

# Stage-1 Logic and Architecture Package

Dual-Platform mRNA Therapy for X-Linked ALD (LNP + EV, Peripheral-First → CNS)

Method2Model

## **Abstract**

This document summarizes the complete Stage-1 outputs for the ALD dual-platform mRNA project within the Method2Model framework. It includes: (i) Use-Case Confirmation Lock (Scope Lock), (ii) an Assumptions Map / Log, (iii) an Input Specification Sheet (ISS), (iv) a reviewable Model Architecture summary, (v) Scenario Pack v1, and (vi) a Verification Plan with architecture-to-code traceability, acceptance tests, and version-lock rules. Together, these components form a review-ready package that can be shared internally (team, collaborators) or externally (sponsors, reviewers) prior to Stage-2 implementation and Stage-3 verification.

## Contents

<b>1</b>	<b>Use-Case Confirmation Lock (Scope Lock)</b>	<b>3</b>
1.1	Chosen Use Cases . . . . .	3
1.2	Scope Boundaries . . . . .	3
1.3	Success Criteria (Stage 1) . . . . .	4
<b>2</b>	<b>Assumptions Map / Log (MRR)</b>	<b>4</b>
2.1	Assumption Table . . . . .	4
2.2	Assumption Maintenance and Versioning . . . . .	6
<b>3</b>	<b>Input Specification Sheet (ISS)</b>	<b>6</b>
3.1	Delivery and Exposure Inputs . . . . .	6
3.2	Expression and Targeting Inputs . . . . .	7
3.3	Disease / PD and Clinical Inputs . . . . .	7
3.4	Safety and CMC Inputs . . . . .	8
<b>4</b>	<b>Model Architecture (Reviewable Summary)</b>	<b>9</b>
4.1	Block Overview . . . . .	9
4.2	Inputs → States → Outputs . . . . .	9
4.3	Fixed vs Fitted Components . . . . .	10
<b>5</b>	<b>Scenario Pack v1 (Definition-Level)</b>	<b>10</b>
5.1	Baseline Scenario (S0) . . . . .	10
5.2	Stress Scenario S1: CNS Delivery-Limited . . . . .	11
5.3	Stress Scenario S2: Overexpression and Targeting Risk . . . . .	11
5.4	Stress Scenario S3: Immune-Sensitive Repeat Dosing . . . . .	12
5.5	Stress Scenario S4: CMC-Driven Potency Variability . . . . .	12
5.6	Stress Scenario S5: Translational Generalization . . . . .	12
<b>6</b>	<b>Verification Plan and Traceability</b>	<b>13</b>
6.1	Architecture-to-Code Traceability . . . . .	13
6.2	Acceptance Tests . . . . .	13
6.3	Scenario Equivalence Checks . . . . .	14
6.4	Version-Lock Rules . . . . .	14
6.5	Approval Gate . . . . .	14

# 1 Use-Case Confirmation Lock (Scope Lock)

## 1.1 Chosen Use Cases

The project explicitly commits to Method2Model Use Cases 1–7 for this Stage-1 architecture:

1. *Before you run it:* scan for protocol blind spots (“what breaks first?”).
2. *Before you commit to sample size:* validate power against real-world noise (at least at the logical/structural level in Stage 1).
3. *Before you measure everything:* identify low-information data and avoid non-essential end-points.
4. *Before you lock a regimen:* stress-test robustness under adherence and biological variability.
5. *Before you generalize:* prevent “works here, fails there” surprises across phenotypes (AMN vs cALD), routes, and platforms.
6. *Before you change a pathway:* simulate CMC, delivery, and safety bottlenecks before large-scale rollout.
7. *Before you submit it:* avoid the IRB/grant/sponsor rewrite loop by providing a coherent, defensible logic and evidence chain.

## 1.2 Scope Boundaries

### In-scope for Stage 1.

- Mechanistic, block-structured architecture (Blocks A–G) that describes:
  - Delivery and exposure (PBPK/transport) for LNP and EV platforms, including FUS-modulated CNS access.
  - Expression and peroxisomal targeting of ALDP from mRNA.
  - Disease and pharmacodynamics: ALDP  $\rightarrow$  peroxisomal  $\beta$ -oxidation  $\rightarrow$  VLCFA  $\rightarrow$  biomarkers (VLCFA, NFL, DTI/Rotarod, Loes) and functional state.
  - Safety and immunogenicity (innate cytokines, complement, hepatotoxicity, anti-PEG).
  - CMC and batch variability (LNP/EV quality attributes and potency).
  - Uncertainty and scenario analysis logic (but not necessarily numerical sampling yet).
  - Defensibility and decision rules (Go/No-Go/Pivot surfaces).
- A first-pass Scenario Pack (baseline + stress scenarios) at the definition level (no coding yet).
- A Verification Plan that describes how Stage 3 will demonstrate that the code matches this architecture.

### Out-of-scope for Stage 1.

- Concrete numerical calibration, parameter estimation, or statistical fitting.
- Regulatory-grade power calculations (these are structurally prepared but numerically deferred).
- Final clinical protocol design; the focus is on model-logic and simulation readiness, not on protocol submission.
- Implementation details of numerical solvers and optimization algorithms.

### 1.3 Success Criteria (Stage 1)

Stage 1 is considered successful if the following criteria are met:

- **Logical completeness:** All high-impact decisions (platform/route choice, AMN vs cALD focus, dosing regimen, key endpoints, primary bottlenecks) are explicitly represented somewhere in the model architecture (Blocks A–G).
- **Traceability:** For each Use Case 1–7, there is a clear path from input data and assumptions, through internal states, to outputs and decisions.
- **Reviewability:** A domain expert (clinician, translational scientist, or CMC lead) can read this package and understand:
  - where each assumption enters,
  - how each assumption can be updated with new evidence, and
  - how scenario results would be interpreted.
- **Verification readiness:** There is a concrete Verification Plan (Section ??) that specifies how Stage 3 will prove code–architecture equivalence, including acceptance tests and scenario equivalence checks.

The approval gate for moving to Stage 2 is formal acceptance of: (i) the Use-Case Confirmation Lock, (ii) the Assumptions Map, (iii) the ISS, (iv) the Architecture summary, (v) Scenario Pack v1, and (vi) the Verification Plan.

## 2 Assumptions Map / Log (MRR)

This section lists key modeling assumptions, each with an identifier, category, evidence level, impact, sensitivity, and a suggested validation path.

### 2.1 Assumption Table

ID	Category	Assumption	Evidence	Impact	Validation Path
A1	Delivery/CNS	Baseline IV-LNP (without FUS) produces negligible direct CNS exposure compared to peripheral exposure.	Moderate	High	Compare model predictions with published biodistribution data; if future tracer studies show higher CNS exposure, update PBPK parameters.
A2	Delivery/CNS	FUS increases effective BBB permeability by a stochastic factor in the range 5–20 for a limited window $\Delta t$ after sonication.	Moderate	High	Parameterize $M_{\text{FUS}}$ from FUS-BBB opening literature; validate against pilot imaging or tracer studies.
A3	CNS cell uptake	EV-IN and EV+FUS routes yield higher uptake in astrocytes than in neurons and oligodendrocytes; oligodendrocyte uptake is modest but non-zero.	Low–Moderate	High	Quantify cellular uptake in ex vivo or in vivo EV biodistribution studies; adjust cell-type uptake functions $f_{k,r,c,\text{cell}}$ .

B1	Targeting	A peroxisomal ALDP level of 1–3.5× wild-type is sufficient to restore most $\beta$ -oxidation capacity without excessive mislocalization.	Moderate	High	Combine in vitro ALDP expression/targeting data with VLCFA rescue experiments; refine window as more data accumulate.
B2	Targeting/Stress	ALDP expression > 3–4× wild-type increases the risk of mislocalization and peroxisome/ER stress.	Low–Moderate	Medium	Monitor stress markers (e.g. CHOP, ATF4) in overexpression experiments; revise overexpression-risk region or stress functions.
C1	PD/VLCFA	Partial restoration of peroxisomal $\beta$ -oxidation in fibroblasts/neurons gives a more-than-linear reduction in VLCFA (steep response at low ALDP).	Moderate	High	Fit the function $g(P_{\text{ALDP,peri}}/P_{\text{ALDP,WT}})$ to in vitro data; update upon availability of new rescue data.
C2	PD/Progression	Cumulative CNS VLCFA burden and inflammation jointly drive axonal injury and, in humans, NFL and Loes dynamics.	Moderate	High	Use longitudinal human data (natural history, HSCT, gene therapy) to constrain $G_{\text{NFL}}$ and $G_{\text{Loes}}$ .
C3	Mouse→Human	Mouse PD relations (VLCFA → axonal injury) are informative but not numerically identical to humans; a scaling factor and uncertainty are required.	Moderate	Medium	Encode explicit uncertainty for translational scaling; where possible, use human metabolic and imaging data to constrain.
D1	Safety/Innate	Innate cytokine and complement responses to LNP and EV can be represented as carrier- and dose-dependent spikes, modulated by nucleoside chemistry and impurities.	High	High	Check against known LNP/EV safety datasets; update dose-response curves as preclinical and clinical safety data become available.
D2	Safety/Repeat dose	Anti-PEG or anti-lipid antibodies after repeated LNP dosing primarily affect clearance and may also contribute to hypersensitivity.	Moderate	Medium–High	Measure anti-PEG titers and associate with PK changes and clinical signs in repeat-dose studies.
E1	CMC/Potency	Batch-to-batch potency variability for in-spec LNP is on the order of 10–30%; for EV, variability in mRNA copies per $10^{10}$ EV can be 30–60%.	Low–Moderate	High	Quantify potency distributions in internal production; refine prior distributions for $\text{Pot}_{k,\text{batch}}$ .
E2	Stability	Under recommended storage ( $-80^{\circ}\text{C}$ ), potency decay is slow (e.g. $\lesssim 20\%$ over months), but freeze-thaw cycles add stepwise losses.	Moderate	Medium	Perform stability studies with potency assays; refine $k_{\text{decay}}$ and $f_{\text{FT}}$ .
F1	Uncertainty	Prior distributions for key parameters (uptake, PD, safety, CMC) can be represented by broad, biologically plausible ranges, not point estimates.	High	High	Maintain the Assumptions Log and update ranges as new evidence is incorporated; document changes with version IDs.
G1	Endpoint hierarchy	The selected endpoints (VLCFA, targeting, PD readouts, safety markers) are sufficient to support Go/No-Go/Pivot decisions without an excessive number of low-information measures.	Moderate	Medium	Re-evaluate via value-of-information analysis in Stage 2; drop endpoints that do not change decisions under realistic scenarios.

## 2.2 Assumption Maintenance and Versioning

Each assumption ID will be:

- Referenced in the codebase (e.g. as comments or metadata),
- Linked to specific parameters or functions, and
- Versioned in an Assumptions Log (MRR) file with entries (**AssumptionID**, description, evidence, change history).

Updates to assumptions after Stage 1 will require:

1. Documentation of new evidence,
2. Re-labeling of architectural versions if assumptions affect formulas or structure,
3. Re-running relevant scenarios and acceptance tests.

## 3 Input Specification Sheet (ISS)

The ISS defines input variables, units, timing, typical ranges, and expected missingness/noise. It is organized by domain.

### 3.1 Delivery and Exposure Inputs

Variable	Definition	Units	Typical Range	Missingness/Noise
Dose <sub>LNP</sub>	mRNA dose per kg for LNP injections	mg/kg or $\mu\text{g/kg}$	0.01–1 mg/kg	Low missingness; measurement error < 10%.
Dose <sub>EV</sub>	EV dose per kg (particles or mRNA copies)	particles/kg or copies/kg	$10^9$ – $10^{12}$ particles/kg	Medium uncertainty in absolute particle count.
Route	Administration route (IV, IN, IV+FUS, EV+FUS)	categorical	<i>n/a</i>	No missingness; design variable.
FUS_params	FUS center frequency, pressure, duration, sonication pattern	structured	device-specific operational ranges	Documented per session; no missingness in protocol.
Body weight	Animal/human body weight at dosing	kg	species- and age-dependent	Measurement error $\sim 2$ –5%.
PBPK parameters	Volumes, flows, mixed clearances, baseline $P_{\text{BBB}}$		literature- and species-based ranges	High uncertainty for CNS-related parameters; encoded as priors.

### 3.2 Expression and Targeting Inputs

Variable	Definition	Units	Typical Range	Missingness/Noise
----------	------------	-------	---------------	-------------------

mRNA_chemistry	Cap, nucleoside modification, UTRs, codon-usage profile	categorical/discrete sets		// Qualitative; encoded in parameter sets.
PEX19/PEX3_levels	Per-cell or per-tissue levels of PEX19/PEX3	relative units	approx. 0.5–1.5× reference	Often not directly measured; inferred from literature.
IF_ALDP_coloc	ALDP–peroxisome colocalization score	0–1	0 (mutant) to near 1 (WT-like)	Image-analysis noise; replicate variability.
PTS_import	PTS1/PTS2 import efficiency	0–1	0.2–1	Assay noise; dependent on cell type.
Protease_prot	Protease protection fraction for ALDP	0–1	0.2–1	Experimental error, fractionation variability.

### 3.3 Disease / PD and Clinical Inputs

Variable	Definition	Units	Typical Range	Missingness/Noise
Phenotype	AMN vs very-early cALD vs early cALD	categorical	<i>n/a</i>	Assigned based on clinical criteria.
VLCFA_plasma	Plasma VLCFA concentration (e.g. C26:0)	$\mu\text{mol/L}$ or normalized ratio	disease-dependent	Analytical CV typically < 10%.
VLCFA_CNS	CNS VLCFA concentration (e.g. C26:0)	$\mu\text{g/g}$ tissue or ratio	model- and data-dependent	Invasive or limited sampling; higher uncertainty.
NfL	Neurofilament light chain (serum/CSF)	pg/mL	elevated vs controls	Assay CV typically < 10%; biological variability.
Loes_score	MRI-based severity score	Loes dimensionless	9–34	Inter-rater variability; only in human branch.
Rotarod_score	Rotarod performance metric (time, distance)	s or arbitrary	baseline and follow-up	Behavioral variability; learning effects.
DTI_metrics	DTI-derived indices (FA, MD, RD)	dimensionless	0–1 or unitless	Scanner noise, processing variability.

### 3.4 Safety and CMC Inputs

Variable	Definition	Units	Typical Range	Missingness/Noise
ALT, AST	Serum values	ALT/AST U/L	normal to > 3× ULN	Lab assay CV < 10%.
Cytokines	IL-6, TNF- $\alpha$ , etc.	pg/mL	baseline to high spikes	High dynamic range, often log-normal.

Complement markers	sC5b-9	units/mL	baseline to elevated	Assay noise; per-sample variability.
Anti_PEG_titer	Antibody against PEG/lipid	titer relative units	0 to high	ELISA-like variability; time-dependent.
LNP_CQA	Size, PDI, RNA integrity, endotoxin, pH, osmolality	EE%, mixed	spec-dependent	QC variability, typically moderate.
EV_CQA	Particle size, EV, competency	count, mRNA/10 <sup>10</sup> translation	production-dependent	Higher uncertainty; batch-to-batch variation.

## 4 Model Architecture (Reviewable Summary)

A full Stage-1 architecture for Blocks A–G was previously specified. Here we provide a condensed, reviewable summary emphasizing: inputs  $\rightarrow$  states  $\rightarrow$  transitions  $\rightarrow$  outputs, dependency mapping, and the distinction between fixed versus fitted components.

### 4.1 Block Overview

- Block A: Delivery & Exposure (PBPK/Transport).
- Block B: Expression & Peroxisomal Targeting.
- Block C: Disease/PD (Mouse branch and Human-projected branch).
- Block D: Safety & Immunogenicity.
- Block E: CMC & Batch Variability.
- Block F: Uncertainty & Scenario Engine.
- Block G: Defensibility & Decision Logic.

### 4.2 Inputs $\rightarrow$ States $\rightarrow$ Outputs

#### Block A.

- Inputs: Dose schedules, routes, FUS parameters, PBPK parameters, batch potency.
- States: Compartmental carrier concentrations  $C_{k,c}(t)$ , uptake states  $U_{k,c,\text{cell}}(t)$ , BBB permeability  $P_{\text{BBB}}(t)$ .
- Outputs: Exposure profiles for each tissue/cell type; CNS exposure classification.

#### Block B.

- Inputs: Cellular uptake from Block A; mRNA chemistry; PEX19/PEX3 capacity.
- States: mRNA  $m(t)$ , total ALDP  $P_{\text{ALDP,tot}}(t)$ , peroxisomal ALDP  $P_{\text{ALDP,peri}}(t)$ , mislocalized ALDP  $P_{\text{ALDP,mis}}(t)$ , targeting efficiency  $\theta_{\text{target}}(t)$ .
- Outputs: ALDP peroxisomal levels, composite  $S_{\text{target}}$ , labels for therapeutic vs sub-therapeutic vs overexpression-risk regions.



### Block C.

- Inputs: Peroxisomal ALDP from Block B; phenotype; species.
- States:  $\beta$ -oxidation capacity  $V_\beta(t)$ , VLCFA concentrations, CNS inflammation, axonal injury indices.
- Outputs: VLCFA trajectories; DTI/Rotarod outcomes (mouse); NfL and Loes (human-projected); Aim-1/Aim-2 success flags.

### Block D.

- Inputs: Dose history; carrier and mRNA properties; immune background.
- States: Cytokine responses, complement activation, ALT/AST, anti-PEG titers.
- Outputs: Safety violation flags and risk profiles.

### Block E.

- Inputs: QC data for LNP/EV batches; storage and handling history.
- States: Quality attributes, potency multipliers, stability decay.
- Outputs: Effective potencies, batch classification (in-spec vs borderline vs out-of-spec).

### Block F.

- Inputs: Parameter prior distributions; scenario definitions.
- States: Scenario sampling indices and parameter instantiations.
- Outputs: First-break maps, robustness curves, generalization matrices.

### Block G.

- Inputs: Outputs from Blocks A–F.
- States: Decision summaries, chain-of-evidence structures.
- Outputs: Go/No-Go/Pivot decisions; reviewer-ready figures and tables; endpoint hierarchies.

## 4.3 Fixed vs Fitted Components

### • Fixed (structural) components:

- Block decomposition (A–G) and their interconnections.
- The fact that mouse and human-projected PD are separate branches, with Loes present only in the human branch.
- Representing key mechanisms (delivery, targeting, VLCFA, safety, CMC) as explicit states, not as opaque parameters.

### • Fitted / tunable components:

- Parameter values and functional forms within the blocks: PBPK rates,  $g$  and  $h$  in Block C, safety dose–response in Block D, potency distributions in Block E.
- Scenario-specific scalings (e.g. translational scaling from mouse to human in Block C).

## 5 Scenario Pack v1 (Definition-Level)

Scenario Pack v1 defines a baseline scenario and a set of stress scenarios at the definition level (without implementation), including: parameter ranges, routes/platforms, and evaluation metrics.

### 5.1 Baseline Scenario (S0)

**Label: S0 – Baseline AMN, Peripheral-first.**

- Phenotype: AMN-like (peripheral VLCFA burden, early or absent CNS lesions).
- Species: mouse for preclinical branch; human-projected for translation.
- Platform and route:
  - IV LNP for peripheral ALDP restoration,
  - Optional EV-IN or EV+FUS exploratory arm for CNS.
- Dosing: weekly LNP for 4 weeks at a mid-range dose (e.g. 0.3–0.5 mg/kg).
- CMC: in-spec LNP batch; no major stability issues.
- Evaluation metrics:
  - Aim-1: in vitro VLCFA reduction and targeting score.
  - Aim-2: plasma C26 reduction  $\geq 35\%$  at week 4; CNS C26 reduction  $\geq 25\%$  if CNS-targeting arm present.
  - Safety: no ALT/AST  $> 3 \times$  ULN; no CRS  $\geq$  Grade 2.
  - Translational: projected stabilization of VLCFA and NfL in the human-projected branch for AMN.

### 5.2 Stress Scenario S1: CNS Delivery-Limited

**Label: S1 – cALD, Low CNS Uptake.**

- Phenotype: very-early cALD.
- Platform/route: IV LNP  $\pm$  EV-IN; FUS not yet used.
- Assumption stress: CNS uptake parameters are set at the lower end of plausible ranges.
- Questions:
  - Does any realistic combination of LNP and EV-IN achieve sufficient CNS ALDP to cross PD thresholds in Block C?
  - Is a pivot to FUS-based strategies necessary for cALD?
- Metrics: CNS VLCFA reduction, NfL/Loes trend (human-projected), first-break mechanism (likely CNS delivery).

### 5.3 Stress Scenario S2: Overexpression and Targeting Risk

**Label: S2 – High Dose, Potential Overexpression.**

- Phenotype: AMN or cALD.
- Platform: LNP and/or EV at higher doses.
- Assumption stress: Parameter combinations that push  $P_{\text{ALDP,tot}}$  beyond  $3\text{--}4\times$  wild-type while keeping peroxisomal capacity high.
- Questions:
  - How often does the system enter the overexpression-risk region?
  - What is the trade-off between additional VLCFA benefit and stress risks?
- Metrics: fraction of scenarios in overexpression-risk region; stress markers (if modeled); safety signals in Block D.

### 5.4 Stress Scenario S3: Immune-Sensitive Repeat Dosing

**Label: S3 – Immune-High, Repeat LNP.**

- Phenotype: AMN or cALD.
- Platform: repeat-dose LNP with high innate sensitivity and a tendency to generate anti-PEG antibodies.
- Assumption stress: steep innate response curves and rapid accumulation of  $Ab_{\text{PEG}}(t)$ .
- Questions:
  - How quickly does safety become dose-limiting?
  - How does accelerated clearance alter efficacy over time?
- Metrics: safety violation rates, effect on exposure and PD outcomes; first-break mechanism (safety vs efficacy).

### 5.5 Stress Scenario S4: CMC-Driven Potency Variability

**Label: S4 – Batch Drift and Stability.**

- Platform: LNP and EV with batch-to-batch potency variability at high end of realistic ranges.
- Assumption stress: potency multipliers  $\text{Pot}_{k,\text{batch}}$  vary widely; some batches borderline in-spec; storage time long with multiple freeze–thaw cycles.
- Questions:
  - How sensitive are outcomes to batch variability?
  - What minimum potency specification is needed for robust success rates?
- Metrics: robustness curves over potency distributions; fraction of scenarios failing due to CMC constraints.

## 5.6 Stress Scenario S5: Translational Generalization

**Label:** S5 – Mouse to Human Projection.

- Input: Parameter combinations that achieve success in the mouse branch.
- Task: propagate these parameter combinations (with uncertainty) into the human-projected branch.
- Questions:
  - Under what conditions do mouse successes correspond to human stabilization of VLCFA and NFL?
  - When does Loes score still progress despite biochemical improvements?
- Metrics: projected trends in NFL and Loes vs natural history; classification into “likely stabilizing”, “borderline”, “unlikely to stabilize”.

## 6 Verification Plan and Traceability

This section defines how Stage 3 will demonstrate that the implemented code matches the Stage 1 architecture.

### 6.1 Architecture-to-Code Traceability

**Block IDs and Code Modules.** Each architecture block (A–G) will have a corresponding code module with a stable identifier:

Block ID	Architecture Label	Code Module (example)
A	Delivery & Exposure	<code>m2m_ald.block_A_pbpk</code>
B	Expression & Targeting	<code>m2m_ald.block_B_expression</code>
C	Disease / PD	<code>m2m_ald.block_C_pd</code>
D	Safety / Immunogenicity	<code>m2m_ald.block_D_safety</code>
E	CMC / Batch	<code>m2m_ald.block_E_cmc</code>
F	Uncertainty / Scenario	<code>m2m_ald.block_F_scenarios</code>
G	Defensibility / Decisions	<code>m2m_ald.block_G_decisions</code>

**Assumption IDs in Code.** Each assumption ID (A1, A2, ..., G1) will be:

- Referenced in code comments near the relevant parameter or function definition.
- Listed in a machine-readable metadata file (e.g. YAML/JSON) linking AssumptionID to code symbol names and default values.

### 6.2 Acceptance Tests

**Unit tests per block.**

- **Block A:** verify that simple PBPK toy cases (e.g. single compartment, no FUS) reproduce analytically expected solutions.
- **Block B:** confirm that constant  $k_{\text{in}}$  and  $k_{\text{deg}}$  generate exponential rise/decay in  $m(t)$  and  $P_{\text{ALDP,tot}}(t)$  and that  $P_{\text{ALDP,mis}} = P_{\text{ALDP,tot}} - P_{\text{ALDP,peri}}$ .

- **Block C:** check that in simple synthetic cases, increasing  $P_{\text{ALDP,peri}}$  monotonically decreases steady-state VLCFA.
- **Block D:** test that safety flags are raised when inputs exceed predefined thresholds.
- **Block E:** validate that potency decay and freeze-thaw rules behave as specified.
- **Block F:** confirm that sampling routines honor prior distributions and scenario definitions.
- **Block G:** verify mapping from state trajectories to Go/No-Go/Pivot decisions given artificial data.

#### Integration tests across blocks.

- End-to-end tests for simplified scenarios (e.g. a single simplified S0-like scenario) where approximate qualitative behavior is known (e.g. increased dose  $\rightarrow$  increased ALDP  $\rightarrow$  decreased VLCFA).
- Consistency tests for mouse vs human-projected branches: ensure that changes in PD parameters propagate as expected.

### 6.3 Scenario Equivalence Checks

For each scenario in Scenario Pack v1 (S0–S5), Stage 3 will:

- Implement a YAML/JSON scenario definition (phenotype, route, dose, batch quality, immune background).
- Run the scenario to produce summary metrics: CNS exposure,  $P_{\text{ALDP,peri}}$  distribution, VLCFA changes, safety violations.
- Compare these outputs against:
  - qualitative expectations from the Stage 1 documentation, and
  - any available analytic or semi-analytic checks (e.g. monotonic effects, sign of changes).
- Document and resolve any discrepancies between conceptual expectations and code behavior.

### 6.4 Version-Lock Rules

We define three levels of versioning:

- **Architecture ID:** ARCH-ALD-DUAL-v1.0 for this Stage-1 specification.
- **Formula/Law ID:** FORM-ALD-PD-v1.0, FORM-ALD-PBPK-v1.0, etc., for specific formula sets within blocks.
- **Code ID:** CODE-ALD-DUAL-v0.x.y for each implementation version.

Rules:

- Any structural change in the architecture (new states, new blocks, changes in connectivity) requires incrementing ARCH-ALD-DUAL.
- Any change in functional forms or formulas that affect PD, PBPK, safety, or CMC semantics must increment the relevant FORM-\* ID and trigger re-validation.
- Any code change affecting numerical behavior but not structure or formulas (e.g. solver choice) increments CODE-\* only and requires rerunning acceptance tests.
- The combination (ARCH-ID, FORM-ID, CODE-ID) must be recorded in every simulation run, along with the Assumption Log version.

## 6.5 Approval Gate

The project proceeds to Stage 2 only after:

1. The Use-Case Confirmation Lock (Section 1) is accepted.
2. The Assumptions Map (Section 2) is approved as an accurate representation of current knowledge gaps and priors.
3. The ISS (Section 3) is accepted as a realistic and sufficient representation of available inputs.
4. The Model Architecture summary (Section 4), consistent with the full architecture specification, is endorsed by the core team.
5. The Scenario Pack v1 (Section 5) is accepted as covering the key decision-relevant regimes.
6. The Verification Plan and traceability rules (Section 6) are accepted as adequate to ensure that Stage 3 can prove architecture–code equivalence.

Once these conditions are met, this Stage-1 package constitutes a review-ready logic and architecture dossier suitable for internal or external sharing.