IntuitionLabs

INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

EMA NAM Data Pilot: Eligibility and Evidence Readiness

> Build practical AI for pharma and biotech with IntuitionLabs. We help life-science teams turn complex information and workflows into useful software, governed knowledge systems and AI tools.

Published by IntuitionLabs · 2026-09-19 · ema nam pilot · ema voluntary data submission · new approach methodologies · nam briefing package · regulatory applicability areas · context of use · evidence readiness · ectd module 2.4 · toxicology

Original article: https://intuitionlabs.ai/articles/ema-nam-voluntary-data-pilot

In this article

Executive Summary

The European Medicines Agency, or **EMA**, opened its New Approach Methodologies, or **NAM**, voluntary data submission pilot on **1 September 2026** for human and veterinary medicines. Applications are expected to remain open until **September 2027** (ema.europa.eu). The mechanism lets companies, contract research organisations, method developers, non-governmental organisations, academic laboratories, consortia, and other stakeholders share NAM data outside a marketing-authorisation application. It is expressly **voluntary, non-binding, and non-decisional** and does not replace established regulatory procedures (ema.europa.eu). Participation therefore does not constitute qualification, acceptance, or permission to rely on a method in a future filing.

Fit begins with **scope**, not method novelty. A dataset should address one of four current regulatory applicability areas: highly specific biologics with no fully relevant species, developmental and reproductive toxicity under ICH S5(R3), safety pharmacology, or hepatotoxicity and drug-induced liver injury (ema.europa.eu). It must also articulate a bounded context of use that links the intended decision, development stage, endpoint, applicability domain, and limitations. That principle is consistent with the FDA definition of validation as establishing accuracy, reliability, and relevance for a specific context of use (fda.gov) and with EFSA's view that qualification depth varies by context (efsa.europa.eu).

Applicants first email the Word application form to the dedicated EMA VDS address or use EudraLink. EMA validates and prioritises the expression of interest, may redirect it to an **Innovation Task Force briefing**, **scientific advice**, or **qualification advice**, and invites selected applicants to submit a full package. A VDS package should explain the regulatory problem, whether the method is standalone or part of a battery or weight-of-evidence approach, its endpoint and mechanism, validation or characterisation, reliability, regulatory gap, advantages, and limitations. Where appropriate, its level of detail should resemble an electronic Common Technical Document **Module 2.4** written summary (ema.europa.eu).

Readiness is best tested before submission. The editorial model in this guide scores **relevance, robustness, reliability, external validation, and limitations disclosure**, each from 0 to 4, for a maximum of **20 points**. A score is a triage device, not an EMA criterion. Teams should also preserve raw and derived data, versions, scripts, controls, deviations, metadata, and an executable path from input to output. OECD guidance calls for an audit trail recording who did what and when (oecd.org), while FDA recommends an inventory of scripts, software versions, and package dependencies (fda.gov). These controls make a package assessable even though they do not guarantee a favourable pilot outcome.

Introduction and Background

The EMA voluntary data submission, or **VDS**, pilot addresses a practical asymmetry. New Approach Methodologies are increasingly used in research and early development, but regulators receive comparatively little method-level data outside product dossiers. The pilot creates a controlled channel for reviewing and discussing those datasets separately from a marketing-authorisation application (ema.europa.eu). Its objective is learning and confidence building, not a shadow approval procedure.

EMA's provisional definition spans **in vitro, in silico, in chemico, ex vivo, and refined in vivo** approaches. The agency notes that the definition remains provisional while a glossary is finalised (ema.europa.eu). This breadth makes the context of use, or **CoU**, decisive. A method is not simply "valid" in the abstract. The evidentiary question is whether it is sufficiently relevant and dependable for a specified purpose, endpoint, population or test system, and development decision.

The regulatory rationale sits within the **3Rs**, replacement, reduction, and refinement of animal use. Article 4 of Directive 2010/63/EU requires a scientifically satisfactory non-live-animal method to be used instead where possible (eur-lex.europa.eu), while separately requiring animal numbers to be minimised without compromising objectives (eur-lex.europa.eu). Those legal principles do not make every NAM submission-ready. Regulatory confidence still depends on purpose, performance, transparency, and limitations.

EFSA expresses the related qualification objective as demonstrating both reliability and relevance (efsa.europa.eu).

For NAM developers, preclinical leaders, contract research organisations, consortia, and regulatory scientists, the immediate decision is narrower: **Does the available dataset fit VDS now, or should it enter another pathway?** The answer requires a scope test, a pathway test, and an evidence-readiness test. IntuitionLabs is an adjacent life-sciences consultancy, not a provider of any EMA pathway. Its documented work connects authoritative sources with permissions, retrieval, citations, evaluation, and accountable operation (intuitionlabs.ai); that governance perspective informs the operational controls below without changing EMA's criteria.

Pilot Definition, Eligibility, and Scope

A bounded learning mechanism

The VDS pilot applies to **human and veterinary medicines** and permits voluntary discussion of data outside marketing-authorisation applications (ema.europa.eu). EMA states separately that it does not use pilot NAM data in its regulatory decision-making process (ema.europa.eu). The practical output is individual scientific feedback on potential relevance and fitness for purpose, not a transferable qualification opinion.

Eligibility has two layers. First, the submitter can be a company, **CRO**, method developer, non-governmental organisation, academic laboratory, consortium, or another relevant stakeholder. Second, the data must fall within a predefined CoU and regulatory applicability area, or **RAA**. A promising technology outside the current RAAs is not made eligible by adding more data.

The application form lists those applicant groups explicitly (ema.europa.eu). The scope test should therefore be completed before a team invests in submission formatting. NC3Rs likewise recommends defining clear NAM contexts of use (nc3rs.org.uk), and FDA asks that a CoU state its intended use and regulatory purpose (fda.gov).

Table 1 turns those requirements into a screening decision.

Question	Ready for VDS when	Redirect or remediate when
Applicant	A developer, company, CRO, laboratory, NGO, consortium, or other stakeholder owns or can lawfully share the package.	Ownership, permission, or a responsible contact is unresolved.
Regulatory applicability area	The problem fits one of the current RAAs described below.	The use case is scientifically interesting but outside the current RAAs. Consider ITF first.
Context of use	The statement specifies intended decision, development stage, endpoint, role, applicability domain, and limits.	The statement is merely “replace animal testing” or names a platform without a regulatory purpose.
Evidence maturity	Method, performance, validation or characterisation, reliability, and limitations can be documented.	Only a concept, prototype, or promotional summary exists. An ITF discussion may be more proportionate.
Regulatory status sought	The goal is non-binding learning about possible future relevance.	The goal is product-specific prospective advice or formal general qualification. Use scientific advice or QoNM.
Disclosure	A sufficiently interpretable package can be transferred securely, with necessary product context.	Commercial or contractual restrictions prevent assessors from understanding the method and results.

The table's central distinction is between **technical promise** and **pilot fit**. ICCVAM similarly describes regulatory acceptability as generally requiring adequate validation studies that characterise usefulness and limitations (ntp.niehs.nih.gov). FDA's current draft framework organises validation around CoU, human biological relevance, technical characterisation, and fitness for purpose (fda.gov). These external frameworks do not govern EMA's pilot, but they expose common weaknesses before application.

Writing a decision-useful context of use

A workable CoU should fit in one or two sentences and answer six questions:

- **Regulatory purpose:** What decision or evidence gap could the output inform? (fda.gov)
- **Development stage:** Discovery, first-in-human planning, clinical development, filing, or lifecycle use? (nc3rs.org.uk)
- **Method role:** Standalone, battery component, prioritisation tool, mechanistic support, or weight of evidence? (fda.gov)
- **Endpoint:** What is measured, and how does it connect to a safety endpoint? (fda.gov)
- **Applicability domain:** Which modalities, compounds, biological systems, populations, or exposure ranges are in scope? (oecd.org)
- **Limitations:** What is out of scope, uncertain, or dependent on another evidence stream? (fda.gov)

A concise template is: **"[Method and version] is intended to [role] for [regulatory problem or endpoint] in [defined domain and development stage], with outputs interpreted [standalone or with specified evidence]; it is not intended for [principal exclusions]."** FDA likewise defines a CoU through a model's specific role and scope (fda.gov), and ICCVAM cautions that several iterations may be necessary (ntp.niehs.nih.gov). Iteration before application is a sign of boundary discipline, not indecision.

The Four Regulatory Applicability Areas

The current annex contains the RAAs below. They are not method types. They are regulatory problem spaces within which different NAMs, batteries, or weight-of-evidence strategies may be considered.

RAA 1: highly specific biologics

RAA 1 concerns biologics with no pharmacologically relevant species or only a partial response (ema.europa.eu). The supporting case may include absent binding to a species ortholog, insufficient homology, or lack of functional activity. The NAM may help bridge limitations in an animal model, interpret existing studies, or address toxicology or pharmacokinetic questions. This is a relevance problem: the applicant must explain why the proposed human-relevant system is informative for the named risk, not merely why the animal comparator is imperfect. FDA's framework asks for predictive performance for the specific CoU (fda.gov).

Evidence preparation should include:

- **Target biology:** ortholog, binding, expression, pathway, and functional activity evidence. (link.springer.com)
- **Model relevance:** why the system reproduces the human mechanism or exposure condition. (fda.gov)
- **Bridging logic:** which uncertainty in the animal or other evidence stream the NAM addresses. (ntp.niehs.nih.gov)
- **Decision boundary:** what the NAM can and cannot support for first-in-human planning or dose selection. (nc3rs.org.uk)

For technical characterisation, ICCVAM recommends recording key data, metadata, and control measurements (ntp.niehs.nih.gov), while FDA asks sponsors to discuss both benefits and limitations (fda.gov).

The broader validation literature makes the same distinction. A peer-reviewed framework identifies five confidence elements: fitness for purpose, human biological relevance, technical characterisation, data integrity and transparency, and independent review (link.springer.com).

RAA 2: developmental and reproductive toxicity

RAA 2 covers developmental and reproductive toxicity, or **DART**, under ICH S5(R3) (ema.europa.eu). NAM evidence may contribute to a structured weight of evidence and may be submitted alongside preliminary in vivo embryo-fetal-development evidence. The package should define the developmental hazard question, timing and exposure relevance, reference set, positive and negative controls, and how discordant results will be interpreted. ICCVAM identifies descriptions of intra-laboratory and interlaboratory studies as useful evidence (ntp.niehs.nih.gov).

A battery should not be presented as a collection of assays. It needs a fixed integration logic. OECD defines a **Defined Approach** as specified information sources whose data are interpreted using a fixed procedure (oecd.org). EFSA likewise describes NAM evaluation within a structured weight-of-evidence approach (efsa.europa.eu). Applicants should therefore predefine how individual results affect the conclusion and show how missing or indeterminate components are handled.

RAA 3: safety pharmacology

RAA 3 includes safety pharmacology for hazard identification, mechanistic understanding, screening, prioritisation, and targeted data generation (ema.europa.eu). It can encompass evidence relevant to cardiovascular, central nervous system, or respiratory risk. The applicant should separate exploratory usefulness from the proposed regulatory role and show endpoint sensitivity, specificity, dynamic range, controls, and exposure coverage. FDA describes technical characterisation as showing that endpoint measurement is robust, reliable, and reproducible (fda.gov).

Technical characterisation should test repeatability, reproducibility, transferability where relevant, and the effect of key variables. OECD's long-standing validation guidance says a method should be robust, transferable, and standardisable (oecd.org). It also calls for documentation of inter-laboratory and intra-laboratory variability over time (oecd.org).

RAA 4: hepatotoxicity and DILI prediction

RAA 4 covers hepatotoxicity and drug-induced liver injury, or **DILI**, for small molecules and biological products (ema.europa.eu). Potential evidence includes human-liver-relevant mechanistic, functional, and exposure-response information; identification of hepatotoxic liabilities; dose-response characterisation; and investigation of mechanisms such as mitochondrial dysfunction or bile-acid-transport inhibition. FDA recommends quantitative comparison of computational quantities of interest with experimental comparators (fda.gov).

A credible package connects assay outputs to a biologically and regulatorily meaningful interpretation. That requires clear model provenance, exposure assumptions, acceptance criteria, uncertainty, and known failure modes. A broad review describes NAM technologies as having a wide range of technical and regulatory

readiness (academic.oup.com). RIVM similarly found that most reviewed nanomaterial NAMs still required validation for regulatory risk assessment (rivm.nl). The lesson is not that validation must be complete before VDS. It is that the package must accurately locate the method on its maturity path.

Choosing VDS, ITF, Scientific Advice, or Qualification

EMA explicitly allows intake validation to redirect an applicant toward another regulatory interaction. The decision turns on **maturity, product specificity, intended output, and regulatory effect**.

Table 2 compares the four pathways for a NAM developer.

Route	Best fit	Typical input	Output and status	Timing or cost signal
VDS pilot	Data fit a current RAA and the objective is regulatory learning outside a product application.	Short application, then invited full package.	Individual, non-binding, non-decisional scientific feedback. No qualification or acceptance is implied.	The published pilot window is time limited. Assessment is expected to follow scientific-advice timing, but may vary with volume, priority, complexity, and expert availability.
Innovation Task Force briefing	The method is early, novel, or its regulatory route is unclear.	Focused background and questions suitable for informal discussion.	Confidential minutes and preliminary views that may not represent EMA scientific committees.	Virtual and free. Initial eligibility feedback is normally expected in **1 to 3 weeks** (ema.europa.eu).
Scientific advice	A sponsor needs prospective answers for development of a particular medicine, including planned NAM use in a future clinical-trial application.	Product-specific questions, positions, and supporting evidence.	Confidential final advice letter. Advice is not legally binding.	Formal procedures are **40 days** without, or **70 days** with, a discussion meeting (ema.europa.eu) (ema.europa.eu). Fees apply.
Qualification of novel methodologies	The method is intended for broader use across programmes and a formal qualification trajectory is sought.	Qualification advice can begin with rationale, preliminary data, and a qualification plan; an opinion requires mature evidence and external validity.	Advice produces a confidential letter. A final opinion accepts the methodology for evidence generation only within the defined CoU.	Early support is free; formal requests are fee-bearing. A granted scoping meeting is scheduled within **6 weeks** (ema.europa.eu).

The routes are complementary rather than hierarchical. **ITF** is appropriate when the development concept or route needs shaping. EMA says these meetings occur earlier than scientific advice (ema.europa.eu). **Scientific advice** is prospective and product-specific, with questions tied to a particular development programme. **Qualification advice** is appropriate when preliminary data and a plan exist for a method intended to have broader use; independent validation or testing is generally a confirmatory step (ema.europa.eu). **VDS** is the safe learning route only when the current pilot scope and package requirements are met.

A useful decision sequence begins with the intended regulatory output. **Product-specific and prospective** needs point to scientific advice. **Formal, reusable qualification** points to qualification advice and potentially an opinion. **An early method or unclear route** points to an ITF briefing. **An assessable dataset within a current RAA, where non-binding feedback is sufficient**, points to VDS. Where several descriptions apply, the interactions should be sequenced and each unresolved question assigned to one route.

Application Flow and Briefing-Package Architecture

From application to invited package

The first step is a completed **Word application form** sent to the EMA VDS mailbox or via EudraLink. EMA describes EudraLink as an encrypted route for confidential information. Applications undergo initial validation, followed by internal validation and prioritisation. Selected applicants are invited to submit a sufficiently documented full package through EudraLink and therefore need an EMA account. Pilot data are stored in EMA's secure SharePoint environment and used for pilot purposes.

EMA's form says EudraLink may be used to send the application (ema.europa.eu). Before transfer, an internal dry run should confirm that another reviewer can navigate the package. NIDDK's reproducibility guidance expects raw validation data to be shared (niddk.nih.gov), and OECD says reproducible reporting should allow others to reproduce or fully reconstruct the study (oecd.org).

The form should do more than name the technology. It should let an intake reviewer answer:

- **Who owns the interaction?** Identify applicant, scientific contact, partners, and data rights. (eur-lex.europa.eu)
- **Which RAA applies?** Name one primary RAA and justify the fit. (nc3rs.org.uk)
- **What is the CoU?** State decision, stage, role, endpoint, domain, and exclusions. (fda.gov)
- **What evidence exists?** Summarise datasets, validation or characterisation, comparators, and external work. (ntp.niehs.nih.gov)
- **Why VDS?** Explain why non-binding learning is the right output now. (efsa.europa.eu)
- **What could be shared?** Flag confidential content and the proposed secure-transfer approach. (niddk.nih.gov)

Package map and Module 2.4 crosswalk

EMA calls the briefing package the primary basis for review. It asks for a high-level RAA and problem description; role as standalone, battery, or weight of evidence; endpoint, mechanism, and safety endpoint; CoU and stage; validation or characterisation; data reliability; intended regulatory application; guidance compatibility; regulatory gap; and advantages and limitations. Relevant product context should be supplied, even where full disclosure is not possible.

The information document specifically asks for measures taken to ensure data reliability (ema.europa.eu). A peer-reviewed confidence framework treats traceability from raw data to final report as a core element (link.springer.com), and OECD asks that software name and version be documented for model fitting (oecd.org).

The requested alignment to **eCTD Module 2.4** is a level-of-detail and written-summary analogy, not an instruction to build a full marketing-authorisation dossier. ICH M4S says the Nonclinical Overview should critically integrate pharmacology, pharmacokinetics, and toxicology (database.ich.org), discuss and justify the nonclinical strategy (database.ich.org), and discuss the scientific validity of alternatives to whole-animal experiments (database.ich.org).

Table 3 combines a package crosswalk with an operational evidence inventory.

Package block	Question to answer	Minimum inventory fields	Module 2.4-style treatment
RAA and problem	What regulatory uncertainty is being addressed, and why does the selected RAA fit?	Owner, approved problem statement, RAA version, source guidance, unresolved assumptions.	Open with the strategy and decision context, not a platform description.
Method and role	Is the NAM standalone, in a battery, or part of weight of evidence?	Method name, version, protocol, software, instrument, model or cell source, role, dependencies.	Integrate the method with the broader nonclinical strategy.
Endpoint and mechanism	What does the output measure, and how does it relate to the safety endpoint?	Endpoint definition, unit, transformation, mechanism map, acceptance criteria, biological controls.	Critically connect mechanistic relevance to the proposed interpretation.
Context of use	What decision, stage, population or domain, and exclusions apply?	CoU version, accountable owner, applicability domain, exclusions, change history.	State boundaries early and use them consistently throughout.
Performance	How accurate, sensitive, specific, repeatable, and robust is the method for the CoU?	Dataset IDs, reference set, comparator, metrics, uncertainty, protocol deviations, analysis code.	Synthesise results and explain discordance rather than listing studies.
Validation and reliability	What internal and external characterisation exists, and can outputs be reproduced?	Laboratory, dates, operators, lots, controls, raw and derived data, audit trail, independent review.	Discuss scientific validity, transferability, and remaining evidence gaps.
Regulatory application	What gap could the NAM fill, and how is it used internally or in submissions?	Decision log, prior interactions, guidance mapping, planned use, fallback strategy.	Explain the overall weight and the consequences of uncertainty.
Advantages and limitations	Where does the approach add information, and where can it fail?	Known limits, failure modes, out-of-domain rules, missing data, mitigation owner and date.	Present a balanced critical assessment, including departures from guidance.

This crosswalk prevents two common package failures: a data dump without an argument, and an argument without traceable data. EURL ECVAM asks for detailed documentation and raw data sufficient for full review (joint-research-centre.ec.europa.eu), while FDA recommends mapping every analysis script to its inputs and outputs (fda.gov). The right package connects both layers.

Data Governance and Reproducible Outputs

A submission-ready evidence repository

The pilot document promises respect for commercially confidential information, but confidentiality does not remove the need for interpretability. Teams should separate the **scientific package**, **transfer package**, and **publication or disclosure plan**, then assign access rights accordingly. Relevant product context should still describe modality, pharmacological class, target, mechanism, indication, and novelty when fuller disclosure is constrained.

Access controls should be explicit: OECD recommends procedures for assigning access rights to each user (oecd.org), and it says access to stored records must be restricted, controlled, and documented (oecd.org).

A practical repository should contain:

- **Source data:** immutable raw files, checksums, acquisition metadata, and data dictionary. (oecd.org)
- **Derived data:** transformation logic, code, parameters, exclusions, and links to source records. (fda.gov)
- **Method control:** approved protocol, standard operating procedures, deviations, and version history. (oecd.org)
- **Model control:** source code or executable, environment, software and package versions, random seeds, and dependencies. (fda.gov)
- **Biological materials:** provenance, identity, passage, lot, supplier, catalogue number, storage, and expiry. (oecd.org)
- **Performance evidence:** reference set, acceptance criteria, repeats, failed runs, sensitivity analyses, and uncertainty. (oecd.org)
- **External evidence:** laboratory, independence, transfer protocol, results, deviations, and reconciliation. (ntp.niehs.nih.gov)
- **Regulatory map:** RAA, CoU, guidance crosswalk, prior advice, decision log, and open issues. (fda.gov)
- **Ownership:** accountable owner, reviewer, approver, location, access group, and retention rule. (oecd.org)

OECD guidance distinguishes records of raw and derived data and study reports (oecd.org), requires controlled documents to carry dated and approved version numbers (oecd.org), and expects reagent metadata such as lot, catalogue, supplier, and expiry (oecd.org). These are useful controls even when a particular study is not formally conducted under Good Laboratory Practice.

Computational and AI-enabled NAMs

For an in silico or artificial-intelligence-enabled method, a static report is insufficient. The reviewer needs to know which executable object produced each result. FDA's model-data guidance recommends a define file for every dataset and associated variables (fda.gov), plus model code, control streams, and output listings (fda.gov). Its computational guidance recommends separating calibration data from validation data (fda.gov) and explicitly stating model limitations (fda.gov).

At minimum, freeze and identify the **model identity**, including semantic version, commit, weights or parameter file, build, and date; the **execution environment**, including operating system, libraries, containers, hardware dependencies, and configuration; and all **data partitions**, including training, calibration, validation, challenge, and external datasets. The package should also record **run controls** such as seed, thresholds, preprocessing, missing-data rules, and post-processing; a trace connecting input identifier, run identifier, log, output, reviewer, and exception; and an independent **reproduction test** with expected tolerances.

FDA also expects the software and version to be named (fda.gov) and instructions sufficient to run submitted code (fda.gov).

NIDDK advises that novel analytical pipelines be documented sufficiently for another investigator to recreate the analysis (niddk.nih.gov), and that metadata should let reusers reproduce findings (niddk.nih.gov). FAIR principles extend that goal by making research objects findable, accessible, interoperable, and reusable for human and machine use (nature.com).

Confidentiality, access, and auditability

Commercially confidential information should be classified at field or artifact level, not used as a blanket label. Define who may access each artifact, which version was transferred, and what redactions change scientific interpretability. GDPR identifies pseudonymisation and encryption as appropriate safeguards for personal data (eur-lex.europa.eu) and requires ongoing confidentiality, integrity, availability, and resilience (eur-lex.europa.eu). Those duties apply when personal data are present; they should not be confused with the separate scientific question of whether the package is complete enough to assess.

IntuitionLabs documents data pipelines, integration, warehousing, and business intelligence as part of its adjacent consulting scope (intuitionlabs.ai). In a VDS-readiness engagement, that capability belongs behind the submission as evidence governance and traceability. It does not make the consultancy a regulatory route, method qualifier, or substitute for applicant accountability.

Data Analysis and Evidence

The pilot has no public acceptance-rate benchmark yet. Its most useful quantitative signals therefore concern timing, evidence scale, reproducibility, and method performance. They should be interpreted as planning anchors, not thresholds imposed by EMA.

First, the operational window is approximately **12 months**, from September 2026 through expected closure in September 2027 (ema.europa.eu). A lessons-learned report is planned after the pilot. The information document says package assessment is expected to follow scientific-advice timing but may vary with submission number, priority, complexity, and expert availability. The formal scientific-advice benchmarks of **40 days** without a discussion meeting and **70 days** with one provide context (ema.europa.eu) (ema.europa.eu), but they are not guaranteed VDS service levels.

Second, external validation can be materially larger than a single-laboratory repeatability exercise. The EU-NETVAL network reports **31 member laboratories** (joint-research-centre.ec.europa.eu), and a described validation design includes method transfer, within-laboratory reproducibility, between-laboratory reproducibility, and relevance assessment (joint-research-centre.ec.europa.eu). The validation process also includes

independent scientific peer review by EURL ECVAM's advisory committee (joint-research-centre.ec.europa.eu). These features show the breadth of review and variability that validation programmes may need to address.

Third, performance metrics must be presented together. A three-laboratory EpiSensA study used **15 coded chemicals** and reported within-laboratory reproducibility of **93.3%, 93.3%, and 86.7%** (analyticalsciencejournals.onlinelibrary.wiley.com). Across **27 chemicals**, between-laboratory reproducibility was **88.9%** (analyticalsciencejournals.onlinelibrary.wiley.com). Against the murine local lymph node assay, the same study reported **92.6% sensitivity, 63.0% specificity, 82.7% accuracy, and 77.8% balanced accuracy** (analyticalsciencejournals.onlinelibrary.wiley.com). The uneven sensitivity and specificity illustrate why a single headline accuracy value is inadequate.

Fourth, transparency affects reproducibility. In an NLM bioinformatics workshop exercise, **no team fully reproduced its assigned paper** (nlm.nih.gov). Participants identified software-version disclosure and availability of underlying raw data as minimum standards (nlm.nih.gov) (nlm.nih.gov). A VDS package should assume that a reviewer who cannot reconstruct the result will discount its evidentiary weight.

Finally, these numbers support three planning conclusions:

- **Do not turn timelines into commitments.** Use the scientific-advice cycle only as a resource-planning analogy.
- **Do not equate repeatability with regulatory relevance.** Both performance and CoU relevance must be demonstrated.
- **Do not report one metric.** Include confusion matrices, uncertainty, prevalence or reference-set composition, exclusions, and out-of-domain behavior.

OECD's QSAR principles require appropriate measures of goodness of fit, robustness, and predictivity (oecd.org). Predictions outside a defined applicability domain are extrapolations and are less likely to be reliable (oecd.org).

FDA likewise asks developers to demonstrate predictive performance for the proposed CoU (fda.gov).

Implementation Guidance and Readiness Scoring

An editorial 20-point score

The following score is an **editorial triage tool**, not an EMA rubric. Score each dimension from **0 to 4**, then add the five values for a maximum of **20**:

- **Relevance:** 0 means no defined regulatory problem; 4 means the CoU, RAA, endpoint, human biology, and decision link are explicit and supported. (efsa.europa.eu)
- **Robustness:** 0 means no controlled performance evidence; 4 means protocol variables, sensitivity, repeatability, controls, and operating range are characterised. (oecd.org)
- **Reliability:** 0 means outputs cannot be traced; 4 means raw-to-result provenance, quality control, deviations, software, and audit trail are complete. (oecd.org)

- **External validation:** 0 means developer-only evidence; 4 means independent, transferable, appropriately powered evidence supports the proposed CoU. (ntp.niehs.nih.gov)
- **Limitations:** 0 means limitations are absent; 4 means applicability boundaries, uncertainty, failure modes, and mitigations are explicit. (fda.gov)

Interpret totals conservatively:

- **0 to 7, discovery:** clarify the problem and create a controlled development plan. ITF may be the better interaction.
- **8 to 12, characterisation:** close critical robustness and traceability gaps before VDS.
- **13 to 16, application candidate:** draft the form, test the CoU, and assemble the package index.
- **17 to 20, review-ready:** run an independent challenge review, freeze artifacts, and prepare secure transfer.

The score should never average away a fatal gap. A total of 17 with a zero for relevance is not ready. EURL ECVAM describes validation as a four-stage process including submission assessment, validation studies, independent peer review, and recommendations (joint-research-centre.ec.europa.eu). That sequence reinforces why independent review should be treated as its own dimension.

EURL ECVAM separately advises an independent data audit to support quality and integrity (joint-research-centre.ec.europa.eu). The confidence framework includes reproducibility, transferability, applicability domain, reference chemicals, and controls (link.springer.com).

Gap remediation by dimension

For relevance gaps:

- Rewrite the CoU until every dataset has a stated role. (fda.gov)
- Map endpoints to biological mechanisms and the named safety question.
- Document the applicability domain and out-of-domain rule.
- Separate current evidence from planned claims.

For robustness gaps:

- Predefine acceptance criteria and analysis before new acquisition. (oecd.org)
- Run variable and perturbation studies around critical method parameters.
- Use coded reference materials where blinding is appropriate. OECD prefers coded reference chemicals to reduce bias (oecd.org).
- Report failed runs and missingness, not only accepted outputs.

For reliability gaps:

- Reconcile raw, processed, analysed, and reported datasets.
- Validate and document scripts that process raw data (oecd.org).
- Freeze software and model versions, dependencies, and run instructions.
- Execute a clean-room reproduction from archived inputs. (niddk.nih.gov)

For external-validation gaps:

- Choose an independent laboratory and define transfer acceptance criteria.
- Predefine how protocol adaptations will be recorded and assessed.
- Preserve the distinction between calibration and validation evidence.
- Explain whether external evidence covers the same CoU and version.

For limitations gaps:

- Create a limitation register tied to claims and mitigation owners.
- Quantify uncertainty where the evidence permits.
- State known confounders, exclusions, and conditions that invalidate output.
- Explain discordance with established evidence without forcing agreement.

NC3Rs calls for clear NAM contexts of use and validation criteria tailored to specific contexts (nc3rs.org.uk). OECD Good In Vitro Method Practices similarly makes defined standards and rigorous, reproducible data conditions for confidence (oecd.org). These principles make remediation more efficient because work is prioritised against the intended decision.

Implications and Future Directions

The VDS pilot creates an evidence-learning channel at a useful point between informal innovation dialogue and formal regulatory advice. Its value will depend less on submission volume than on whether the incoming packages let assessors compare methods, contexts, evidence maturity, and recurring gaps. EMA plans a lessons-learned report after the pilot, so application patterns may influence future operating practice without guaranteeing acceptance of any individual method.

Three implications follow for applicants. First, **portfolio evidence can be more informative than a single showcase study** when it reveals domain boundaries and performance across mechanisms, compounds, or laboratories. Second, **versioned evidence is essential** because changes to cells, protocols, software, or training data can break the connection between earlier validation and the submitted artifact. FDA advises assessing differences and their impact when validation evidence was generated with another model version (fda.gov). Third, **limitations are part of the evidence**, not a concession. A bounded, reproducible claim is more reviewable than a broad claim supported by selective results.

The pilot also points toward greater international convergence around CoU, reliability, relevance, and transparent reporting. EFSA says qualification aims to demonstrate both reliability and relevance (efsa.europa.eu), while RIVM identifies FAIR data management as important to validation and acceptance (rivm.nl). Alignment is not identity: applicants should still map each package to the precise EMA RAA and pathway.

The wider literature supports the same cautious convergence: FAIR principles target machine-actionable reuse (nature.com), while a cross-technology review describes a broad spectrum of technical and regulatory readiness (academic.oup.com).

Operationally, teams should maintain a living **regulatory evidence product** rather than assemble a one-off document. That product includes the CoU, evidence inventory, executable model or controlled protocol, performance ledger, limitations register, and decision log. It can support VDS now and reduce rework if the method later enters scientific advice or qualification.

OECD's current research-data guidance frames readiness as a lifecycle from generation and reporting through evaluation and regulatory integration (oecd.org). NC3Rs also identifies transparent reporting as part of building confidence in NAMs (nc3rs.org.uk).

Frequently Asked Questions (FAQs)

Who is eligible for the EMA NAM pilot?

Companies, CROs, method developers, NGOs, academic laboratories, consortia, and other relevant stakeholders may apply. Eligibility still depends on whether the data fit a predefined CoU and one of the current RAAs.

The applicant categories come from the current official form (ema.europa.eu).

How does an applicant apply?

Complete the official Word application form and send it to the dedicated VDS mailbox, using EudraLink where appropriate for protected transfer. EMA validates and prioritises applications, then invites selected applicants to send a full package.

The form confirms that EudraLink may be used for transfer (ema.europa.eu).

The two-stage pattern is also common in method validation: EURL ECVAM requires presubmission and, if invited, a complete submission (joint-research-centre.ec.europa.eu). This is an analogy for package planning, not a statement that the two programmes share criteria.

What belongs in the briefing package?

Include the RAA and problem, method role, endpoint and mechanism, safety endpoint, CoU, development stage, validation or characterisation, reliability, external evidence, regulatory gap, guidance mapping, advantages, limitations, and relevant product context. Organise it as an integrated argument with traceable supporting data.

EMA specifically asks for measures taken to ensure data reliability (ema.europa.eu).

Does participation mean EMA has accepted or qualified the NAM?

No. The pilot is non-binding and non-decisional. Qualification is a separate procedure, and product-specific prospective questions generally belong in scientific advice.

The official information document characterises the mechanism in those terms (ema.europa.eu).

How long will EMA assessment take?

EMA expects assessment to follow the scientific-advice timeline, but expressly says timing may vary. Applicants should not treat the published scientific-advice procedure lengths as VDS commitments.

Is external validation mandatory before applying?

The pilot package should describe internal and external validation or characterisation where available and explain reliability. The required maturity depends on the CoU and claim. Absence of external validation should be disclosed and reflected in the proposed use, limitations, and remediation plan.

EFSA similarly says the degree of qualification varies with the specific CoU (efsa.europa.eu), while ICCVAM asks for the data and information a sponsoring agency needs to assess suitability (ntp.niehs.nih.gov).

Can AI or computational models be submitted?

Yes, the provisional NAM scope includes in silico approaches. A computational package should identify data partitions, code or executable, software and model versions, parameters, environment, scripts, outputs, applicability domain, uncertainty, and reproduction instructions. OECD states that transparent algorithm reporting should let others reproduce a QSAR model (oecd.org).

What happens to confidential information?

EMA states that pilot assessment and outcome publication will respect commercially confidential information. Applicants should use the specified secure route, classify sensitive artifacts, minimise personal data, document access, and preserve enough scientific context for assessment.

The commitment appears in the pilot information document (ema.europa.eu).

Conclusion

The **EMA NAM voluntary data pilot** is a bounded learning mechanism for assessable datasets, not a shortcut to regulatory acceptance. As of publication, it is open to a broad group of stakeholders, covers human and veterinary medicines, and has a time-limited application window. Fit requires both a current regulatory applicability area and a precise context of use.

The most effective application starts with a route decision. Early concepts and route uncertainty point to ITF. Product-specific prospective questions point to scientific advice. A broadly reusable methodology on a formal acceptance path points to qualification advice. VDS fits when an applicant has a sufficiently documented dataset and seeks non-binding feedback outside a marketing-authorisation application.

Evidence readiness is then an exercise in disciplined integration. The package should connect regulatory problem, method role, endpoint, mechanism, performance, validation, reliability, application, and limitations at roughly a Module 2.4 written-summary level where appropriate. Its supporting repository should preserve data provenance, versions, controls, scripts, metadata, audit trails, and reproducible outputs.

That repository design is consistent with the FAIR goal of reusable research objects (nature.com).

The editorial **20-point readiness score** can prioritise work, but it cannot confer eligibility or predict EMA's view. The decisive standard is narrower and more useful: can an assessor understand exactly what the NAM is intended to do, reproduce how the evidence was generated, see where it performs, and identify where it does not? If the answer is yes within a current RAA, the dataset is a credible VDS candidate. If not, the gaps should determine the next experiment, governance control, or regulatory interaction.

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://www.ema.europa.eu/en/human-regulatory-overview/research-development/ethical-use-animals-medicine-testing/regulatory-acceptance-new-approach-methodologies-nams-reduce-animal-use-testing/new-approach-methodologies-replace-reduce-animal-testing-voluntary-data-submission-pilot
2. https://www.ema.europa.eu/en/documents/other/voluntary-data-submission-vds-pilot-information-document_en.pdf
3. https://www.ema.europa.eu/en/documents/template-form/voluntary-data-submission-vds-pilot-application-form_en.docx
4. https://www.fda.gov/media/191589/download
5. https://www.efsa.europa.eu/en/topics/new-approach-methodologies-nams
6. https://www.oecd.org/en/publications/guidance-document-on-good-in-vitro-method-practices-givimp_9789264304796-en/full-report/component-15.html
7. https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/model-data-format
8. https://www.ema.europa.eu/en/human-regulatory-overview/research-development/ethical-use-animals-medicine-testing/regulatory-acceptance-new-approach-methodologies-nams-reduce-animal-use-testing
9. https://eur-lex.europa.eu/eli/dir/2010/63/oj/
10. https://intuitionlabs.ai/
11. https://nc3rs.org.uk/sites/default/files/2025-09/Incorporating%20new%20approach%20methodologies%20in%20the%20development%20of%20new%20medicines.pdf
12. https://ntp.niehs.nih.gov/whatwestudy/niceatm/resources-for-test-method-developers/submissions
13. https://intuitionlabs.ai/articles/preclinical-vs-clinical-research
14. https://www.fda.gov/media/154985/download
15. https://www.oecd.org/content/dam/oecd/en/publications/reports/2014/09/guidance-document-on-the-validation-of-quantitative-structure-activity-relationship-q-sar-models_g1ghcc68/9789264085442-en.pdf
16. https://ntp.niehs.nih.gov/sites/default/files/2024-03/VWG_Report_27Feb2024_FD_508.pdf
17. https://link.springer.com/article/10.1007/s00204-022-03365-4
18. https://www.oecd.org/en/publications/guideline-no-497-defined-approaches-on-skin-sensitisation_b92879a4-en.html
19. https://www.oecd.org/content/dam/oecd/en/publications/reports/2005/08/guidance-document-on-the-validation-and-international-acceptance-of-new-or-updated-test-methods-for-hazard-assessment_2695fa12/e1f1244b-en.pdf
20. https://www.fda.gov/media/87586/download
21. https://academic.oup.com/toxres/article/13/2/tfae044/7634804
22. https://www.rivm.nl/en/weblog/new-approach-methodologies-for-risk-assessment-and-safe-and-sustainable-by-design-of
23. https://www.ema.europa.eu/en/human-regulatory-overview/research-development/innovation-task-force-briefing-meetings

24. https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/european-medicines-agency-guidance-applicants-seeking-scientific-advice-protocol-assistance_en.pdf
25. https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-novel-methodologies-drug-development-guidance-applicants_en.pdf
26. https://www.niddk.nih.gov/research-funding/research-resources/data-management-sharing/guidance-writing-dms-plan
27. https://www.oecd.org/en/publications/guidance-document-on-good-in-vitro-method-practices-givimp_9789264304796-en/full-report/component-14.html
28. https://eur-lex.europa.eu/eli/reg/2016/679/2016-05-04/eng/pdf
29. https://www.oecd.org/en/publications/guidance-document-on-good-in-vitro-method-practices-givimp_9789264304796-en/full-report/component-13.html
30. https://database.ich.org/sites/default/files/M4S_R2_Guideline.pdf
31. https://joint-research-centre.ec.europa.eu/scientific-activities/reference-and-measurement/european-union-reference-laboratories/eu-reference-laboratory-alternatives-animal-testing-eurl-ecvam/alternative-methods-toxicity-testing/validation-and-submission-process/eurl-ecvam-test-method-submission_en
32. https://www.oecd.org/en/publications/guidance-document-on-good-in-vitro-method-practices-givimp_9789264304796-en/full-report/component-7.html
33. https://intuitionlabs.ai/articles/21-cfr-part-58-glp-guide
34. https://intuitionlabs.ai/articles/pharma-ai-validation-evidence-fda-ema
35. https://www.fda.gov/files/drugs/published/Physiologically-Based-Pharmacokinetic-Analyses-%E2%80%94-Format-and-Content-Guidance-for-Industry.pdf
36. https://www.nature.com/articles/sdata201618
37. https://joint-research-centre.ec.europa.eu/scientific-activities/reference-and-measurement/european-union-reference-laboratories/eu-reference-laboratory-alternatives-animal-testing-eurl-ecvam/alternative-methods-toxicity-testing/european-union-network-laboratories-validation-alternative-methods_en
38. https://joint-research-centre.ec.europa.eu/projects-and-activities/reference-and-measurement/european-union-reference-laboratories/eu-reference-laboratory-alternatives-animal-testing-eurl-ecvam/alternative-methods-toxicity-testing/validation-and-submission-process_en?prefLang=es
39. https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/jat.4559
40. https://www.nlm.nih.gov/od/osi/documents/Short_Report_on_Reproducibility_Workshop.pdf
41. https://www.oecd.org/en/publications/guidance-document-on-good-in-vitro-method-practices-givimp-second-edition_5ba6777b-en.html
42. https://www.oecd.org/en/publications/guidance-document-on-the-generation-reporting-and-use-of-research-data-for-regulatory-assessments_8d49ec1d-en.html

THE PUBLISHER

About IntuitionLabs

IntuitionLabs is an AI consulting, custom software development and data engineering firm serving pharmaceutical, biotechnology, medical-device and other life-science organizations. We work with clinical, regulatory, medical-affairs, commercial, quality and IT teams to connect technology decisions with the work people need to accomplish.

AI consulting and adoption

Our AI enablement services cover readiness assessments, use-case selection, governance and policies, team workshops, adoption measurement and ongoing advisory support. We help organizations structure the information layer behind AI: source material, context, permissions and maintained knowledge that make generated answers useful and reviewable. Private LLM inference and hosted AI options support teams evaluating how to operate AI with appropriate control over their data and infrastructure.

Software, data and life-science workflows

IntuitionLabs develops custom software for pharma and biotech, integrates enterprise systems, and builds data engineering and business intelligence solutions. Areas of focus include AI agents, regulatory research, medical writing, medical affairs, CMC information, competitive intelligence and clinical-document workflows. Our eTMF intelligence work includes cross-system reconciliation and inspection-readiness support.

Enterprise platforms and regulated delivery

We provide Veeva services, application support, managed services, integrations and custom applications, alongside enterprise content work involving platforms such as Egnyte. For regulated workflows, our services include GxP enablement, computer-system validation and software development addressing 21 CFR Part 11 requirements. The applicable controls, validation responsibilities and acceptance criteria are defined for each engagement.

Work with IntuitionLabs

Explore AI enablement, pharma and biotech software development, data engineering and BI, and Veeva services. Contact IntuitionLabs to discuss your workflow, information sources and implementation needs.

IntuitionLabs publishes educational research to help life-science teams make informed technology decisions. Coverage of a product or organization does not imply a client relationship, endorsement or partnership.

Website: https://intuitionlabs.ai