Public Comment

Docket No. FDA-2025-D-6131

Draft Guidance for Industry: General Considerations for the Use of New Approach Methodologies in Drug Development

Submitted by: Precigenetics, Inc.

Public comment — no confidential business information included.

Table of Contents

Comment 14. Reviewer rubrics and specialist reviewer pairings

Closing

Appendix A. NAM Context-of-Use Briefing Template

Dear Sir or Madam:

Precigenetics, Inc. ("Precigenetics") respectfully submits this comment in response to the draft guidance for industry "General Considerations for the Use of New Approach Methodologies in Drug Development," issued by the Center for Drug Evaluation and Research (CDER) on March 18, 2026 (Docket No. FDA-2025-D-6131; 91 Fed. Reg. 13313, March 19, 2026). We appreciate FDA's continued leadership in establishing a clear, scientifically rigorous framework for the integration of New Approach Methodologies (NAMs) into nonclinical drug development.

About Precigenetics

Precigenetics is a U.S.-based platform company in nonclinical drug development. Of particular relevance to this guidance, our work integrates biophotonics, microphysiological disease modeling, mechanism-based (physics-grounded) computational modeling of biology, and machine-learning analysis to generate human-relevant, longitudinal, multi-endpoint, information-rich nonclinical evidence. Our broader platform falls within the class of human biology–based nonclinical test methods recognized in section 505(z) of the Federal Food, Drug, and Cosmetic Act (FD&C Act), as amended by section 3209 of the Food and Drug Omnibus Reform Act of 2022. We are an early-stage company preparing to use platform-level, multi-endpoint, longitudinal measurement methods to support future nonclinical regulatory submissions, including under section 351(k) of the Public Health Service Act for biosimilar comparability — a use case that will become increasingly important as major reference biologics reach patent expiration in the late 2020s. The specific technical implementation of our platform is outside the scope of this public comment and may be the subject of a separate confidential technical submission to the Agency at a later date.

This comment does not seek FDA endorsement of any proprietary company or technology. It is offered from the perspective of a sponsor that intends to generate, submit, and have evaluated NAM evidence under the framework this guidance establishes — a perspective that, we believe, may be useful to the Agency in assessing how the draft will function in practice. We request targeted, technology-neutral clarifications that would make the final guidance more useful for sponsors, technology developers, rare-disease programs, and FDA review divisions evaluating integrated NAM evidence.

General Endorsement

We strongly endorse the four-pillar validation framework established in Section III of the draft guidance — Context of Use (COU), Human Biological Relevance, Technical Characterization, and Fit-for-Purpose. These principles are well-aligned with the scientific realities of contemporary NAM development and provide a coherent, durable foundation for both sponsor planning and FDA review. We further endorse the framework's explicit alignment with the recommendations of the Science Board to the Food and Drug Administration (Potential Approaches to Drive Future Integration of New Alternative Methods for

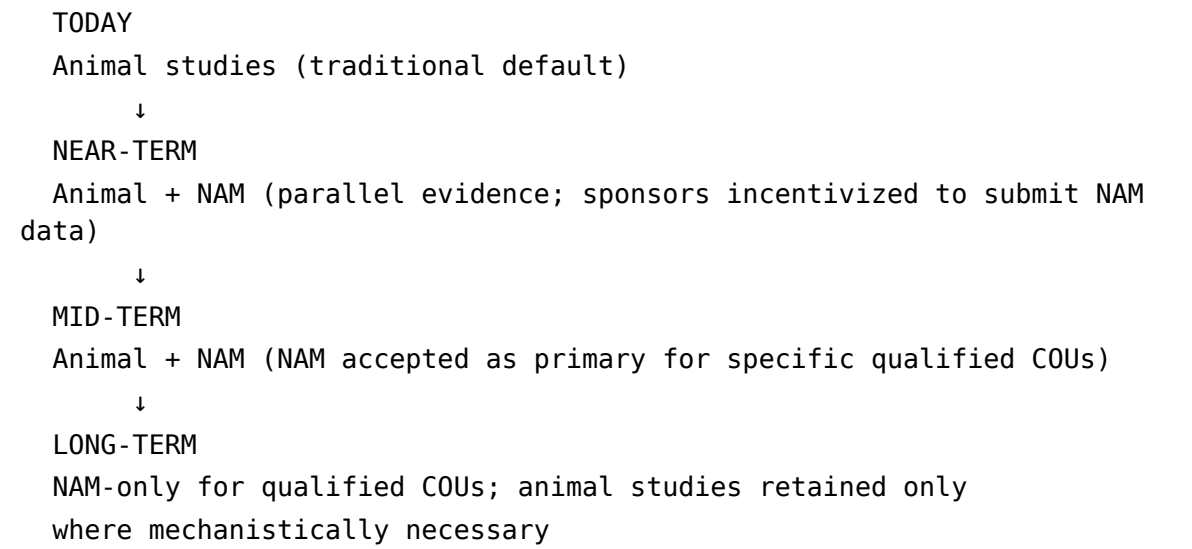
Regulatory Decision-Making, October 2024), with CDER's Roadmap to Reducing Animal Testing in Preclinical Safety Studies (April 2025), and with the ICCVAM Validation Workgroup's 2024 report on Validation, Qualification, and Regulatory Acceptance of NAMs.

We particularly commend the clarification, at lines 34–40 of the draft, that **a NAM does not necessarily need to be formally validated to be considered for review**, and that a fit-for-purpose NAM may adequately address specific toxicological concerns within a weight-of-evidence (WoE) framework even where formal validation has not yet been completed. This clarification meaningfully lowers the practical barrier to entry for emerging NAM categories. It is consistent with the broader WoE approach FDA has employed in related regulatory contexts, and it appropriately recognizes that validation is itself an evolving and accumulating process rather than a binary precondition. We respectfully urge FDA to retain this language unchanged in the final guidance.

We also endorse the guidance's emphasis on early and structured engagement with the appropriate FDA review division (lines 76–82). For sponsors developing emerging NAM categories, early engagement is essential, and the explicit signal in the guidance that the Agency welcomes such engagement is meaningful. We intend to seek such engagement with the relevant CDER review division as our own NAM-based programs progress.

Parallel-evidence, phased-reduction framing

Read together, we understand the draft to enable a pragmatic transition pathway for the field. In the near term, the guidance creates space for sponsors to submit NAM evidence alongside traditional animal studies, with NAM data contributing to an accumulating weight of evidence and to reviewer familiarity. Over time, as cumulative evidence demonstrates that NAMs reliably address specific contexts of use, traditional animal studies for those contexts can be reduced or eliminated where they would be redundant. This is, we believe, the right trajectory: it preserves safety by not requiring sponsors to immediately abandon evidence types FDA reviewers are familiar with, while incentivizing the generation of NAM evidence that will form the basis for eventual reduction of redundant animal studies. The schematic below captures this framing:



```
Each transition is driven by accumulated weight of evidence,
not by calendar timeline; COU-specific.
```

The comments below are intended to make this trajectory operationally credible. None of our comments asks FDA to lower the evidentiary standard. Each asks FDA to make the standard more precise, more actionable, and more consistently applied so that the parallel-evidence-then-phased-reduction trajectory the guidance enables can be realized in practice.

Part A. Framework and Definitions

Comment 1. Expressly recognize integrated NAM systems, and clarify that a Context of Use may be defined at the level of an integrated measurement method.

Reference: Section I; Section II; Section III.A, lines 134–147.

The draft appropriately describes NAMs as including complex in vitro, 2D in vitro, in chemico, and in silico studies. The final guidance should go one step further and clarify, as a definitional matter, that a NAM may be an integrated system: a human-relevant biological model together with the physical, analytical, and computational infrastructure used to expose, measure, and interpret that model.

This clarification matters because many next-generation NAMs are not merely "models." They are measurement systems. A patient-derived organoid, an organ-on-chip, a spheroid, or a 2D human-cell system becomes more decision-relevant when paired with controlled dosing, perfusion, time-course exposure, non-destructive sensing, and computational analysis. Conversely, a biologically sophisticated model can still produce weak evidence if the readout is static, destructive, poorly controlled, or non-reproducible. A technology-neutral definition would avoid unintentionally privileging model catalogs over platforms that improve the quality, continuity, and interpretability of human-relevant evidence across many model classes.

Recommendation: clarify that NAMs may include integrated systems, and that a COU may be defined at the level of a measurement method supporting multiple drug development decisions through an accumulating weight-of-evidence approach.

Suggested addition to Section I or Section II:

> *For purposes of this guidance, NAMs may include integrated systems comprising a human-relevant biological model, a physical culture or exposure platform, measurement or readout technologies, sensors, and computational or machine-learning analysis used to generate interpretable endpoints for a defined context of use. A NAM need not be limited to the biological model itself; the measurement and analysis method may also be central to the NAM and should be described, characterized, and validated as appropriate for the intended context of use.*

This framing has direct consequences for the COU section. The five illustrative COUs in the draft (lines 140–147) are each framed around a single assay measuring a single endpoint to inform a single decision. This framing is appropriate for many traditional NAMs but does not cleanly accommodate the integrated systems described above, whose scientific value lies precisely in informing multiple drug development decisions simultaneously from one integrated experimental design — patient-derived organoid panels, single-cell and spatial multi-omic methods, multiplexed unbiased phenotypic methods, and information-rich continuous measurement platforms among them. Separating such an integrated experiment into per-endpoint sub-studies would discard much of the information it provides, and in many cases is analytically inappropriate because the endpoints are jointly derived from a single underlying dataset.

The final guidance should clarify, either through additional language in Section III.A or through an explanatory footnote:

> *A COU may be defined at the level of an individual assay supporting a single drug development decision, or at the level of a measurement method or platform supporting multiple decisions through an accumulating weight-of-evidence approach. In the latter case, sponsors should articulate the platform-level COU, describe the set of decisions the method is intended to inform, and provide validation evidence appropriate to each decision.*

Taken together, these two clarifications would preserve the scientific rigor of the four-pillar framework while accommodating the integrated character of next-generation NAMs, and would be fully consistent with the WoE principle articulated at lines 34–40.

Comment 2. Permit modular validation strategies for NAMs composed of separable components.

Reference: Section III, cross-cutting.

The draft guidance correctly focuses validation on the intended context of use. Integrated NAM systems, however, are often modular: a sponsor may use one biological model, one exposure-control device, one measurement modality, and one analysis pipeline. Some components are technically characterized once and re-used across many applications, while others require disease-specific or endpoint-specific bridging. Requiring every integrated NAM to be revalidated from first principles for every disease model, drug class, or endpoint would slow adoption, create unnecessary cost, and disadvantage sponsors of integrated platforms relative to sponsors of single-purpose assays.

Recommendation: allow sponsors to justify a modular validation strategy, in which technically characterized components need not be revalidated from first principles for each new application, provided the sponsor explains prior evidence boundaries and proposed bridging studies.

Suggested language for Section III:

> *Where a NAM is composed of separable components — for example, a biological model, an exposure device, a measurement modality, a sensor system, and an analysis algorithm — sponsors may justify a modular validation strategy. Under such a strategy, elements*

that have been technically characterized across the relevant operating conditions need not be revalidated from first principles for each new disease model or drug program, provided the sponsor explains the boundaries of the prior evidence, any changes to the system, and the specific bridging studies needed to support the proposed context of use.

A modular approach would preserve scientific rigor while encouraging scalable validation. It would also help FDA distinguish between two issues that often require different kinds of evidence: the introduction of a new biological context (which may require new biological-relevance evidence) and the introduction of a new measurement method (which may require new technical-characterization evidence).

Comment 12. Add a glossary of key terms with cited definitions.

Reference: Cross-cutting.

The draft guidance uses several closely related terms — "platform," "device," "chip," "model," "method," "assay," and "test" — in ways that, read across the document, are partially overlapping rather than rigorously distinguished. The categories of method the guidance addresses (NAM, MPS, organ-on-chip, organoid, in vitro, in chemico, in silico) are similarly used without explicit definition.

Recommendation: include a brief glossary at the front of the final guidance that adopts established cited definitions where they exist, to support consistent interpretation across sponsors, reviewers, and international harmonization counterparts.

Proposed starting definitions, drawn from authoritative sources:

Term	Suggested definition	Cited source
New Approach Methodology (NAM)	Any technology, methodology, approach, or combination thereof that can be used to provide information on chemical hazard and risk assessment that replaces, reduces, or refines the use of animals.	ICCVAM (Harrill et al., 2024), Validation, Qualification, and Regulatory Acceptance of NAMs
Microphysiological system (MPS)	An in vitro platform containing human or animal cells, organ- or tissue-derived explants, and/or self-assembling organoids within microenvironmental niches capable of facilitating physiologically meaningful biochemical, mechanical, and electrical tissue- or organ-level responses.	FDA Alternative Methods Working Group, as cited in the FDA Roadmap to Reducing Animal Testing (April 2025)
Organ-on-chip (OoC)	A subset of MPS, comprising microfluidic devices with tissue constructs that incorporate fluid flow. All organ-on-chips are MPS, but not all MPS are organ-on-chips; static three-dimensional culture systems (organoids, spheroids) are MPS without being organ-	GAO Report on Organ-on-a-Chip Technology (GAO-25-107335, 2025)

Term	Suggested definition	Cited source
	on-chips.	
Context of Use (COU)	A clear statement of the specific regulatory purpose for which a NAM is intended, including the drug development decision(s) the NAM is intended to inform.	Draft Guidance, Section III.A; ICCVAM 2024
Validation	The process by which the accuracy, reliability, and relevance of a procedure are established for a specific context of use.	ICCVAM 2024, as cited in the draft guidance at line 103
Qualification	A determination that a drug development tool and its proposed COU can be relied upon to have a specific interpretation and application in drug development and regulatory review.	FD&C Act section 507(e)(7); draft guidance line 104
Reference cell standard	A characterized, lot-controlled biological preparation (for example, a pooled-donor primary cell lot) made available as a common substrate against which NAM performance can be evaluated across laboratories and platforms.	Proposed; analogous to USP biological reference standards and NIST reference materials
Weight of evidence (WoE)	An integrative approach in which multiple lines of evidence are evaluated together, considering quality, relevance, and consistency, to support a regulatory determination.	Consistent with usage in ICCVAM 2024 and OECD frameworks

Adoption of a glossary at the front of the final guidance, with cited sources, would reduce interpretive ambiguity, improve cross-document consistency, and facilitate international harmonization with EMA, PMDA, and OECD counterparts that use overlapping but not identical terminology.

Part B. Validation Considerations by Pillar

Comment 3. Information richness over time: validation considerations for longitudinal and continuous measurement methods.

Reference: Section III.B and III.C, lines 196–249.

The Technical Characterization section provides clear and detailed guidance applicable to most NAM categories. We particularly note the inclusion of platform-specific considerations for organ chips at lines 234–249 — flow, matrix, shear stress, scaffold biocompatibility, and test-article/platform interactions —

which is an excellent model for how the framework can accommodate emerging method classes through targeted, platform-specific bullets that supplement (rather than replace) the general bullets above them.

Many drug effects are not well captured by a single endpoint: cells may die, adapt, recover, resist, change metabolic state, or enter a delayed toxicity trajectory. Destructive endpoint assays can miss that timing and force sponsors to infer a dynamic process from separate samples. When properly controlled, longitudinal same-sample data can provide a more biologically informative trajectory than a terminal snapshot, particularly for hepatotoxicity, cardiotoxicity, neurotoxicity, immune activation, mitochondrial and oxidative stress, resistance, recovery, and dose-schedule decisions.

Recommendation: add a set of platform-specific bullets to Section III.C for continuous and longitudinal measurement methods, paralleling the existing organ chip bullets.

This class includes any method that captures biological information about the same sample at multiple time points across an extended observation window — for example, repeated measurements of cellular state, function, chemistry, morphology, or activity over hours, days, or longer — rather than producing a single terminal endpoint per sample. We propose:

- Measurement frequency and temporal resolution, with justification in relation to the biological kinetics observed. Sampling that is too sparse may miss biologically meaningful transitions; sampling that is too dense may introduce confounding perturbation.
- Signal stability over the observation window, including drift correction or normalization procedures applied.
- Evidence that the measurement process itself does not perturb the biological system in ways that confound the endpoint. Depending on the measurement, relevant perturbation modes may include effects of any energy or material introduced by the measurement (for example, light, electrical, mechanical, thermal, or chemical), as well as cumulative effects of repeated sampling over an extended window. Sponsors should describe the controls used to identify and bound these effects.
- Documentation of the analytical pipeline that converts raw temporal signals into the reported endpoint, including filtering, smoothing, trajectory-fitting, or feature-extraction procedures, and sensitivity of the reported endpoint to those analytical choices.
- For methods producing high-dimensional output at each time point: characterization of inter-experiment and inter-batch reproducibility of the full output, not only of the specific reported endpoint.

In Section III.B, we suggest general language such as:

> *FDA recognizes that longitudinal, non-destructive measurements in the same biological system may be especially informative for temporal, metabolic, adaptive, recovery-related, or exposure-dependent toxicities. Sponsors should describe how same-sample trajectories are generated, whether the measurement process may perturb the biological system, and how trajectory features relate to toxicological mechanisms, clinical biomarkers, or other relevant endpoints for the proposed context of use.*

Comment 4. Strengthen Human Biological Relevance for non-destructive methods, expanded endpoints, and mechanism-led cellular composition.

Reference: Section III.B, lines 149–194.

The Human Biological Relevance section provides excellent illustrative examples for established in vitro NAM categories — the neurotoxicity, hepatotoxicity, and respiratory toxicity vignettes (lines 163–194) are clear, concrete, and well-chosen. We offer three connected refinements.

Recommendation: strengthen Section III.B in three ways — (4A) address physiological-state maintenance for non-destructive methods; (4B) clarify that NAM endpoints need not be limited to clinical biomarker analogs; (4C) shift cellular-composition guidance from canonical-checklist to mechanism-led justification.

4A. Maintenance of physiological state for non-destructive methods.

For NAMs in which cells are destroyed at the endpoint, the question of whether the measurement itself altered the biological system is largely moot — the measurement is the endpoint. For NAMs that observe living cells continuously or longitudinally, however, the validity of the measurement depends on demonstrating that the cells remain in a physiologically relevant state throughout the observation period.

> *For methods that observe living cells over an extended observation window, sponsors should describe how the physiological state of the cells is maintained and characterized throughout the window, including any reference markers used to confirm continued viability, identity, and function. Sponsors should also address whether the measurement modality itself affects the physiological state of the cells, and should provide evidence that any such effects do not confound the endpoint of interest.*

4B. Endpoints need not be limited to clinical biomarker analogs.

The hepatotoxicity example at lines 181–186 names albumin, urea, ALT, AST, and CYP450 expression — endpoints with direct clinical-blood-biomarker analogs. Read together, these examples may imply that in vitro NAM endpoints should be limited to measurements that look like clinical chemistry panels. In vitro and ex vivo systems can in fact measure information unobservable in clinical settings: intracellular metabolite kinetics, real-time mitochondrial membrane potential, lipid droplet formation dynamics, bile canaliculus contractility, organelle-level stress responses, and single-cell heterogeneity. Restricting NAM endpoints to clinical analogs would discard the principal informational advantage of in vitro measurement.

> *NAM endpoints need not be limited to direct analogs of clinical biomarkers. Sponsors may use mechanistically-linked in vitro endpoints — including intracellular, organelle-level, and single-cell measurements not available in clinical settings — provided the sponsor establishes the mechanistic linkage between the NAM endpoint and the clinical outcome of interest.*

4C. Mechanism-led cellular composition.

The hepatotoxicity example specifies that the model "should contain the relevant cell types (e.g., hepatocytes, stellate cells, Kupffer cells)." Read as guidance, this can inadvertently become a checklist. For some hepatotoxic drug classes, Kupffer cells are essential (immune-mediated DILI); for others,

hepatocytes alone are sufficient (direct metabolic toxicity). A mechanism-led standard would be more scientifically rigorous than a checklist.

> *Cellular composition of the NAM should be justified based on the cell types mechanistically relevant to the toxicity or pharmacology of interest, rather than by reference to canonical anatomical composition alone. Sponsors may use orthogonal evidence — including single-cell transcriptomic data, single-cell drug-sensitivity data, or published mechanistic literature — to support the relevance of the chosen composition for the proposed context of use.*

Comment 5A. Expand Technical Characterization for information-rich, multimodal, AI/ML-enabled, and mechanism-grounded NAMs.

Reference: Section III.C, lines 196–249.

Section III.C includes useful technical-characterization criteria covering dose and frequency, test-substance preparation, detection methods, instrumentation, materials compatibility, assay variability, statistical methods, working duration, cell source, biological variability, reference compounds, culture media, and organ-chip considerations.

Recommendation: expand Section III.C to address information-rich, multimodal, and computationally-mediated NAMs, including instrument calibration, assay-induced artifacts, throughput, stratified per-class performance, pre-registration of validation studies, AI/ML governance with explicit attention to interpretability and black-box risk, multimodal class-imbalance considerations, and mechanism-grounded computational modeling.

Proposed additional bullets:

- Instrument and sensor calibration, including sensitivity, specificity where applicable, limit of detection or equivalent performance metrics, signal-to-noise, drift, batch effects, and reproducibility across instruments, operators, sites, and runs.
- For information-rich NAMs that capture broad biological content per sample — including methods that probe cellular chemistry, whole-cell or whole-tissue phenotype in an unbiased manner, multi-omic readouts, or any combination of these — sponsors should document the parameters controlling data richness (the volume and dimensionality of biological information produced per sample), informational richness (the biologically meaningful content extracted from that data after analysis), and the temporal and spatial axes over which information is captured. Each of these axes (chemistry, phenotype, omics, time, space) is relevant to mechanism of action determination, and sponsors should describe which axes the NAM addresses and at what resolution. Where data are inherently sparse along one or more axes (for example, sparse temporal sampling or sparse spatial coverage), sponsors should describe how that sparsity is handled in analysis and how it bounds the interpretability of the result. Finer technical specifics of any particular measurement modality may be appropriate for confidential technical submission rather than the public docket.

- Preprocessing, quality control, and perturbation controls. For any NAM whose output passes through a preprocessing or feature-extraction pipeline before producing the reported endpoint, sponsors should document that pipeline, the quality-control thresholds applied at each stage, the sensitivity of the reported endpoint to those preprocessing choices, and any controls used to demonstrate that the measurement process itself does not materially perturb the biological response for the proposed context of use. This is consistent with the principle that the measurement and the analysis pipeline together constitute the NAM, as discussed in Comments 1 and 2.
- Explicit documentation of how the assay itself may introduce artifacts. This extends the existing language at lines 213–218 on incubator conditions and assay variability to cover, for example, edge effects in multi-well systems, evaporation gradients, dye- or label-induced perturbations, mechanical agitation effects, and interactions between the test article and components introduced by the assay (such as solvents, surfactants, or carriers). Sponsors should describe the controls used to identify and bound these artifacts.
- For microfluidic or controlled-exposure systems used with longitudinal measurement: flow rate, perfusion, shear stress, dosing gradients, residence time, evaporation, adsorption, leaching, material compatibility, exposure history, and how those parameters are monitored and controlled over the observation window.
- For multimodal systems: how data streams are synchronized, normalized, versioned, and linked to samples, wells, devices, lots, donors, time points, and dosing conditions.
- For machine-learning or statistical models used to convert measurements into endpoints: training and validation data composition, model architecture, prespecified performance metrics, uncertainty estimation, procedures to detect and prevent data leakage and overfitting, sensitivity to biological and technical variability, distribution-shift handling between training and submission data, and procedures for locking or updating models used in regulatory submissions.
- Interpretability and black-box risk. Models whose outputs cannot be meaningfully interpreted — either through inherently interpretable architectures, post-hoc interpretability methods of established quality, or convergent corroboration with mechanism-grounded models — should be held to a correspondingly higher bar on validation data volume, throughput, replication, and uncertainty quantification. Black-box models without compensating evidence of validation rigor should be disfavored. Interpretability is not a substitute for performance, and performance is not a substitute for interpretability; sponsors should address both, with the appropriate balance depending on the COU.
- Multimodal models and class imbalance across modalities. Where a NAM combines multiple measurement modalities (for example, mapping between a phenotypic assay and a single-cell transcriptomic readout, or fusing imaging with electrophysiological data), sponsors should characterize not only sample-level class imbalance but also modality-level data sufficiency. A multimodal model in which one modality is well-represented and another is sparsely represented may produce outputs that appear unified but in fact rely disproportionately on the data-rich modality. Sponsors should report per-modality data volume, per-modality reproducibility, and evidence that the model is not exploiting modality-specific artifacts. Multimodal foundation models in particular warrant explicit treatment of inter-modality balance, alignment validation, and the failure modes that arise when modalities are unequally informative for a given task.

- Mechanism-grounded (physics-based) computational modeling. Where a NAM incorporates mechanism-based computational models — for example, physiologically based pharmacokinetic models, biophysical or biochemical reaction-network simulations, or first-principles models of cellular processes — sponsors should describe the underlying mechanistic assumptions, parameter sources, calibration and validation procedures, sensitivity analyses, and known boundary conditions. Mechanism-grounded models and data-driven models are complementary; sponsors using both should describe how the two are integrated and how disagreements between them are resolved.
- Throughput reporting. Sponsors should report the throughput (number of independent biological replicates per drug, number of distinct drugs evaluable per unit time per system) at which the claimed performance characteristics were obtained. Throughput is relevant to both the statistical strength of validation data and the practical utility of the NAM in a real submission workflow.
- Stratified performance reporting. Sponsors should report predictive performance disaggregated by drug class and toxicity mechanism, in addition to any aggregate measures. A NAM's performance for direct hepatotoxic mechanisms may differ from its performance for immune-mediated DILI; aggregate performance numbers can be misleading where class composition of the validation set differs from class composition of intended use.
- Pre-registration and minimum reporting standards. To support field-wide rigor and reduce the risk of selective reporting, sponsors are encouraged (and FDA may consider whether they should be required) to pre-register NAM validation studies on a public registry, with prespecified compound sets, endpoints, statistical analyses, and performance thresholds. All pre-specified endpoints should be reported regardless of outcome.
- For raw and processed NAM data: metadata sufficient for FDA to understand the provenance, processing, and interpretation of the evidence, including data dictionaries and audit trails where appropriate.

These additions are consistent with the OECD GIVIMP principles referenced at line 201 of the draft, with established machine-learning governance practice in other FDA-regulated settings, and with the broader scientific-integrity infrastructure used in clinical research.

Comment 5B. Strengthen cell source characterization, reference materials, and accounting for necrosis across NAM categories.

Reference: Section III.C, lines 220–226.

The current draft addresses cell source and biological variability briefly at lines 220–226. The cell source is, in our view, often the single largest variance term in MPS reproducibility, and it is the variance term most commonly under-reported.

Recommendation: expand the cell-source provisions of Section III.C to address reference cell standards, disclosure of immortalized cell line use, and consistent accounting for necrosis and equivalent canonical endpoints across NAM categories.

5B.i. Reference cell standards.

Primary hepatocyte preparations from different donors, immortalized hepatic cell lines (HepG2, HepaRG), and iPSC-derived hepatocyte preparations from different maturation protocols can produce meaningfully different results on nominally identical platforms. The final guidance should encourage — and, in time, FDA work with NIST, USP, and ICCVAM to establish — standardized pooled-donor reference cell preparations that any NAM developer claiming class-specific performance must demonstrate against. This is analogous to USP reference standards for analytical methods and to NIST's reference materials program.

5B.ii. Disclosure and justification of cell line use.

Where a NAM uses immortalized cell lines rather than primary or iPSC-derived cells, sponsors should disclose this and justify the choice in relation to the COU, including any known phenotypic departures from the primary cell type the line is intended to model (for example, reduced or altered CYP450 expression in HepG2 relative to primary hepatocytes). This is not a prohibition on immortalized lines, which remain appropriate for many mechanistic studies; it is a transparency and justification requirement.

5B.iii. Accounting for necrosis across organoid and MPS systems.

Necrosis is a canonical and mechanistically meaningful toxicity endpoint, but it is currently accounted for in different ways across organoid, spheroid, and organ-chip systems — some report LDH release, some report viability dye exclusion, some report image-based morphological criteria, and some do not report necrosis at all as a distinct endpoint. Comparative performance claims across NAM categories can be difficult to interpret as a result. The final guidance should recommend that sponsors explicitly account for necrosis (and equivalent canonical endpoints such as apoptosis, mitochondrial dysfunction, and oxidative stress) in their NAM submissions, describing how each is measured or, where it is not measured, justifying the omission in relation to the proposed COU. Over time, ICCVAM and the IQ Consortium's Microphysiological Systems Affiliate are appropriate convening bodies for developing common reporting expectations across NAM categories.

Comment 6. Add a fourth Fit-for-Purpose objective for integrated platform NAMs; clarify fit-for-purpose evidence when traditional comparators are weak or unavailable; and acknowledge that no nonclinical method is ground truth for human biology.

Reference: Section III.D, lines 251–289.

Section III.D enumerates three drug development objectives that a fit-for-purpose NAM may fulfill: (i) replacing or offering an alternative approach to traditional methods, (ii) filling a data gap where traditional models are unavailable or insufficient, and (iii) confirming or complementing findings from traditional methods. These three objectives capture the principal historical roles of NAMs in nonclinical development.

Recommendation: (i) add a fourth fit-for-purpose objective recognizing the distinctive value of platform-level NAMs that inform multiple decisions through a single integrated measurement; (ii) clarify that fit-for-purpose evidence does not require comparison to a traditional method where no suitable traditional comparator exists; and (iii) acknowledge that traditional comparator methods are

themselves imperfect proxies for human biology, and that disagreements between a NAM and a traditional method should be evaluated on the merits of both rather than reflexively resolved in favor of the historical method.

Suggested fourth objective:

> *Informs multiple drug development decisions through a single integrated measurement, for example, by providing data that simultaneously supports dose selection, mechanistic understanding, and risk prioritization, in a manner consistent with the weight-of-evidence framework. In such cases, the fit-for-purpose case rests on the coherence of the integrated measurement and the appropriateness of each decision it informs, rather than on a one-to-one mapping to a single traditional method.*

In many important contexts, the traditional comparator is unavailable, poorly predictive, non-human, or not relevant to the specific decision — particularly in rare disease, patient-derived systems, immune-specific biology, species-restricted mechanisms, and mechanistic questions that animal studies cannot answer:

> *When traditional comparator methods are unavailable, insufficiently informative, or not expected to be human-relevant for the proposed context of use, sponsors should justify alternative approaches to demonstrating fit-for-purpose performance. Such approaches may include reference compounds, orthogonal biomarkers, clinical concordance where available, mechanistic plausibility, human biological relevance, prospective-retrospective validation, inter-laboratory reproducibility, and uncertainty analysis. A NAM need not be shown to be broadly predictive across all toxicities or drug classes if it is reliable for the specific decision and context of use proposed by the sponsor.*

Third, the final guidance should acknowledge that no nonclinical method — animal-based or NAM-based — is ground truth for human biology. Traditional animal models have well-documented translation failures (the literature on cross-species toxicology discordance, the basis for Comment 13 below, is extensive). Established in vitro assays have known artifacts and known boundary conditions. Clinical biomarkers themselves are imperfect proxies for clinical outcomes. The four-pillar framework reads most coherently when each method, including each traditional comparator, is itself evaluated for its own validation status, biological relevance, technical characterization, and fit-for-purpose. When a NAM is compared to a traditional method and the two disagree, the disagreement should be evaluated on the merits of both methods rather than reflexively resolved in favor of the historical method. Suggested clarifying language:

> *Comparison of a NAM to a traditional comparator method is most informative when the validation status, biological relevance, technical characterization, and fit-for-purpose of the traditional method are themselves explicit. Where a NAM and a traditional method disagree, sponsors should provide evidence sufficient for FDA to evaluate which method is more reliable for the specific question and context of use, rather than presuming the traditional method to be definitive.*

Comment 7. Clarify weight-of-evidence considerations in rare disease, patient-derived, and other low-data, mechanistically inaccessible settings.

Reference: Sections I, II, and III.D.

In rare and ultra-rare diseases, large clinical datasets may be unavailable, natural-history data may be limited, and animal models may not reproduce the human disease. The same is true for diseases where animal models are known to be poor predictors of human biology for mechanistic reasons — certain neurodegenerative diseases, several immune-mediated conditions, and a range of diseases involving uniquely human cellular biology.

Recommendation: clarify that in low-data and mechanistically inaccessible settings, NAMs may play a greater role in the overall weight of evidence supporting drug development decisions when supported by strong scientific rationale, appropriate technical characterization, human biological relevance, and a clearly defined context of use.

> *FDA recognizes that in rare diseases, ultra-rare diseases, and other low-data or mechanistically inaccessible settings — including diseases for which animal models are known to be poor predictors of human biology — traditional sources of evidence may be limited or unavailable. In such cases, NAMs may play a greater role in the overall weight of evidence supporting drug development decisions when supported by a strong scientific rationale, appropriate technical characterization, human biological relevance, and a clearly defined context of use.*

We also suggest the final guidance encourage sponsors, in such contexts, to explain how NAM endpoints connect to clinical biomarkers, pharmacodynamic measures, safety monitoring, dose selection, or patient stratification — bridging nonclinical NAM evidence to trial design rather than treating NAMs as isolated preclinical assays. This recommendation is consistent with comments already filed on this docket (see, e.g., FDA-2025-D-6131-0020).

Part C. Operational and Cross-Cutting

Comment 8. Publish a priority matrix of drug class × NAM-type pairings with worked validation examples and accelerated ISTAND consideration.

Reference: Cross-cutting; expanding on the structure of Section III.

Recommendation: publish, alongside the final guidance, a priority matrix of drug-class-by-NAM-type pairings where validation evidence is most operationally mature, paired with worked examples and accelerated ISTAND consideration for submissions targeting prioritized pairings.

8A. Worked validation examples.

Worked examples are well-precedented in FDA guidance documents and serve a function that principles alone cannot: they show sponsors how the principles are intended to interact in a realistic submission, and they provide reviewers with a shared reference point. A useful worked example would walk through a

hypothetical NAM submission and articulate, for each of the four pillars, the COU and decision(s) the data are intended to inform; the features establishing Human Biological Relevance, with explicit attention to how the choice of biological model relates to the COU; the elements of Technical Characterization expected at submission; and the Fit-for-Purpose argument.

The final guidance should include at least one worked example for a traditional NAM category, at least one for an integrated platform-level or longitudinal NAM, and at least one illustrating how NAM evidence can support reconsideration of a drug development program where prior nonclinical findings in animal models may not have been human-relevant.

8B. Priority matrix of drug class × NAM type.

Hepatotoxicity × hepatic MPS is the natural first cell, given the prevalence of drug-induced liver injury in late-stage clinical and post-market failures, the relative maturity of human hepatic NAMs, and FDA's acceptance of an Organ-Chip into ISTAND for DILI prediction (CDER, September 24, 2024). The Liver-Chip identified the majority of hepatotoxic drugs that caused liver injury in patients in a head-to-head study (Ewart et al., Nature Communications). Cardiotoxicity × cardiomyocyte MPS is a natural second cell given the Comprehensive In Vitro Proarrhythmia Assay (CiPA) precedent. Specific exposure-route contexts — topical/dermal absorption, first-pass hepatic metabolism, intrathecal and nasal isolated-organ models — are additional natural cells where the systemic complexity of in vivo studies adds limited value relative to the human-relevant in vitro approach.

Schematic of the proposed pathway:

```
   1. Priority pairing identified
      (e.g., hepatotox × hepatic MPS; cardiotox × cardiomyocyte MPS;
       topical absorption; first-pass hepatic metabolism; intrathecal;
nasal)
         ↓
   2. Reference protocol published
      (FDA + ICCVAM + IQ MPS Affiliate develop jointly)
         ↓
   3. Sponsor submits against priority protocol
         ↓
   4. Accelerated ISTAND consideration
         ↓
   5. Qualification or structured RFE (see Comment 10)
```

Submissions targeting prioritized pairings should be eligible for accelerated consideration under the existing ISTAND program. Sponsors developing methods in less-prioritized pairings should remain eligible for standard ISTAND review without prejudice; the priority matrix is intended to coordinate field-wide effort, not to restrict it.

Comment 9. Add a structured NAM Context-of-Use briefing template to support early engagement.

Reference: Section II, lines 76–82.

Recommendation: include in the final guidance an optional structured NAM briefing template (see Appendix A) that sponsors can bring to a pre-IND or early consultative meeting to reduce ambiguity, especially for small sponsors and first movers.

A template of this kind would help review divisions provide more consistent advice across submissions and across the range of NAM categories the guidance is intended to support. We propose the appendix at the end of this comment as a starting point.

Comment 10. Operationalize structured Requests for Information for NAM-supported submissions.

Reference: Cross-cutting; references to existing CDER Information Request mechanism.

CDER review divisions already issue Information Request (IR) letters to sponsors during IND review to surface specific scientific or technical concerns. For NAM-supported submissions, however, IRs are issued informally and without structured COU specification, which can leave sponsors uncertain about what additional evidence would resolve the concern.

Recommendation: convert the existing CDER Information Request mechanism, for NAM-supported submissions, into a structured form that specifies the four-pillar element at issue and the least burdensome additional information that would resolve the concern.

Schematic of the proposed mechanism:

```
Sponsor submits NAM evidence
     ↓
FDA review of submission
     ↓
Evidence sufficient?
     ↓
  YES → Submission proceeds
  NO  → Structured Information Request issued, specifying:
          • Pillar at issue (COU / HBR / TC / FFP)
          • Specific scientific or technical concern
          • Least burdensome additional information needed
            (further NAM data, bridging data, orthogonal assays,
             revised COU, or traditional nonclinical data)
                ↓
```

Sponsor responds with specified information

Suggested language:

> *When FDA determines that submitted NAM data are not sufficient for the sponsor's proposed context of use, FDA should, where practicable, issue a structured Information Request that identifies the specific scientific or technical concern, indicates which of the four validation considerations the concern relates to, and describes the least burdensome additional information that would resolve it. Such additional information may include further NAM data, bridging data, orthogonal assays, modified COU framing, or traditional nonclinical data where necessary.*

The reviewer time required to issue a structured IR is small relative to the time saved by reducing iteration cycles between sponsor and review division.

Comment 11. Costs, cost savings, and incentive structures under Executive Order 14192.

Reference: Federal Register notice, Background, paragraph following Section I.

Recommendation: the cost analysis accompanying the final guidance should explicitly reflect (i) cost savings from accepting fit-for-purpose NAM data without prior formal validation, (ii) larger savings from integrated platform NAMs that substitute one experiment for several, (iii) savings from worked examples, briefing templates, structured IRs, reviewer rubrics, and the cross-species discordance registry, (iv) tiered incentive structures for sponsors contributing NAM data to public evidence pools, (v) Phase 0 bridge pathways for NAM-supported submissions, and (vi) a public NAM-performance feedback registry.

First, the clarification at lines 34–40 — that a NAM does not necessarily need to be formally validated to be considered for review — is the single most important cost-relevant element of the guidance. The historical pattern under which sponsors withhold NAM data from submissions out of concern of rejection has imposed substantial costs on the system. We strongly support retaining this language.

Second, the cost savings of integrated, platform-level NAMs are materially larger than those of single-endpoint NAMs of equivalent quality, because a single integrated experimental design substitutes for what would otherwise be a series of separate studies. The clarifications proposed in Comments 1, 2, and 6 would directly enable sponsors to capture these savings.

Third, worked examples (Comment 8), briefing templates (Comment 9), structured Information Requests (Comment 10), reviewer rubrics and specialist reviewer pairings (Comment 14), and the cross-species discordance registry (Comment 13) generate cost savings together by reducing iteration cycles. The cost of sponsor uncertainty is borne disproportionately by smaller sponsors, academic developers, rare-disease groups, and patient-cell translational programs.

Fourth, FDA should consider tiered regulatory incentive structures: accelerated review timelines, reduced PDUFA fees, or modest priority-review benefits for sponsors who voluntarily contribute de-identified,

standardized NAM datasets to a public or cross-sponsor evidence pool. This mirrors the Priority Review Voucher mechanism and would not require new statutory authority.

Fifth, FDA should consider whether NAM-supported submissions could be paired with Phase 0 exploratory IND pathways under 21 CFR 312 Subpart E and the 2006 FDA Phase 0 guidance, creating a graduated review model where NAM evidence supports microdose human studies and the resulting human data informs subsequent traditional review.

Sixth, FDA may also consider supporting development, within the Complement-ARIE NAM Data Hub being established by NIH, of a public NAM-performance feedback registry that tracks, in a de-identified manner, cases in which NAM predictions did or did not align with subsequent clinical outcomes.

Comment 13. Support development of a public registry of documented cross-species toxicology discordances.

Reference: Cross-cutting; relates to Sections III.B and III.D.

A substantial fraction of compounds discontinued during preclinical development are discontinued because of findings in animal models that may or may not translate to human biology. Conversely, a substantial fraction of clinical and post-market drug failures involve toxicity mechanisms not predicted by the animal models used in preclinical development. The underlying scientific knowledge exists in the toxicology literature but is dispersed, inconsistently curated, and not readily accessible as a basis for regulatory decision-making.

Recommendation: FDA should support, within the Complement-ARIE NAM Data Hub, the development of a curated public registry of documented cross-species toxicology discordances — compounds and compound classes where animal findings did not translate to humans, or where human-specific toxicity was missed by animal models — with mechanism documented to the extent known.

Such a registry would:

- Provide sponsors with a defensible scientific basis for fit-for-purpose NAM claims that target species-specific concerns.
- Provide FDA reviewers with cumulative evidence for evaluating NAM-supported submissions in drug classes where cross-species discordance is well-documented.
- Support the worked-example approach proposed in Comment 8 by providing a concrete population of cases.
- Contribute to international harmonization by providing a shared reference resource.

The registry need not be created de novo: the relevant primary literature exists, and a coordinated curation effort under the Complement-ARIE NAM Data Hub would produce a usable v0 within a tractable timeframe.

Comment 14. Strengthen reviewer capacity through structured evaluation rubrics and NAM specialist reviewer pairings.

Reference: Cross-cutting; Section II.

The four-pillar framework establishes principles but does not specify how reviewers should weight evidence within and across pillars when evaluating a NAM-supported submission. The result, in practice, is that evaluation depends substantially on individual reviewer experience with the specific NAM category in question.

Recommendation: develop structured evaluation rubrics for prioritized COUs, and formally extend CDER's existing specialty-reviewer pairing model (used for biologics, gene therapy, and cell therapy) to NAM-supported submissions.

14A. Structured evaluation rubrics for prioritized COUs.

For the prioritized drug-class-by-NAM-type pairings identified in Comment 8 (beginning with hepatotoxicity × hepatic MPS), FDA should develop, with input from the NAM developer community and from international counterparts, structured evaluation rubrics that specify what "adequate evidence" looks like for each pillar at each level of intended regulatory weight. The closest existing precedent is FDA's bioequivalence guidance for generic drugs, which provides reviewers with explicit quantitative criteria rather than leaving evaluation to reviewer judgment alone. Structured rubrics would not eliminate reviewer judgment — they would anchor it, improve consistency, and reduce the structural tendency to err on the side of rejection in the face of unfamiliar evidence types.

14B. NAM specialist reviewer pairings within existing review divisions.

CDER already pairs primary reviewers with specialty reviewers for complex submission categories — most notably for biologics, gene therapies, and cell therapies. The final guidance — or an accompanying CDER policy — should formally extend this model to NAM-supported submissions, by establishing a NAM specialist reviewer capability within existing pharm/tox review divisions, with training in microphysiological systems, computational toxicology, sensor-based and AI/ML-enabled measurement methods, and longitudinal data analysis. Specialist reviewers would be paired with the primary pharm/tox reviewer rather than replacing them; this approach strengthens existing review capacity rather than creating parallel structures. Initial recruitment could draw from NIH ORIVA, academic NAM developers, and industry NAM scientists. The Complement-ARIE Technology Development Centers and the IQ Consortium Microphysiological Systems Affiliate are natural partners for joint training and exchange.

Closing

Precigenetics appreciates the opportunity to contribute to the development of this important guidance. We support FDA's broader effort to facilitate the integration of human-relevant nonclinical methods into drug development, and we are committed to engaging constructively with CDER as our own NAM-based programs progress toward future regulatory submissions.

The four-pillar framework established in this draft is, in our view, the right framework. The fourteen comments above are intended to extend that framework to fully accommodate the emerging class of integrated, measurement-rich, human-relevant NAMs that will increasingly characterize the nonclinical landscape over the next several years. None of our comments asks FDA to lower the evidentiary standard. Each asks FDA to make the standard more precise, more actionable, and more consistently applied — through clearer definitions, structured rubrics, worked examples, reference standards, briefing templates, structured feedback mechanisms, and cumulative-evidence infrastructure.

Finally, we note that just as every clinical trial, every nonclinical study, and every drug development program is evaluated on its own facts, **every NAM — and every NAM-supported submission — should likewise be evaluated case by case, on the specific COU it is intended to inform and the specific weight of evidence it offers**. The four-pillar framework, made more precise through the additions proposed above, is precisely the instrument that enables this kind of disciplined, evidence-led, case-by-case evaluation at scale.

We welcome the opportunity to discuss any aspect of these comments with the Agency, and we look forward to seeing the final guidance.

Respectfully submitted,

Precigenetics, Inc.
San Carlos, California
nams@precigenetics.com
May 18, 2026

Appendix A. NAM Context-of-Use Briefing Template

The following is offered as a starting point for the structured briefing template proposed in Comment 9. It is intentionally generic and technology-neutral, and is designed to be useable across organoid, organ-chip, in silico, sensor-based, optical, longitudinal, and integrated NAM categories.

Element	Information to include
Regulatory decision context	The specific decision the NAM is intended to inform — for example, dose selection, clinical safety monitoring, species selection, mechanistic interpretation of an adverse event, or support for a weight-of-evidence package.
Context of use	Drug class, indication, development stage, organ system, endpoint(s), and whether the NAM is exploratory, supportive, confirmatory, or intended as an alternative to a traditional method.
Biological model	Cell type, tissue model, donor source, disease relevance, species, maturity, phenotype, genetic features, and known biological variability. Justification for cellular composition based on mechanistic relevance to the toxicity or pharmacology of interest.
Exposure system	Dosing method, duration, frequency, perfusion or flow conditions, device materials, exposure gradients, exposure history, and compatibility with the test article.
Measurement modality	Readouts, instrumentation, calibration, reproducibility, sensitivity and specificity (where applicable), throughput, measurement-induced perturbation controls, assay-induced artifact controls, and raw and processed data outputs.
Computational analysis	Feature generation, statistical or machine-learning model used, validation data, model versioning, uncertainty estimation, locked analysis plan, and interpretability relevant to the COU.
Controls and benchmarks	Positive and negative controls, reference compounds, reference cell standards (where available), orthogonal assays, comparator data (where applicable), and acceptance criteria.
Performance characteristics	Stratified performance by drug class and toxicity mechanism, throughput at which performance was obtained, pre-registration status of validation studies, and complete reporting of pre-specified endpoints.
Data package	Summary of raw data, processed data, metadata, quality-control thresholds, missing-data handling, and audit trail or data-provenance plan.

Element	Information to include
Limitations	Known failure modes, boundary conditions, unsupported uses, and any proposed bridging studies.
Fit-for-purpose rationale	How the NAM evidence will contribute to the overall decision, and why it is reliable enough for that purpose under the four-pillar framework.