed probabilities of 52% female and 48% male. These proportions were used as population-level modeling assumptions rather than estimates calibrated against a specific clinical cohort.

Each virtual patient was further characterized by four latent immune-state variables representing baseline inflammatory status, overall immune responsiveness, innate immune reactivity, and adaptive immune potential. Baseline inflammatory status was represented as a normalized latent variable bounded between 0 and 1. In the absence of cohort-specific data suitable for direct calibration, a Beta (2,6) distribution (27) was selected a priori to encode a predominantly low-inflammatory virtual population while retaining inter-individual heterogeneity. This distribution has a mean of 0.25, a mode of approximately 0.17, and a standard deviation of approximately 0.14. Its right-skewed shape concentrates most virtual patients within the lower portion of the normalized range while preserving a smaller subgroup with moderately elevated baseline inflammatory status. The selected parameters should therefore be interpreted as explicit population-level modeling assumptions rather than as estimates derived from a specific clinical cohort.

Overall immune responsiveness, innate immune reactivity, and adaptive immune potential were independently sampled from normal distributions centered at 1.0 (immune responsiveness and adaptive potential: $\mu = 1.0$, $\sigma = 0.20$; innate reactivity: $\mu = 1.0$, $\sigma = 0.25$). To avoid implausibly low values, these variables were truncated at a minimum value of 0.2. Unless otherwise specified, the virtual-patient attributes were sampled independently; consequently, the current implementation does not reproduce demographic or immunological covariance structures observed in clinical populations.

For descriptive purposes, virtual patients were stratified into three inflammatory-risk classes according to their baseline inflammatory status (Low: <0.25 ; Intermediate: $0.25-0.50$; High: ≥ 0.50). This classification was used to visualize the distribution of simulated immune backgrounds and to support descriptive comparison of the resulting synthetic response states.

The virtual-patient profiles were computationally combined with the candidate RNA-LNP formulations to construct the patient–formulation design matrix. No new patient-specific UISS simulations were performed. Instead, the four latent immune-state variables contributed to generation of the synthetic response states and, together with the seven formulation parameters, constituted the 11 predictors used by the Random Forest surrogate. Age and sex were retained for cohort description but were not included as model predictors in the current implementation.

Virtual-patient generation was implemented in R version 4.6.0 (R Core Team, 2026) (28) using independent random sampling from prespecified parametric distributions. Age was generated using `stats::runif()` and rounded with `base::round()`, biological sex was assigned by probability-weighted sampling using `base::sample()`, baseline inflammatory status was generated using `stats::rbeta()`, and the three remaining latent immune-state variables were generated using `stats::rnorm()`. Lower truncation of the normally distributed immune variables was implemented using `base::pmax()`, while inflammatory-risk classes were assigned using `dplyr::case_when()` (dplyr version 1.2.1). No Latin hypercube sampling, resampling, or empirical population-calibration procedure was applied.

A fixed random seed (seed = 123) was used throughout the virtual-cohort generation procedure to support computational reproducibility.

Parameter	Value
Number of virtual patients	200
Random seed	123
Age	Uniform distribution, 18–85 years
Sex	Female 52%; Male 48%
Baseline inflammatory status	Beta (2,6)
Overall immune responsiveness	Normal ($\mu = 1.0$, $\sigma = 0.20$)
Innate immune reactivity	Normal ($\mu = 1.0$, $\sigma = 0.25$)
Adaptive immune potential	Normal ($\mu = 1.0$, $\sigma = 0.20$)
Lower truncation for immune variables	0.2
Inflammatory risk classes	Low (<0.25), Intermediate (0.25–0.50), High (≥ 0.50)

Table 3. Parameters used for the generation of the synthetic virtual patient cohort. Demographic and immunological variables, associated probability distributions, and parameter values used to simulate inter-individual variability across the virtual population

2.3 Construction of the RNA-LNP formulation design space

A virtual library of candidate RNA-LNP formulations was generated by systematically sampling the principal physicochemical properties known to influence RNA delivery, biodistribution,

intracellular trafficking, and immunological safety. The objective was to computationally explore a broad, yet biologically plausible formulation space representative of lipid nanoparticles commonly investigated for RNA therapeutics.

Each formulation was characterized by seven physicochemical descriptors, including nanoparticle size, surface charge (zeta potential), PEGylation density, ionizable lipid fraction, helper lipid fraction, cholesterol fraction, and the presence or absence of a targeting ligand (Table 4). These variables were selected because they are known to affect circulation time, cellular uptake, endosomal escape, lipid nanoparticle stability, tissue targeting, and innate immune recognition (9).

A total of 500 candidate RNA-LNP formulations were generated by independently sampling each formulation parameter from predefined biologically plausible distributions. The three major lipid components (ionizable lipid, helper lipid, and cholesterol) were subsequently normalized to ensure that their relative fractions summed to unity while preserving the independently sampled PEGylation density and targeting-ligand status. This normalization step generated chemically consistent formulations while maintaining broad exploration of the formulation design space(10,12). A fixed random seed (seed = 123) was used to ensure reproducibility of the formulation library. The resulting virtual design space enabled systematic evaluation of diverse RNA-LNP candidates before downstream mechanistic interpretation, synthetic multi-omics generation, and AI-based optimization.

Parameter	Distribution Sampling	/ Range	Unit
Number of formulations	Fixed	500	—
Random seed	Fixed	123	—
Particle size	Uniform	50–150	nm
Surface charge (zeta potential)	Uniform	–10 to +10	mV
PEGylation density	Uniform	0.10–0.50	mol%
Ionizable lipid fraction*	Uniform	0.35–0.60	relative fraction
Cholesterol fraction*	Uniform	0.25–0.45	relative fraction
Helper lipid fraction*	Uniform	0.05–0.20	relative fraction
Targeting ligand	Bernoulli distribution	0 = absent; 1 = present (P = 0.40)	binary

Table 4. Parameters used for the generation of the virtual RNA-LNP formulation library.

* After sampling, the ionizable lipid, cholesterol, and helper lipid fractions were normalized so that their combined proportion summed to one.

2.4 Mechanistic immune descriptors derived from previously published UISS simulations

The mechanistic component of the proposed framework is based on the Universal Immune System Simulator (UISS), an agent-based platform developed to reproduce the temporal dynamics of the human immune system under physiological and pathological conditions(29). UISS explicitly represents immune cells, cytokines, antibodies, and molecular mediators as autonomous interacting agents governed by biologically defined rules. Through these interactions, the model reproduces coordinated innate and adaptive immune behavior following antigenic or therapeutic stimulation and has been applied to infectious-disease and vaccination modeling (30).

Rather than conducting a new candidate-specific UISS simulation campaign, the present study repurposed temporal immune-response trajectories generated in our previous RNA-LNP optimization study (20) as a pre-existing mechanistic knowledge layer. UISS therefore did not directly simulate or optimize each virtual patient–formulation combination evaluated in the current workflow. Instead, previously generated activation profiles for major immune compartments, including epithelial/target cells, helper and cytotoxic T lymphocytes, B lymphocytes, memory T lymphocytes, and immunoglobulin production, were summarized into mechanistic descriptors used to construct the Mechanistically Informed Immune Signatures (MIIS) and constrain the downstream synthetic biological representations.

For each immune compartment, temporal activation trajectories are summarized into quantitative descriptors describing complementary aspects of the simulated immune response, including peak activation, mean activation level, activation persistence, time-to-peak, and residual activation at the end of the simulation. Rather than representing isolated simulation outputs, these descriptors provide a compact and biologically interpretable representation of immune dynamics that bridges the agent-based simulations with the downstream synthetic multi-omics generation module.

The extracted descriptors are subsequently integrated into the synthetic multi-omics workflow, where they guide the generation of coherent transcriptomic, cytokine, immune-cellular, and pathway-level representations used for synthetic immune-landscape characterization and retrospective biological consistency assessment.

2.5 Construction of Mechanistically Informed Immune Signatures

The compartment-level descriptors defined in Section 2.4 were organized according to predefined functional mappings to construct five higher-order Mechanistically Informed Immune Signatures (MIIS). This aggregation reduced the dimensionality of the UISS-derived outputs while retaining complementary information on response magnitude, cumulative activation, kinetics, persistence, and residual immune activity.

The resulting MIIS represented early innate activation, cellular adaptive immunity, humoral response, immune memory, and global immune persistence. Each signature integrated information from biologically related immune compartments and temporal descriptors, providing a compact and interpretable representation of the previously simulated immune dynamics.

The MIIS therefore constitute the intermediate mechanistic layer connecting the previously published UISS trajectories with the downstream synthetic multi-omics module. They were used as biological constraints for synthetic data generation rather than as direct predictors of the Random Forest surrogate or as objectives of the Pareto analysis.

2.6 Synthetic multi-omics generation

The five Mechanistically Informed Immune Signatures (MIIS) defined in Section 2.5 were used as mechanistic constraints for generating synthetic multi-omics representations comprising transcriptomic profiles, cytokine and chemokine signatures, immune-cell population states, and pathway-level activity scores. Rather than attempting to reconstruct individual molecular measurements, this module was designed to generate internally coherent synthetic profiles that preserved the functional organization of the underlying immune dynamics and were suitable for exploratory characterization and retrospective biological consistency assessment (31)(32).

Within this framework, MIIS constitute an intermediate functional representation connecting previously published mechanistic simulations with synthetic molecular data. The four synthetic modalities were generated to maintain explicit relationships with the corresponding mechanistic signatures while representing complementary aspects of the simulated immune response.

Synthetic transcriptomic profiles comprised 50 synthetic gene-level features organized into five biologically relevant immune programs, with 10 features assigned to each program: interferon response, inflammatory response, antigen presentation, adaptive immunity, and cellular stress response. Baseline expression values were sampled from a normal distribution with a mean of 6.0 and a standard deviation of 0.4 and were subsequently modulated according to the corresponding MIIS.

Each immune program was modulated by one or more MIIS according to predefined functional mappings. Early innate activation primarily influenced the interferon and inflammatory programs, whereas cellular adaptive immunity and humoral response predominantly modulated antigen-presentation and adaptive-immunity programs. Global immune persistence and the temporal properties retained by the mechanistic descriptors additionally contributed to the cellular stress-response program.

To introduce biological variability while preserving mechanistic consistency, Gaussian noise was added during the generation of transcriptomic, cytokine, and immune-cellular data. This procedure

produced internally consistent synthetic datasets that retained predefined relationships among molecular features without attempting to reproduce exact experimental transcript abundances.

Pathway-level scores were derived from the synthetic transcriptomic matrix rather than generated independently. For each sample s , the score of immune program p was calculated as the arithmetic mean of the expression values of the 10 synthetic gene-level features assigned to that program:

$$P_{p,s} = \frac{1}{10} \sum_{g \in G_p} X_{g,s}$$

where G_p denotes the set of 10 synthetic gene-level features assigned to program, p and $X_{g,s}$ denotes the synthetic expression value of feature g in sample s . This derivation ensured internal consistency between the transcriptomic and pathway-level layers and provided the functional representation used for retrospective comparison with the independent GSE171964 dataset.

The complete mapping between the MIIS and the downstream synthetic multi-omics layers is summarized in Table 5.

Mechanistic biological signature	Synthetic transcriptomic programs	Additional synthetic layers
Early innate activation	Interferon response; inflammatory response	Cytokine signatures
Cellular adaptive immunity	Adaptive immunity; antigen presentation	Immune-cell populations
Humoral response	Adaptive immunity	Immunoglobulin-associated features
Immune memory	Adaptive immunity	Memory-cell dynamics
Global immune persistence	Cellular stress response	Global pathway activity

Table 5. Mapping between mechanistic biological signatures and synthetic multi-omics components.

Although the current implementation uses predefined mappings between mechanistic biological signatures and immune programs, the modular architecture allows future integration of more sophisticated mapping strategies, including network-based propagation models, graph neural networks, or foundation-model embeddings.

The synthetic multi-omics representations were used for exploratory characterization of the simulated immune landscape and for retrospective biological consistency assessment. They were not included among the predictors of the Random Forest surrogate model, which was trained separately using formulation parameters and latent immune-state variables.

Unsupervised analysis of the synthetic transcriptomic landscape.

To characterize the global organization of the generated synthetic immune states, the transcriptomic matrix was transposed into a samples-by-genes representation and standardized

gene-wise across samples using Z-score transformation. Principal component analysis (PCA) was performed on the resulting standardized 50-gene matrix using the `prcomp` function in R. Because the matrix had already been standardized, no additional centering or scaling was applied during PCA. The first two principal components were retained for visualization, and the proportion of variance explained by each component was calculated from the corresponding eigenvalues. Synthetic samples were colored according to their inflammatory-risk class, and normal-distribution ellipses encompassing 80% of the observations within each class were included for descriptive visualization.

Uniform Manifold Approximation and Projection (UMAP) (33) was independently applied to the same standardized 50-gene expression matrix using the `umap` function from the R package `uwot`. UMAP was computed using 30 nearest neighbors, a minimum distance of 0.20, and Euclidean distance. Computation was restricted to a single thread, and a fixed random seed of 123 was used to support reproducibility. The resulting two-dimensional coordinates were visualized according to inflammatory-risk class to examine the local neighborhood structure and continuity of the MIIS-driven synthetic immune-response landscape. PCA and UMAP coordinates were used exclusively for exploratory visualization and were not included as predictive features in the downstream artificial intelligence models or in the computation of the Safety-by-Design score.

For pathway-level visualization, each of the five immune-program scores was independently z-standardized across samples. Samples were then ordered by agglomerative hierarchical clustering of the resulting five-dimensional score matrix, using Euclidean distance and Ward's minimum-variance method (`ward.D2`). The resulting dendrogram order was used to arrange the columns of the pathway-activity heatmap, and inflammatory-risk class was displayed as an accompanying sample-annotation strip.

2.7 Composite Safety-by-Design score

The composite Safety-by-Design (SbD) score was designed to summarize complementary efficacy- and safety-related response dimensions for each simulated patient–formulation record. In the candidate-prioritization branch, five response variables were considered: delivery efficiency, adaptive immune activation, innate immune activation, cytokine burden, and off-target immune activation. These variables were generated from the RNA-LNP formulation parameters and latent virtual-patient immune-state variables. UISS-derived descriptors, Mechanistically Informed Immune Signatures (MIIS), and synthetic multi-omics variables were not used directly in calculating the composite score. Patient-level scores were subsequently averaged across the virtual cohort to obtain formulation-level summary scores for ranking and visualization.

Before calculating the composite score, each endpoint x_j was normalized to the $[0,1]$ interval using the empirical minimum and maximum values estimated from the complete set of 10,000 simulated patient–formulation records:

$$x_j^* = \min\{1, \max[0, (x_j - m_j)/(M_j - m_j)]\}$$

where m_j and M_j denote the empirical minimum and maximum, respectively, of endpoint j in the complete simulated design space. The asterisk denotes the corresponding normalized value. This normalization produced five dimensionless variables: normalized delivery efficiency D^* , adaptive immune activation A^* , innate immune activation I^* , cytokine burden C^* , and off-target immune activation O^* .

Fixed coefficients were assigned a priori to the five normalized components: 0.30 for delivery efficiency, 0.20 for adaptive immune activation, 0.20 for innate immune activation, 0.15 for cytokine burden, and 0.15 for off-target immune activation. For the three endpoints to be minimized, directional transformation was performed by calculating their complements. The composite score was therefore defined as:

$$SbD = 0.30 D^* + 0.20 A^* + 0.20(1 - I^*) + 0.15(1 - C^*) + 0.15(1 - O^*)$$

The coefficients sum to one. Following normalization and directional transformation, each component ranged from 0 (least favorable) to 1 (most favorable), ensuring that higher composite scores consistently represented a more favorable efficacy–safety profile. Because the normalization parameters were derived from the synthetic design space, the resulting Safety-by-Design score should be interpreted as a relative prioritization metric rather than as an externally calibrated clinical safety scale.

Higher Safety-by-Design scores therefore identified candidate RNA-LNP formulations predicted to combine favorable delivery and adaptive immune activation with reduced innate activation, cytokine burden, and off-target immune activation. The composite score was modeled by the Random Forest surrogate and subsequently used for post hoc ranking and visualization of Pareto-optimal candidates; it was not included as a separate Pareto-dominance objective.

The components, weights, optimization directions, and empirical normalization parameters used to compute the composite Safety-by-Design score are summarized in Table 6.

Component	Weight	Contribution	Empirical minimum	Empirical maximum
Delivery efficiency	0.30	Positive	0.000	3.270
Adaptive immune activation	0.20	Positive	0.000	4.903
Innate immune activation	0.20	Negative	0.498	3.094
Cytokine burden	0.15	Negative	0.494	5.093
Off-target immune activation	0.15	Negative	0.000	6.139

Table 6. Components, weights, directions, and empirical normalization parameters of the composite Safety-by-Design score.

2.8 Explainable artificial intelligence framework

Predictive machine-learning models were trained to estimate the composite Safety-by-Design score from RNA-LNP formulation characteristics and virtual-patient immune attributes. The analysis dataset comprised 10,000 simulated patient–formulation records representing 200 virtual patients and 500 RNA-LNP formulations. Eleven predictors were included: baseline inflammatory status, overall immune responsiveness, innate immune reactivity, adaptive immune potential, nanoparticle size, zeta potential, PEGylation density, ionizable lipid fraction, cholesterol fraction, helper lipid fraction, and targeting-ligand status. Age, sex, UISS-derived descriptors, MIIS, and synthetic multi-omics variables were not used as model inputs.

Random Forest regression was implemented using the ranger R package (version 0.18.0) with 300 trees(34,35). Predictive performance was evaluated using five-fold cross-validation grouped by RNA-LNP formulation. All simulated records associated with the same formulation were assigned to a single fold, ensuring that no RNA-LNP candidate occurred in both the training and validation subsets. This formulation-grouped design was selected as the primary validation strategy because it evaluates generalization to previously unseen RNA-LNP candidates and prevents information leakage across repeated patient–formulation evaluations (36,37).

Out-of-fold predictions from the five validation folds were pooled to calculate the coefficient of determination:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$$

Model performance was summarized using R^2 , root mean squared error (RMSE) and mean absolute error (MAE). Fold-specific metrics and fold assignments were also retained. As a sensitivity analysis, the same procedure was repeated using five-fold cross-validation grouped by virtual patient, thereby evaluating generalization to previously unseen immune backgrounds. After validation, a final Random Forest model was fitted to all 10,000 records, and permutation-based feature importance was calculated to quantify the contribution of each predictor. A fixed random seed (seed = 123) was used for fold assignment and model fitting.

The main parameters of the Random Forest model and the cross-validation strategy are summarized in Table 7.

Parameter	Value
Machine-learning algorithm	Random Forest regression
Implementation	ranger package, version 0.18.0
Predictors	11 formulation and latent immune-state variables
Analysis dataset	10,000 sampled patient–formulation records; 500 LNP formulations; 200 virtual patients
Primary validation	Five-fold cross-validation grouped by RNA-LNP formulation

Training/validation observations per fold	Approximately 80%/20%; exact counts vary with formulation-group size
Number of trees	300 (num.trees = 300)
Random seed	123
Performance metrics	R ² , RMSE, MAE
Feature importance	Permutation importance from the final model

Table 7. Parameters of the explainable artificial intelligence module.

2.9 Multi-objective prioritization of RNA-LNP formulations

Candidate RNA-LNP formulations were prioritized using a Pareto-based multi-objective optimization strategy designed to identify formulations that maximize predicted delivery and adaptive immune activation while minimizing undesirable innate, cytokine, and off-target immune activation (16).

For each candidate formulation, mean values of the predicted Safety-by-Design score, delivery efficiency, adaptive immune activation, innate immune activation, cytokine burden, and off-target immune activation were computed across the virtual patient population. These aggregated formulation-level metrics were used for candidate assessment. Pareto dominance, however, was evaluated using five non-redundant endpoint objectives: delivery efficiency, adaptive immune activation, innate immune activation, cytokine burden, and off-target immune activation. The composite Safety-by-Design score was retained as a summary metric for subsequent ranking and visualization but was not included as a separate dominance objective.

The Pareto analysis simultaneously sought to:

- maximize mean delivery efficiency;
- maximize mean adaptive immune activation;
- minimize mean innate immune activation;
- minimize mean cytokine burden;
- minimize mean off-target immune activation.

The composite Safety-by-Design score was not treated as an additional Pareto objective because its constituent efficacy- and safety-related endpoints overlap with the five endpoint objectives above. Instead, it was used after Pareto filtering to rank non-dominated candidates and to support visualization of the delivery–SbD trade-off.

Candidate formulations were compared using Pareto-dominance analysis across five formulation-level objectives: maximization of delivery efficiency and adaptive immune activation, and

minimization of innate immune activation, cytokine burden, and off-target activation. Multi-objective analysis was selected because improvements in delivery and adaptive activation may be accompanied by increased innate or inflammatory responses, preventing identification of a single formulation that is optimal across all criteria. Pareto analysis preserves these competing trade-offs without requiring additional a priori weighting of the individual objectives(38). A formulation was classified as Pareto-optimal when no other candidate was at least as favorable across all five objectives and strictly more favorable in at least one. The resulting Pareto front therefore represents the set of non-dominated RNA-LNP candidates providing the most favorable compromises among efficacy- and safety-related endpoints. Pareto-optimal formulations were subsequently ranked according to their mean composite Safety-by-Design score across virtual patients, which served as a secondary decision criterion for identifying the leading candidates. The 20 highest-ranking Pareto-optimal formulations were retained for downstream interpretation and visualization.

The complete multi-objective optimization strategy is summarized in Table 8.

Component	Description
Optimization strategy	Pareto dominance analysis
Optimization unit	Mean formulation performance across the virtual patient population
Dominance objectives	Five endpoint objectives: delivery efficiency, adaptive immune activation, innate immune activation, cytokine burden, and off-target immune activation
Objectives to maximize	Mean delivery efficiency; mean adaptive immune activation
Objectives to minimize	Mean innate immune activation; mean cytokine burden; mean off-target immune activation
Safety-by-Design score	Not used as a separate dominance objective; retained for post hoc ranking and visualization
Pareto-optimal candidates	21 of 500 formulations
Ranking strategy	Pareto filtering followed by ranking according to mean Safety-by-Design score
Final output	Ranked list of Pareto-optimal RNA-LNP candidates

Table 8. Multi-objective optimization strategy.

2.10 Retrospective biological consistency validation

The biological plausibility of the immune programs generated by the framework was assessed through retrospective validation using the human single-cell RNA-seq dataset GSE171964, which includes peripheral blood mononuclear cells collected from healthy volunteers at multiple time points following BNT162b2 mRNA vaccination (21).

Single-cell count data were aggregated into pseudo-bulk transcriptomic profiles at the patient–time point level (39). For each pseudo-bulk profile, pathway-level signature scores were computed for five immune programs represented in the synthetic multi-omics layer: interferon response, inflammatory response, antigen presentation, adaptive immunity, and stress response.

Rather than treating validation as a direct quantitative prediction task, the analysis was designed as a biological consistency assessment. Observed immune activation was defined as the deviation of each pathway-level signature score from the corresponding Day 0 baseline:

$$\Delta S_{i,t} = S_{i,t} - S_{i,0}$$

where $S_{i,t}$ represents the observed score of immune-program (i) at time point (t), and $S_{i,0}$ represents the baseline value.

Synthetic predictions were compared with observed post-vaccination deviations from baseline at the immune-program level. The primary validation objective was to determine whether the framework preserved the relative hierarchy of vaccine-induced immune programs rather than reproducing the absolute magnitude of transcriptomic responses.

2.11 Biological Program Concordance Index

To quantify the preservation of immune-program organization between synthetic simulations and human transcriptomic observations, we defined a Biological Program Concordance Index. The BPCI integrates three complementary measures of biological agreement: rank concordance, direction agreement, and top-program agreement.

First, rank concordance was computed using Spearman correlation between predicted immune-program activation scores and observed baseline deviations. Spearman values were rescaled to the 0–1 interval:

$$R = \frac{\rho_s + 1}{2}$$

Second, direction agreement was computed as the proportion of immune programs showing concordant activation direction relative to baseline:

$$D = \frac{N_{\text{direction agreement}}}{N}$$

Third, top-program agreement was computed as the overlap between the three highest-ranked immune programs in the synthetic predictions and in the observed data:

$$T = \frac{|P_{\text{top3}} \cap O_{\text{top3}}|}{3}$$

The final BPCI was defined as a weighted composite score:

$$BPCI = 0.4R + 0.4D + 0.2T$$

The final score ranges from 0 to 1, where higher values indicate stronger biological concordance. This metric was used to evaluate whether the mechanistically informed synthetic framework reproduced the functional organization of vaccine-induced immune responses observed in independent human single-cell data.

2.12 Statistical analysis

Descriptive statistics were computed for all simulated variables and generated endpoints. Continuous variables were summarized using mean, standard deviation, median, and interquartile range where appropriate.

Machine-learning performance was evaluated using the coefficient of determination and root mean squared error. Associations between predicted and experimentally observed pathway activation scores were assessed using Spearman correlation coefficients. Rank-based concordance measures were used for biological program comparisons because the validation objective focused on preservation of immune-program hierarchy rather than absolute transcriptomic magnitude.

Unless otherwise specified, all analyses were performed in R(28) version 4.6.0 (R Core Team, 2026), and figures were generated using ggplot2 (40).

3. Results

3.1 Construction of the virtual patient–formulation design space

To enable systematic exploration of RNA therapeutic formulations under heterogeneous biological conditions, a virtual design space was generated by combining synthetic patient cohorts with a broad landscape of candidate RNA–lipid nanoparticle (RNA-LNP) formulations (Figure 2).

Figure 2. Construction of the virtual patient–formulation design space

Virtual patients and RNA–LNP formulation candidates generated for downstream safety-by-design analysis.

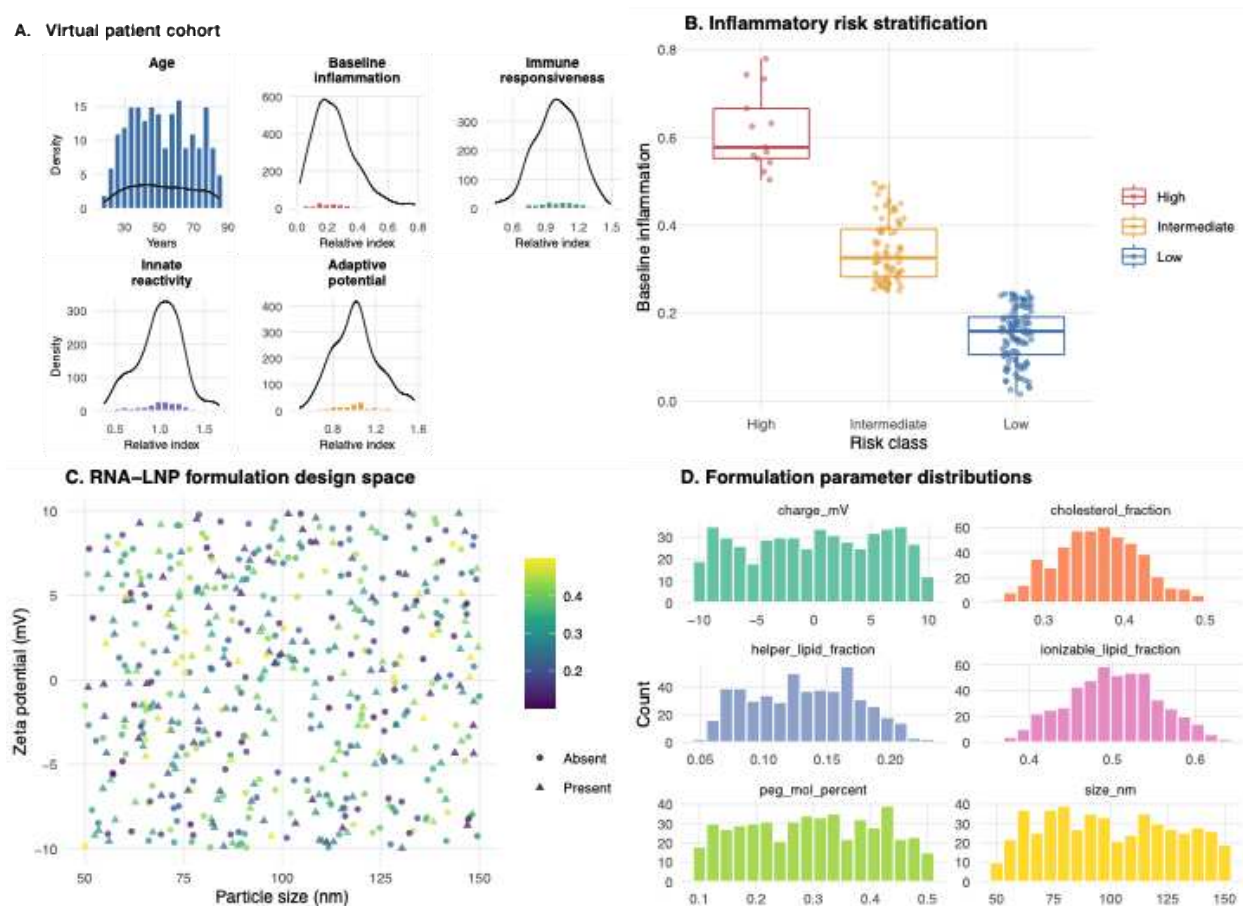


Figure 2. Construction of the virtual patient and RNA–LNP formulation design space. (A) Distributions of the demographic and latent immune-state variables used to represent inter-individual heterogeneity. (B) Descriptive stratification of virtual patients into low-, intermediate-, and high-inflammatory-risk classes according to baseline inflammatory status. (C) Sampled RNA–LNP formulation space defined by nanoparticle size, zeta potential, PEG molar fraction, and targeting-ligand status. (D) Marginal distributions of the principal physicochemical formulation variables, showing coverage of the prespecified parameter ranges.

The virtual patient cohort comprised individuals with heterogeneous demographic and immunological characteristics, including age, baseline inflammatory status, innate immune reactivity, adaptive immune potential, and overall immune responsiveness (Figure 2A). These variables were sampled from predefined probability distributions chosen to reproduce realistic inter-individual variability while maintaining biologically plausible parameter ranges. The resulting cohort covered a broad spectrum of immune phenotypes, providing a representative substrate for downstream mechanistically informed analyses.

To facilitate downstream safety analyses, patients were additionally stratified into three baseline inflammatory-risk classes (low, intermediate, and high) according to their simulated inflammatory profile (Figure 2B). This stratification enabled candidate formulations to be evaluated across individuals presenting distinct predispositions toward excessive inflammatory activation.

In parallel, a virtual formulation library was generated by sampling the multidimensional RNA-LNP design space (Figure 2C). Candidate formulations spanned a wide range of physicochemical properties, including particle size, surface charge (zeta potential), lipid composition, PEG-lipid content, and the presence or absence of targeting ligands. Uniform sampling across the predefined parameter ranges ensured broad coverage of the formulation landscape without introducing a priori optimization bias.

The distributions of the principal formulation variables are summarized in Figure 2D. Together, the virtual patient cohort and the RNA-LNP formulation library constitute the computational design space used throughout subsequent synthetic multi-omics generation, explainable AI analysis, and multi-objective optimization.

3.2 Mechanistically Informed Immune Signatures (MIIS)

Previously published UISS simulation outputs provided a high-dimensional mechanistic representation of immune-response dynamics, capturing the temporal behavior of multiple immune compartments and molecular mediators. Rather than comparing individual state variables directly with transcriptomic measurements, these outputs were transformed into biologically interpretable descriptors summarizing the major functional components of the immune response. This feature-engineering strategy generated Mechanistically Informed Immune Signatures (MIIS), providing a common biological representation for integrating mechanistic knowledge with synthetic and experimentally derived transcriptomic programs.

The descriptor library included profiles for epithelial/target cells, helper T cells, cytotoxic T cells, B cells, memory T cells, and immunoglobulin production. For each compartment, temporal trajectories from the published UISS simulations were summarized using complementary quantitative descriptors, including peak activation, mean activity, area under the response curve (AUC), time-to-peak, and final activation state. These descriptors capture both response magnitude and persistence without implying that new UISS trajectories were generated for every virtual patient–formulation pair in the present study.

The extracted descriptors were aggregated into five higher-order Mechanistically Informed Immune Signatures: early innate activation, cellular adaptive immunity, humoral response, immune memory, and global immune persistence. These five MIIS summarize complementary properties of the previously simulated immune trajectories and are distinct from the downstream transcriptomic comparison programs. Through the predefined mapping described in Methods, the

MIIS subsequently informed the generation of interferon-response, inflammatory-response, antigen-presentation, adaptive-immunity, and stress-response scores.

The resulting MIIS constitute the mechanistic knowledge layer used throughout the remainder of the study. The downstream synthetic immune-program scores provide the common functional space used for comparison with human transcriptomic data; consequently, retrospective validation evaluates whether MIIS-informed synthetic profiles preserve biological program organization rather than testing the five MIIS directly against gene-expression measurements.

The process used to derive the Mechanistically Informed Immune Signatures (MIIS) from the previously published UISS simulations is illustrated in Figure 3.

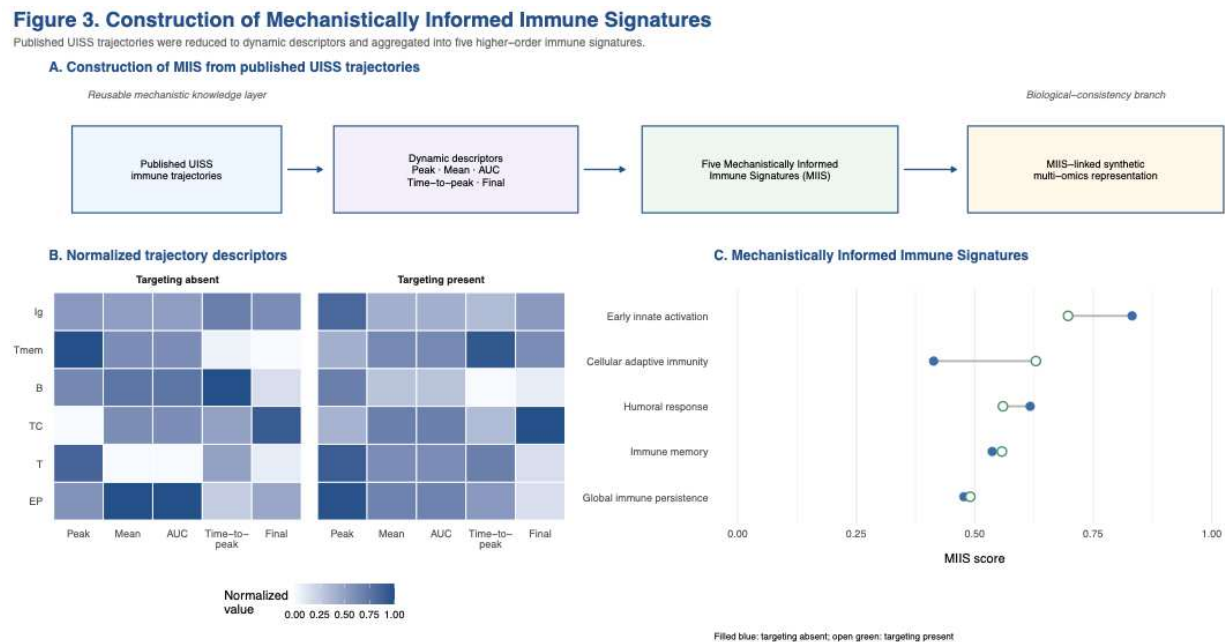


Figure 3. Construction of Mechanistically Informed Immune Signatures (MIIS) from previously published UISS trajectories. (A) Published temporal activation profiles from six immune compartments were summarized into five dynamic descriptors and aggregated into higher-order MIIS that constrain the synthetic multi-omics layer. (B) Normalized peak, mean, area-under-the-curve, time-to-peak, and final-activation descriptors for targeting-absent and targeting-present conditions. (C) MIIS scores for the two targeting conditions. Filled blue points indicate targeting absent; open green points indicate targeting present.

3.3 Synthetic multi-omics landscape

The MIIS were subsequently translated into synthetic multi-omics profiles to evaluate whether the framework could generate biologically structured immune-response landscapes rather than isolated molecular features. The resulting synthetic dataset integrated transcriptomic profiles with pathway-level immune-program scores representing interferon response, inflammatory response, antigen presentation, adaptive immunity, and stress response.

Figure 4. Organization of the synthetic multi-omics landscape

Synthetic molecular profiles preserve biologically coherent immune-program organization across virtual patients and risk classes.

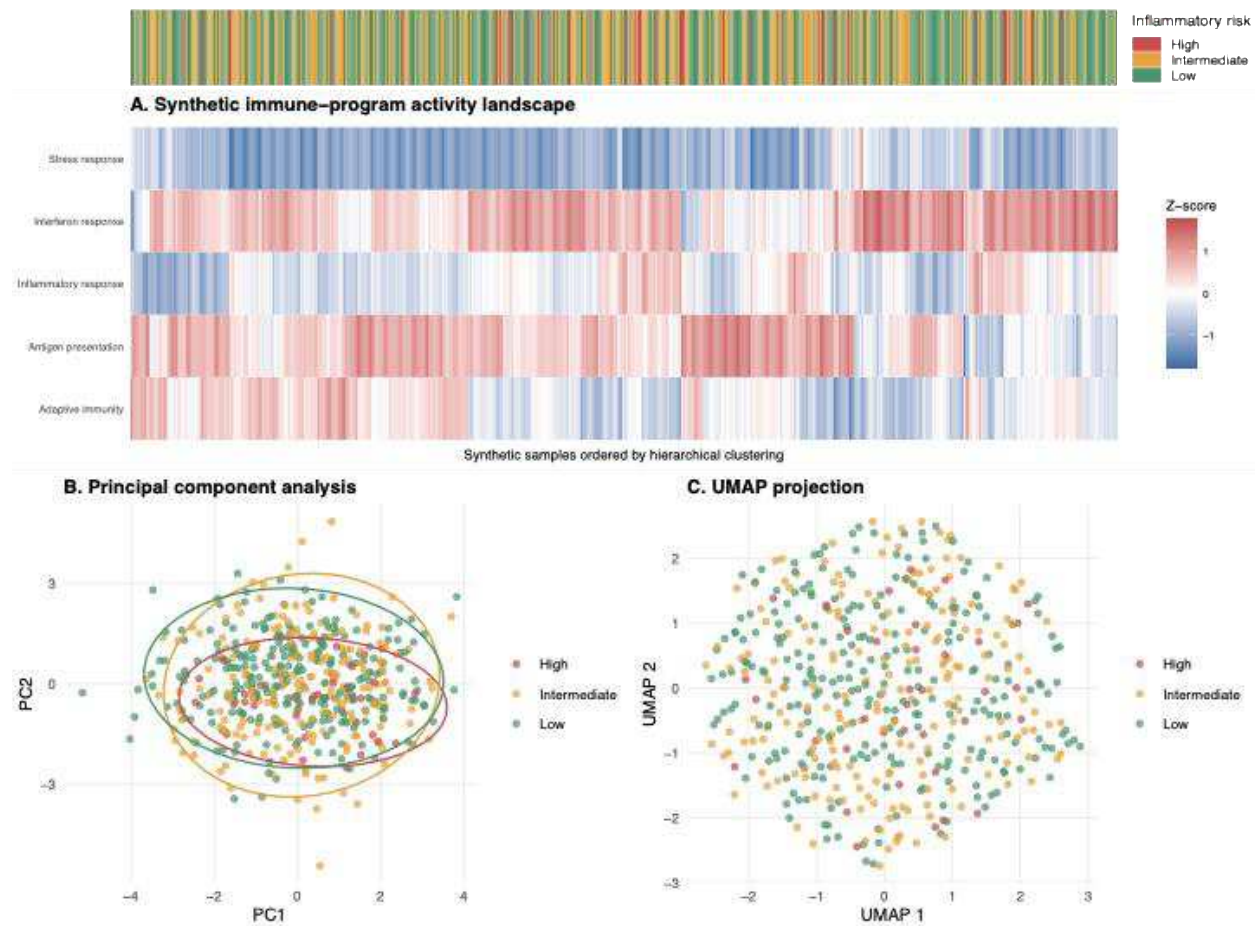


Figure 4. Organization of the MIIS-constrained synthetic transcriptomic landscape. (A) Heatmap of standardized activity scores for five downstream immune programs across synthetic samples ordered by hierarchical clustering; the upper annotation indicates inflammatory-risk class. (B) Principal component analysis of the standardized 50-feature synthetic transcriptomic matrix, colored by inflammatory-risk class. Ellipses encompass 80% of each class under a multivariate normal approximation. (C) UMAP projection of the same matrix, illustrating a continuous synthetic immune-state landscape and local neighborhood structure. These visualizations are exploratory and were not used as predictors of the Random Forest surrogate.

Hierarchical clustering of pathway-level scores revealed a structured synthetic immune-program landscape, with coordinated patterns of activation across biologically related immune programs (Figure 4A). This organization indicates that the generated profiles preserved coherent relationships among immune processes rather than reflecting random or independent variation.

Dimensionality-reduction analyses further supported the presence of structured biological variability. Principal component analysis showed that synthetic transcriptomic profiles occupied a continuous multidimensional space, with the first two components explaining a substantial fraction of the total variance (Figure 4B). Samples were distributed across inflammatory-risk classes without complete artificial separation, suggesting that the synthetic generator captured gradual immune-state variation across virtual patients.

Consistently, UMAP projection confirmed the presence of a continuous immune-response manifold, with local structure reflecting heterogeneous inflammatory-risk profiles (Figure 4C). Overall, these results indicate that the MIIS-driven synthetic multi-omics generation step produced biologically coherent and heterogeneous immune landscapes suitable for biological characterization and retrospective consistency assessment.

These analyses characterize the biological organization of the MIIS-constrained synthetic profiles. The synthetic multi-omics variables were not subsequently used as Random Forest predictors; the surrogate-modeling analysis described below represents a complementary candidate-prioritization branch based on formulation and latent immune-state variables.

3.4 Explainable AI modeling of the Safety-by-Design score

Under the primary formulation-grouped cross-validation scheme, the Random Forest model explained 88.7% of the variance in out-of-fold Safety-by-Design scores ($R^2 = 0.887$), with an RMSE of 0.055 and an MAE of 0.044 (Figure 5A). Performance was consistent across the five folds, with a mean fold-specific R^2 of 0.886 (SD = 0.013) and a mean RMSE of 0.055 (SD = 0.002). These results indicate that the surrogate model retained strong predictive performance when evaluated on RNA-LNP formulations that were entirely absent from the corresponding training sets. In the patient-grouped sensitivity analysis, the model achieved a pooled R^2 of 0.973 and an RMSE of 0.027. The higher performance observed under patient-grouped validation is expected because RNA-LNP formulations remain represented in the training data; this analysis

therefore assesses generalization across immune backgrounds rather than extrapolation to unseen formulations.

Figure 5. Explainable AI prediction of the Safety-by-Design score

Random Forest performance was evaluated using formulation-grouped cross-validation to prevent information leakage across RNA-LNP candidates.

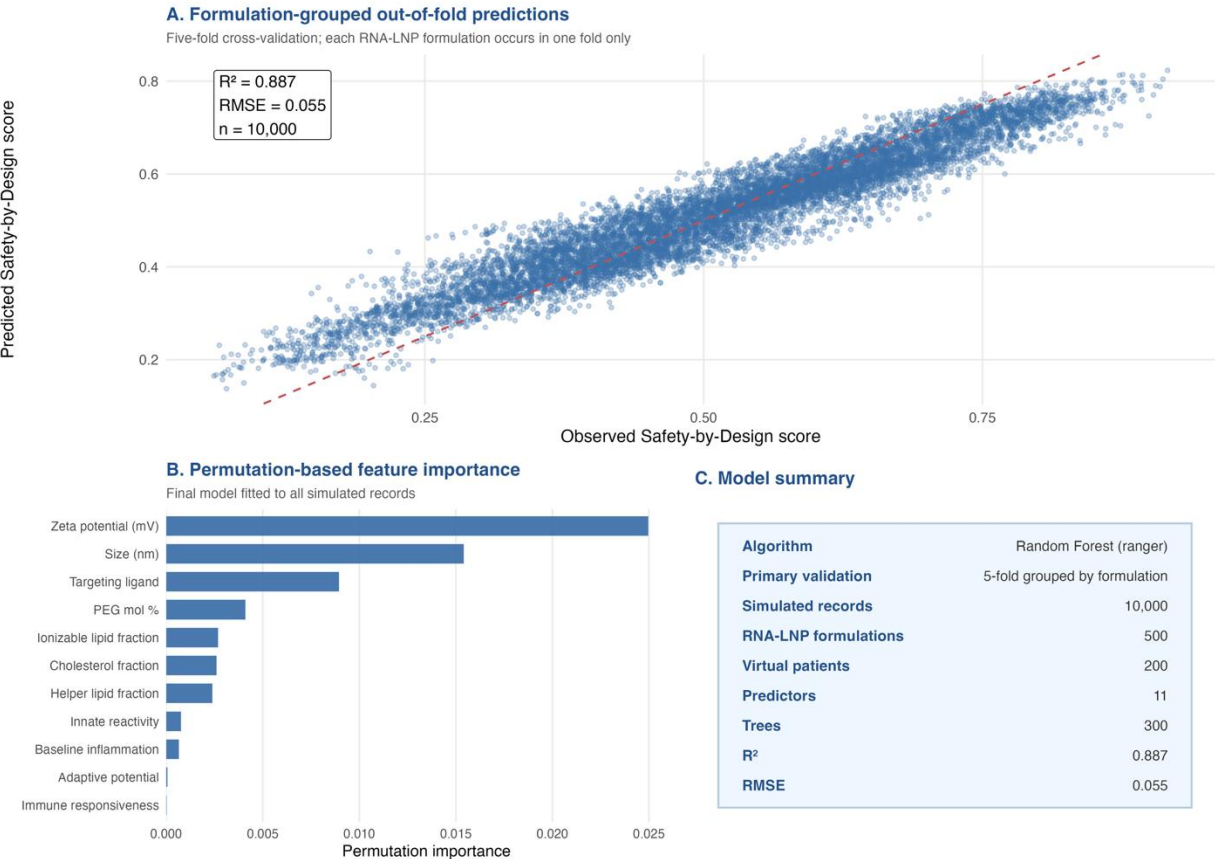


Figure 5. Random Forest prediction and interpretation of the composite Safety-by-Design score. (A) Observed versus out-of-fold predicted scores under five-fold cross-validation grouped by RNA-LNP formulation. The pooled out-of-fold performance across 10,000 simulated patient–formulation records was $R^2 = 0.887$, $RMSE = 0.055$, and $MAE = 0.044$. (B) Permutation-based predictor importance from the final model fitted to all records. (C) Summary of the model architecture and primary validation design. The final Random Forest model used 300 trees and 11 formulation and latent immune-state predictors.

Permutation-based feature importance from the final model identified zeta potential and nanoparticle size as the dominant predictors, followed by targeting-ligand status and PEGylation density (Figure 5B). The remaining lipid-composition variables and virtual-patient immune attributes contributed more modestly to prediction. Together, these findings indicate that the composite score can be approximated by an interpretable surrogate model while preserving a strict separation between candidate formulations during primary performance assessment. These results show that the surrogate model captures non-linear associations between RNA-LNP physicochemical characteristics, heterogeneous latent immune states, and the composite SbD score while retaining an interpretable predictor structure.

3.5 Pareto-based prioritization of RNA-LNP candidates

Multi-objective optimization identified a restricted subset of RNA-LNP formulations simultaneously maximizing mean delivery efficiency and adaptive immune activation while minimizing mean innate immune activation, cytokine burden, and off-target immune activation. Among the 500 candidate formulations explored, 21 were classified as Pareto-optimal, indicating that no alternative formulation improved all five endpoint objectives simultaneously (Figure 6A).

Figure 6. Pareto-based prioritization of RNA-LNP formulations

Multi-objective optimization identifies lead RNA-LNP candidates balancing delivery efficiency and Safety-by-Design performance.

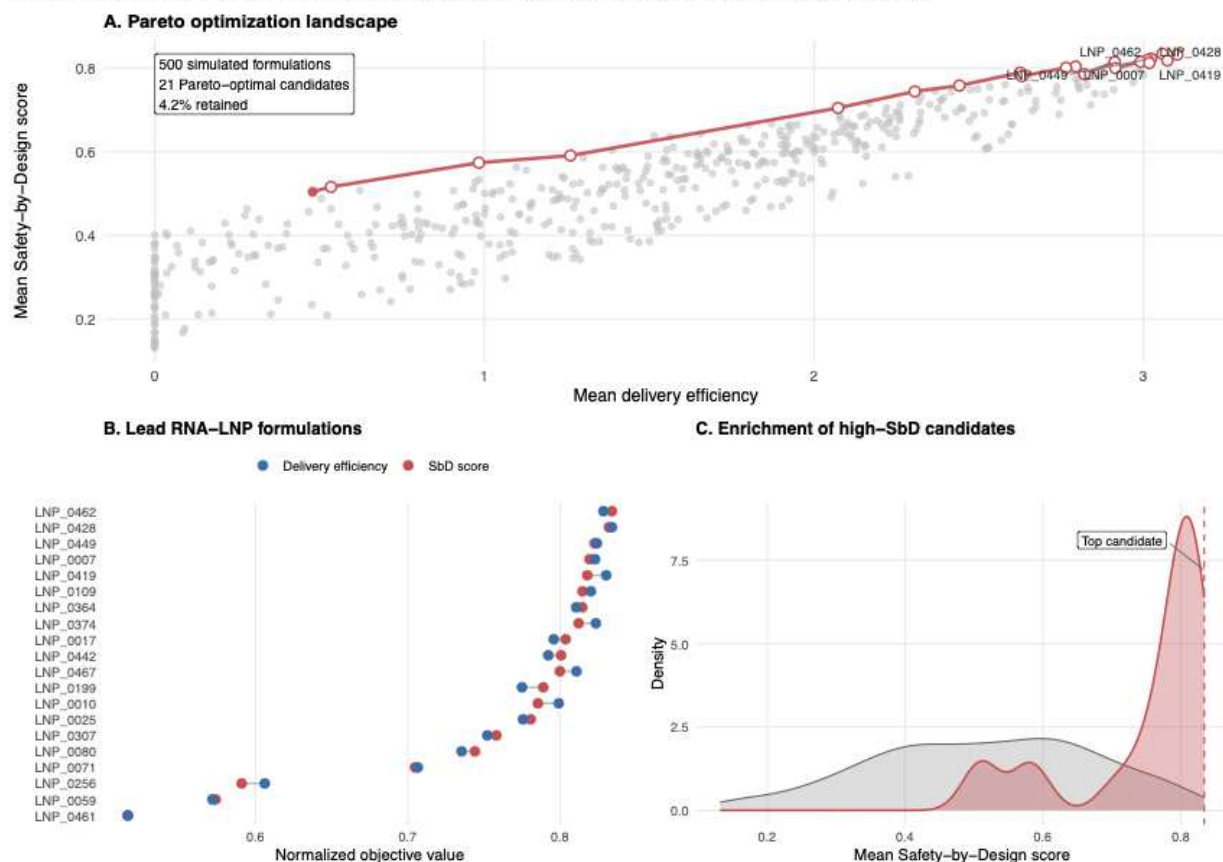


Figure 6. Pareto-based prioritization of non-dominated RNA-LNP formulations. (A) Two-dimensional visualization of all 500 formulations by mean delivery efficiency and mean Safety-by-Design score. Red points denote the 21 candidates classified as Pareto-optimal using the five endpoint objectives; the connecting line is included only to aid visualization of their projection in this two-dimensional space. (B) Normalized delivery efficiency and mean Safety-by-Design score for the highest-ranking Pareto-optimal candidates. (C) Distribution of mean Safety-by-Design scores for all formulations and for the Pareto-optimal subset, showing enrichment of higher-scoring candidates after five-objective Pareto filtering and score-based post hoc ranking.

The 21 non-dominated candidates defined the multi-objective trade-off surface, while their subsequent ranking according to mean Safety-by-Design score identified a smaller set of formulations combining high predicted delivery with favorable overall safety profiles (Figure 6A–B). The distribution of Safety-by-Design scores showed that Pareto-optimal formulations were enriched in the upper tail of the optimization landscape (Figure 6C), while retaining diversity across the five endpoint objectives.

3.6 Retrospective biological consistency assessment of MIIS-informed immune programs

To assess the biological plausibility of the synthetic multi-omics representations generated under Mechanistically Informed Immune Signature (MIIS) constraints, we retrospectively compared five downstream immune-program scores with the independent human single-cell RNA-seq dataset GSE171964, which describes the temporal immune response following BNT162b2 vaccination. The comparison therefore evaluates the biological organization emerging from the MIIS-informed framework rather than a direct one-to-one correspondence between the five MIIS and individual transcriptomic measurements.

Rather than treating validation as an absolute transcriptomic prediction task, pathway activation at each post-vaccination time point was expressed as the deviation from the corresponding Day 0 baseline. This transformation isolated vaccine-associated changes from constitutive immune activity and provided a common functional space for comparing synthetic and experimental responses. Inspection of the observed transcriptomic trajectories revealed a coherent temporal progression of the five downstream comparison programs (Figure 7A). Interferon signaling showed the strongest early perturbation, followed by antigen-presentation and inflammatory programs, whereas adaptive immunity developed more gradually and the stress-response program remained comparatively modest. This pattern is consistent with the expected transition from early innate activation toward a progressively established adaptive response after mRNA vaccination.

Figure 7. BPCI-based assessment of systems-level immune-program consistency

Retrospective comparison between MIIS-informed synthetic profiles and human vaccination transcriptomics.

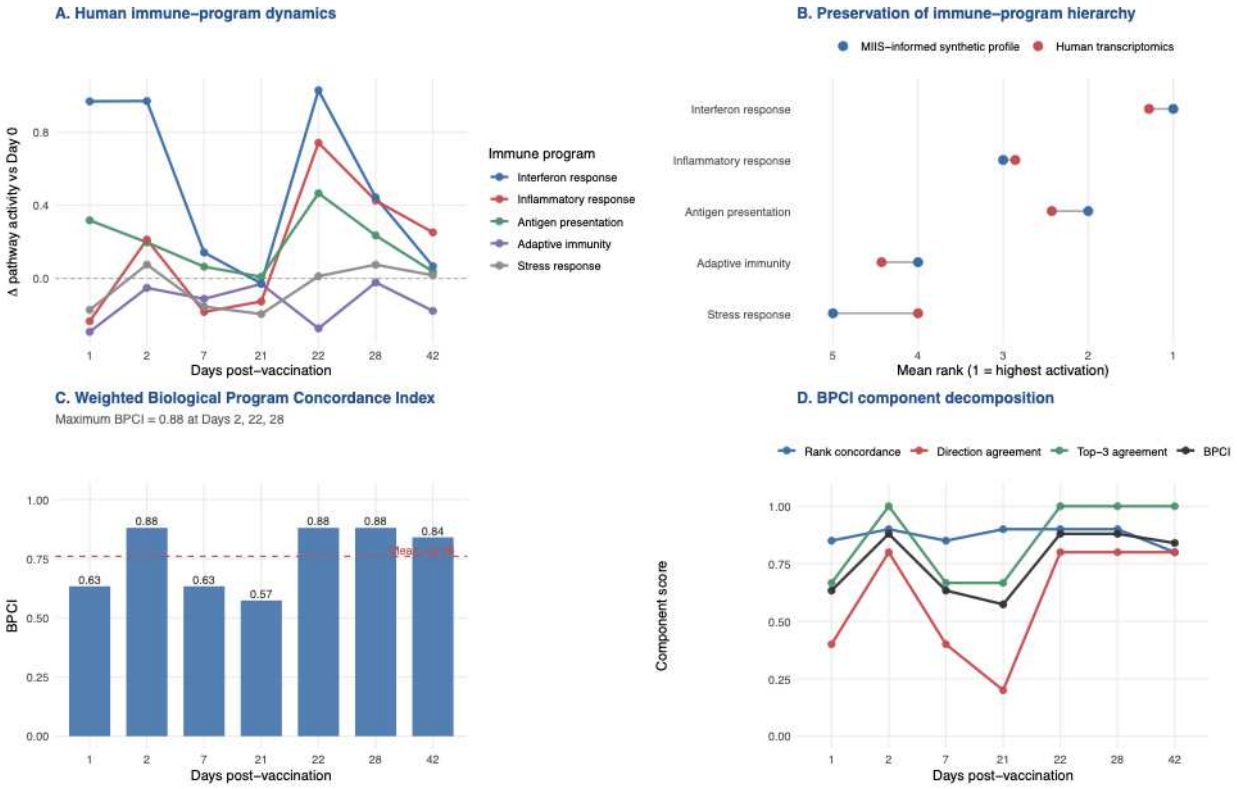


Figure 7. BPCI-based assessment of systems-level immune-program consistency. A) Mean deviations of five immune-program scores from the corresponding Day 0 baseline in GSE171964. (B) Comparison of mean immune-program ranks derived from the MIIS-informed synthetic profiles and human transcriptomic observations. (C) Weighted Biological Program Concordance Index (BPCI) across seven post-vaccination time points. Values ranged from 0.57 to 0.88 (mean = 0.76), with the highest concordance observed at Days 2, 22, and 28; the dashed line indicates the mean BPCI. (D) Temporal decomposition of the BPCI into rescaled rank concordance, activation-direction agreement, and top-three program overlap. BPCI was calculated as $(0.4R + 0.4D + 0.2T)$, where (R) , (D) , and (T) denote rescaled rank concordance, direction agreement, and top-three program overlap, respectively.

The relative hierarchy of the five immune programs was broadly preserved between the MIIS-informed synthetic profiles and human transcriptomic observations (Figure 7B). Spearman rank concordance across the post-vaccination trajectory ranged from 0.60 to 0.80, corresponding to rescaled rank-component values of 0.80–0.90. Agreement among the three highest-ranked programs ranged from 0.67 to 1.00, whereas activation-direction agreement showed greater temporal variability, ranging from 0.20 to 0.80.

To quantify this systems-level agreement, we applied the Biological Program Concordance Index (BPCI), which integrates three complementary components: preservation of immune-program ranking, agreement in activation direction, and overlap among the three highest-ranked programs. The resulting composite score ranges from 0, indicating no concordance, to 1, indicating complete concordance. Using the prespecified weighted formulation, BPCI values ranged from 0.57 to 0.88 across the seven post-vaccination time points, with a mean of 0.76 (SD = 0.14; Figure 7C). The highest concordance was observed at Days 2, 22, and 28 (BPCI = 0.88), whereas the lowest value occurred at Day 21 (BPCI = 0.57). Decomposition of the index showed that rank concordance remained consistently high, while temporal variation in the overall BPCI was driven mainly by activation-direction agreement (Figure 7D).

Overall, these results indicate that the MIIS-informed synthetic framework preserves key aspects of the systems-level organization of human vaccine-induced immunity despite not being trained directly on experimental transcriptomic measurements. This retrospective assessment supports interpretation of the framework as a functional representation of coordinated immune behavior rather than as a predictor of exact experimental gene-expression values.

4. Discussion

The present study introduces a proof-of-concept Safety-by-Design architecture for the early computational prioritization of RNA-LNP formulations, structured around two complementary but methodologically distinct branches (Figure 1). In the candidate-prioritization branch, formulation and virtual-patient variables are used to calculate a composite Safety-by-Design score, train a Random Forest surrogate, and identify non-dominated formulations through Pareto-based multi-objective analysis. In parallel, the biological-consistency branch transfers mechanistic immune knowledge from previously published UISS simulations into MIIS-constrained synthetic multi-omics representations and evaluates their systems-level concordance with independent human transcriptomic data. This separation enables efficacy–safety prioritization and biological plausibility assessment to contribute complementary forms of computational evidence without conflating their inputs, objectives, or inferential roles.

Within the latter branch, UISS acts as a pre-existing mechanistic knowledge source rather than as a newly executed simulation engine. Temporal immune trajectories generated in previously published UISS simulations were summarized into dynamic descriptors and subsequently condensed into five Mechanistically Informed Immune Signatures (MIIS): early innate activation, cellular adaptive immunity, humoral response, immune memory, and global immune persistence (20), (41). These signatures constrain the generation of synthetic molecular profiles, illustrating how previously generated mechanistic knowledge can be transferred into downstream data-driven

analyses. However, this approach does not constitute candidate-specific UISS simulation of every virtual patient–formulation pair evaluated in the present study.

The five mechanistic MIIS should be distinguished from the five downstream transcriptomic comparison programs—interferon response, inflammatory response, antigen presentation, adaptive immunity, and stress response. The former provide the mechanistic constraints used during synthetic multi-omics generation, whereas the latter constitute the shared functional space used for exploratory characterization and comparison with human transcriptomic observations. As formally represented in Figure 1, neither the MIIS nor the synthetic multi-omics variables were used as predictors of the Random Forest surrogate or as objectives of the Pareto analysis.

Figure 8 provides a complementary conceptual synthesis of the framework. By placing MIIS at the center, the figure illustrates how previously published mechanistic immune trajectories are translated into synthetic biological representations, interpreted alongside the parallel candidate-prioritization branch, and compared with independent human transcriptomic observations. The connections shown in Figure 8 therefore represent complementary analytical relationships rather than a single sequential predictor flow.

Figure 8. MIIS provide an interpretable bridge between mechanistic simulations and human transcriptomics

MIIS provide a shared interpretive context linking published UISS trajectories, synthetic multi-omics, parallel candidate prioritization, and independent human transcriptomics.

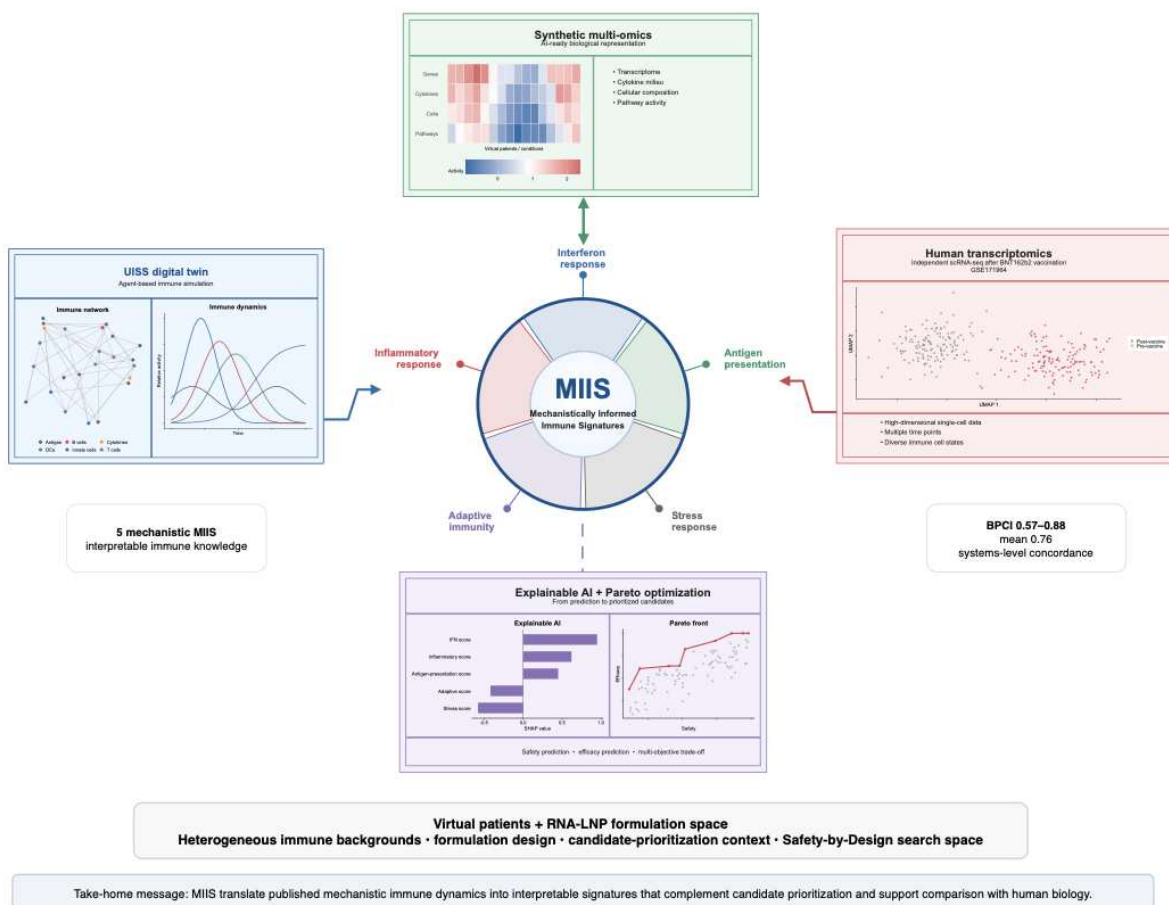


Figure 8. Mechanistically Informed Immune Signatures provide an interpretable bridge across the proposed computational framework. Previously published UISS immune trajectories are summarized into five MIIS that constrain the generation of synthetic multi-omics representations and support systems-level comparison with independent human transcriptomic data. The explainable artificial intelligence and Pareto-optimization module represents the parallel candidate-prioritization branch and is shown to contextualize its complementary role within the overall framework; MIIS and synthetic multi-omics variables are not used directly as Random Forest predictors or Pareto-dominance objectives. Together, the two analytical branches provide complementary evidence for computational prioritization and biological consistency assessment.

The surrogate-modeling results indicate that the composite Safety-by-Design score can be approximated accurately within the simulated design space when formulation-level information leakage is controlled. Five-fold cross-validation grouped by RNA-LNP formulation yielded a mean R^2 of 0.886 (SD = 0.013) and a mean RMSE of 0.055 (SD = 0.002). The subsequent five-objective Pareto analysis retained 21 of the 500 candidate formulations. These findings support the computational utility of the workflow for candidate prioritization but should be interpreted as internal performance within a synthetic, assumption-constrained environment rather than as evidence of clinical predictive accuracy.

Retrospective analysis using GSE171964 provided an independent systems-level biological consistency assessment (21). The MIIS-informed synthetic profiles preserved the broad hierarchy of vaccine-associated immune programs, with the interferon response emerging as the dominant program and antigen-presentation and inflammatory responses occupying intermediate positions. Using the prespecified weighted formulation, BPCI values ranged from 0.57 to 0.88 and reached their highest values at Days 2, 22, and 28. This degree of agreement supports biological plausibility at the immune-program level without implying direct reconstruction or prediction of experimental transcript abundances.

A second methodological contribution is the Biological Program Concordance Index. Rather than relying on a single global correlation coefficient, the BPCI integrates immune-program rank concordance, activation-direction agreement, and overlap among the three highest-ranked programs. The index therefore summarizes complementary aspects of biological organization, although its weighting scheme and interpretation remain methodological choices that require evaluation in additional datasets and biological contexts.

More broadly, the results suggest that validation of mechanistically informed computational frameworks may benefit from complementing transcript-level metrics with assessments of coordinated biological organization. Biological systems are dynamic, stochastic, and heterogeneous, and different molecular trajectories may converge on similar functional states. Preservation of systems-level organization can therefore provide information not captured by direct gene-wise comparisons, consistent with the principles of systems biology (41). It should be regarded as a complementary validation perspective rather than a replacement for molecular and experimental validation.

The virtual patient component extends the framework by representing heterogeneous latent immune backgrounds during candidate evaluation. Combined with explainability-oriented

modeling and Pareto-based prioritization, this structure permits exploration of trade-offs across multiple formulations and immune contexts. Nevertheless, the virtual cohort reflects prespecified probability distributions rather than a calibrated clinical population, and the resulting heterogeneity should be interpreted as structured scenario exploration.

The framework is aligned with the objectives of New Approach Methodologies and Model-Informed Drug Development (17,18) by providing a transparent computational layer for hypothesis generation and early candidate prioritization. Its role is to complement laboratory studies and help focus experimental resources; it cannot establish formulation safety, efficacy, biodistribution, or clinical benefit independently.

The architecture is intentionally modular and could incorporate alternative mechanistic simulators, richer multi-omics generators, other artificial-intelligence models, or additional optimization criteria. Future work should dynamically couple candidate-specific mechanistic simulations with experimentally calibrated molecular data, prospectively test prioritized formulations, and evaluate MIIS and BPCI across additional RNA platforms, cohorts, dosing regimens, and disease contexts. Such extensions will determine whether the proposed representation and validation strategy can generalize beyond the current proof-of-concept setting.

5. Limitations

Several limitations define the scope of this proof-of-concept study.

First, MIIS were derived from previously published UISS simulations rather than from candidate-specific simulations of every virtual patient–formulation pair. Formulation-dependent feedback on mechanistic immune dynamics is therefore represented indirectly. Future implementations should dynamically couple candidate-specific UISS simulations with the optimization pipeline.

Second, the synthetic multi-omics generator is deliberately simplified and assumption driven. Its 50-gene transcriptomic layer, predefined immune programs, and imposed stochastic variation should be interpreted as functional synthetic representations rather than predictions of real gene expression, cell-type-specific responses, or full biological complexity.

Third, the virtual patient and formulation libraries were generated from prespecified distributions and parameter ranges rather than calibrated against clinical cohorts or experimental formulation databases. Moreover, empirical min–max normalization makes the Safety-by-Design score relative to the simulated design space rather than to an externally calibrated clinical scale. Although normalization parameters were fixed and reused across downstream subsets, external calibration and sensitivity analyses will be required to assess robustness and applicability beyond the current ranges.

Fourth, retrospective biological consistency assessment relied on a single human scRNA-seq dataset describing responses to BNT162b2 vaccination. The BPCI weighting scheme has not yet been externally validated; its generalizability should therefore be evaluated across additional cohorts, RNA platforms, disease settings, and experimental datasets.

Finally, the framework focuses on immune-related efficacy and safety surrogates and does not explicitly represent pharmacokinetics, biodistribution, dose–exposure relationships, repeated dosing, manufacturing constraints, or patient-specific genetic determinants. The strong Random Forest performance reflects internal prediction within a synthetic design space and does not establish prospective or clinical validity. Accordingly, the framework is intended for hypothesis generation and early candidate prioritization, while experimental confirmation remains essential.

6. Conclusions

We developed a proof-of-concept Safety-by-Design framework for early computational prioritization of RNA-LNP formulations by integrating synthetic virtual patients, a virtual

formulation design space, mechanistic descriptors from previously published UISS simulations, MIIS-constrained synthetic multi-omics, explainability-oriented artificial intelligence, Pareto-based prioritization, and retrospective biological consistency assessment.

MIIS provide the central interface between the pre-existing mechanistic knowledge layer and downstream synthetic molecular representations. The five mechanistic signatures are distinct from the five transcriptomic comparison programs that they constrain, allowing mechanistic interpretability to be maintained while generating biologically interpretable synthetic data suitable for exploratory and validation analyses.

Within the simulated design space, grouped Random Forest validation showed strong surrogate-model performance and Pareto analysis identified a restricted set of non-dominated RNA-LNP candidates. These results demonstrate the utility of the architecture for structured computational screening while remaining conditional on the assumptions used to generate the virtual cohort, formulations, and response variables.

Retrospective comparison with independent human single-cell transcriptomic data showed preservation of the relative organization of vaccine-associated immune programs, with BPCI values ranging from 0.57 to 0.88 across the analysed time points (mean = 0.76) and the highest concordance observed at Days 2, 22, and 28. The BPCI therefore offers a transparent systems-level summary of biological consistency, but its broader validity must be established through external and prospective evaluation.

The framework is not intended to replace experimental studies or to provide direct clinical predictions. Its immediate value lies in organizing mechanistic knowledge, synthetic data generation, explainable modeling, and multi-objective prioritization within a reproducible workflow that can guide hypothesis generation and focus subsequent laboratory testing.

Future work should incorporate candidate-specific mechanistic simulations, empirically calibrated virtual populations and formulation priors, richer molecular representations, and validation across additional datasets and RNA platforms. These developments will determine the translational value of MIIS and the BPCI within model-informed drug development and New Approach Methodologies.

References

1. Sahin U, Karikó K, Türeci Ö. mRNA-based therapeutics — developing a new class of drugs. *Nat Rev Drug Discov*. 2014 Oct 19;13(10):759–80. doi:10.1038/nrd4278
2. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine*. 2020 Dec 31;383(27):2603–15. doi:10.1056/NEJMoa2034577
3. Walsh EE, Frenck RW, Falsey AR, Kitchin N, Absalon J, Gurtman A, et al. Safety and Immunogenicity of Two RNA-Based Covid-19 Vaccine Candidates. *New England Journal of Medicine*. 2020 Dec 17;383(25):2439–50. doi:10.1056/NEJMoa2027906
4. Pardi N, Hogan MJ, Porter FW, Weissman D. mRNA vaccines — a new era in vaccinology. *Nat Rev Drug Discov*. 2018 Apr 12;17(4):261–79. doi:10.1038/nrd.2017.243
5. Kulkarni JA, Witzigmann D, Thomson SB, Chen S, Leavitt BR, Cullis PR, et al. The current landscape of nucleic acid therapeutics. *Nat Nanotechnol*. 2021 Jun 31;16(6):630–43. doi:10.1038/s41565-021-00898-0
6. Qin S, Tang X, Chen Y, Chen K, Fan N, Xiao W, et al. mRNA-based therapeutics: powerful and versatile tools to combat diseases. *Signal Transduct Target Ther*. 2022 May 21;7(1):166. doi:10.1038/s41392-022-01007-w
7. Damase TR, Sukhovshin R, Boada C, Taraballi F, Pettigrew RI, Cooke JP. The Limitless Future of RNA Therapeutics. *Front Bioeng Biotechnol*. 2021 Mar 18;9. doi:10.3389/fbioe.2021.628137
8. Habrant D, Peuziat P, Colombani T, Dallet L, Gehin J, Goudeau E, et al. Design of Ionizable Lipids To Overcome the Limiting Step of Endosomal Escape: Application in the Intracellular Delivery of mRNA, DNA, and siRNA. *J Med Chem*. 2016 Apr 14;59(7):3046–62. doi:10.1021/acs.jmedchem.5b01679
9. Hou X, Zaks T, Langer R, Dong Y. Lipid nanoparticles for mRNA delivery. *Nat Rev Mater*. 2021 Aug 10;6(12):1078–94. doi:10.1038/s41578-021-00358-0
10. Song D, Zhao Y, Wang Z, Xu Q. Tuning Lipid Nanoparticles for RNA Delivery to Extrahepatic Organs. *Advanced Materials*. 2024 Nov 5;36(44). doi:10.1002/adma.202401445
11. Schoenmaker L, Witzigmann D, Kulkarni JA, Verbeke R, Kersten G, Jiskoot W, et al. mRNA-lipid nanoparticle COVID-19 vaccines: Structure and stability. *Int J Pharm*. 2021 May;601:120586. doi:10.1016/j.ijpharm.2021.120586

12. Ndeupen S, Qin Z, Jacobsen S, Bouteau A, Estanbouli H, Igyártó BZ. The mRNA-LNP platform's lipid nanoparticle component used in preclinical vaccine studies is highly inflammatory. *iScience*. 2021 Dec;24(12):103479. doi:10.1016/j.isci.2021.103479
13. Zhang Y, Sun C, Wang C, Jankovic KE, Dong Y. Lipids and Lipid Derivatives for RNA Delivery. *Chem Rev*. 2021 Oct 27;121(20):12181–277. doi:10.1021/acs.chemrev.1c00244
14. Nakamura T, Kawai M, Sato Y, Maeki M, Tokeshi M, Harashima H. The Effect of Size and Charge of Lipid Nanoparticles Prepared by Microfluidic Mixing on Their Lymph Node Transitivity and Distribution. *Mol Pharm*. 2020 Mar 2;17(3):944–53. doi:10.1021/acs.molpharmaceut.9b01182
15. Li B, Raji IO, Gordon AGR, Sun L, Raimondo TM, Oladimeji FA, et al. Accelerating ionizable lipid discovery for mRNA delivery using machine learning and combinatorial chemistry. *Nat Mater*. 2024 Jul 13;23(7):1002–8. doi:10.1038/s41563-024-01867-3
16. Handl J, Kell DB, Knowles J. Multiobjective Optimization in Bioinformatics and Computational Biology. *IEEE/ACM Trans Comput Biol Bioinform*. 2007 Apr;4(2):279–92. doi:10.1109/TCBB.2007.070203
17. Wang Y, Zhu H, Madabushi R, Liu Q, Huang S, Zineh I. Model-Informed Drug Development: Current US Regulatory Practice and Future Considerations. *Clin Pharmacol Ther*. 2019 Apr;105(4):899–911. doi:10.1002/cpt.1363
18. Schmeisser S, Miccoli A, von Bergen M, Berggren E, Braeuning A, Busch W, et al. New approach methodologies in human regulatory toxicology – Not if, but how and when! *Environ Int*. 2023 Aug;178:108082. doi:10.1016/j.envint.2023.108082
19. Oh Y, Bedingfield SK, Schneebeli ST, Li J, Ardekani AM, Colston KJ, et al. Challenges and opportunities in computational studies for lipid nanoparticle development. *npj Drug Discovery*. 2025 Sep 1;2(1):23. doi:10.1038/s44386-025-00024-3
20. Di Salvatore V, Cernuto F, Russo G, Pappalardo F. A computational framework for optimizing mRNA vaccine delivery via AI-guided nanoparticle design and in silico gene expression profiling. *Front Immunol*. 2025 Dec 5;16. doi:10.3389/fimmu.2025.1628583
21. Arunachalam PS, Scott MKD, Hagan T, Li C, Feng Y, Wimmers F, et al. Systems vaccinology of the BNT162b2 mRNA vaccine in humans. *Nature*. 2021 Aug 19;596(7872):410–6. doi:10.1038/s41586-021-03791-x
22. van de Poel I, Robaey Z. Safe-by-Design: from Safety to Responsibility. *Nanoethics*. 2017 Dec 22;11(3):297–306. doi:10.1007/s11569-017-0301-x
23. Schmutz M, Borges O, Jesus S, Borchard G, Perale G, Zinn M, et al. A Methodological Safe-by-Design Approach for the Development of Nanomedicines. *Front Bioeng Biotechnol*. 2020 Apr 2;8. doi:10.3389/fbioe.2020.00258

24. Bastús NG, Puntos V. Nanosafety: Towards Safer Nanoparticles by Design. *Curr Med Chem*. 2018 Dec 3;25(35):4587–601. doi:10.2174/0929867324666170413124915
25. Allen R, Rieger T, Musante C. Efficient Generation and Selection of Virtual Populations in Quantitative Systems Pharmacology Models. *CPT Pharmacometrics Syst Pharmacol*. 2016 Mar 17;5(3):140–6. doi:10.1002/psp4.12063
26. Moingeon P, Chenel M, Rousseau C, Voisin E, Guedj M. Virtual patients, digital twins and causal disease models: Paving the ground for in silico clinical trials. *Drug Discov Today*. 2023 Jul;28(7):103605. doi:10.1016/j.drudis.2023.103605
27. Famoye F, Lee C, Eugene N. Beta-Normal Distribution: Bimodality Properties and Application. *Journal of Modern Applied Statistical Methods*. 2004 May 1;3(1):85–103. doi:10.22237/jmasm/1083370200
28. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2026. Available from: <https://cran.r-project.org/manuals.html> doi:10.32614/R.manuals
29. Pappalardo F, Pennisi M, Motta S. Universal immune system simulator framework (UISS). In: *Proceedings of the First ACM International Conference on Bioinformatics and Computational Biology* [Internet]. New York, NY, USA: ACM; 2010. p. 649–50. Available from: <https://dl.acm.org/doi/10.1145/1854776.1854900> doi:10.1145/1854776.1854900
30. Russo G, Crispino E, Maleki A, Di Salvatore V, Stanco F, Pappalardo F. Beyond the state of the art of reverse vaccinology: predicting vaccine efficacy with the universal immune system simulator for influenza. *BMC Bioinformatics*. 2023 Jun 5;24(1):231. doi:10.1186/s12859-023-05374-1
31. van Breugel B, Liu T, Oglic D, van der Schaar M. Synthetic data in biomedicine via generative artificial intelligence. *Nature Reviews Bioengineering*. 2024 Oct 8;2(12):991–1004. doi:10.1038/s44222-024-00245-7
32. Reel PS, Reel S, Pearson E, Trucco E, Jefferson E. Using machine learning approaches for multi-omics data analysis: A review. *Biotechnol Adv*. 2021 Jul;49:107739. doi:10.1016/j.biotechadv.2021.107739
33. McInnes L, Healy J, Saul N, Großberger L. UMAP: Uniform Manifold Approximation and Projection. *J Open Source Softw*. 2018 Sep 2;3(29):861. doi:10.21105/joss.00861
34. Wright MN, Ziegler A. ranger : A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *J Stat Softw*. 2017;77(1). doi:10.18637/jss.v077.i01
35. Breiman L. Random Forests. *Mach Learn*. 2001 Oct;45(1):5–32. doi:10.1023/A:1010933404324

36. Kapoor S, Narayanan A. Leakage and the reproducibility crisis in machine-learning-based science. *Patterns*. 2023 Sep;4(9):100804. doi:10.1016/j.patter.2023.100804
37. Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Arroita G, et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*. 2017 Aug 3;40(8):913–29. doi:10.1111/ecog.02881
38. Marler RT, Arora JS. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*. 2004 Apr 1;26(6):369–95. doi:10.1007/s00158-003-0368-6
39. Squair JW, Gautier M, Kathe C, Anderson MA, James ND, Hutson TH, et al. Confronting false discoveries in single-cell differential expression. *Nat Commun*. 2021 Sep 28;12(1):5692. doi:10.1038/s41467-021-25960-2
40. Wickham H. *ggplot2* [Internet]. Cham: Springer International Publishing; 2016. (Use R!). Available from: <http://link.springer.com/10.1007/978-3-319-24277-4> doi:10.1007/978-3-319-24277-4
41. Kitano H. Systems Biology: A Brief Overview. *Science* (1979). 2002 Mar;295(5560):1662–4. doi:10.1126/science.1069492

Data availability

The human single-cell RNA-sequencing dataset used for retrospective biological consistency validation is publicly available from the NCBI Gene Expression Omnibus under accession number GSE171964. The virtual patient data, RNA-LNP formulation library, UISS-derived mechanistic descriptors, synthetic multi-omics dataset, and processed results generated or analysed in this study are available from the corresponding author upon reasonable request.

Code availability

All R scripts used to generate the virtual patient cohorts, construct the RNA-LNP formulation design space, derive the synthetic multi-omics representations, compute the Safety-by-Design score, train and evaluate the machine-learning models, perform Pareto-based optimization, conduct retrospective transcriptomic validation, calculate the Biological Program Concordance Index, and reproduce the figures are available from the corresponding author upon reasonable request.

Ethics statement

This study involved computational simulations, synthetic data, and secondary analysis of a publicly available, de-identified human transcriptomic dataset. No new human participants or

animals were recruited or experimentally investigated as part of this study. Consequently, additional institutional ethics approval and informed consent were not required. Ethical approval and informed-consent procedures associated with the original human dataset are reported in the corresponding primary publication and GEO record.

Consent for publication

Not applicable

Clinical trial number

Not applicable

Author contributions

VDS: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Visualization, and Writing – original draft. **GR:** Mechanistic modeling, Validation, Resources, and Writing – review and editing. **FP:** Supervision, Project administration, Funding acquisition, and Writing – review and editing. All authors reviewed and approved the final manuscript.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

Not applicable.

Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors used OpenAI to support English-language editing and manuscript organization. All AI-assisted text was critically reviewed, edited, and validated by the authors. The authors take full responsibility for the content, analyses, interpretations, and conclusions presented in this work.