WHITE PAPER

AI-Digital-Physical Convergence: The Future of DMTA in Drug Discovery & Development

By Revvity Signals

August, 2026



Embracing Digitalization & AI in Drug Discovery

In the rapidly evolving landscape of technology, AI-augmented innovation plays a critical role in driving the digital and physical worlds towards a synergistic future. This transformative power of AI influences a multitude of domains—from general advancements to specific industries such as chemicals—and impacts the entire R&D and innovation lifecycle.

While some pharmaceutical discovery organizations are fully digitalized, the majority are fragmented environments where scientists are responsible for the transposition of information between systems throughout the design-make-test-analyze (DMTA) cycle.

As we delve into the various facets of this phenomenon, we will explore how using AI and digital augmentation changes the dynamics of drug discovery, prioritization, and operational productivity; as well as the specialized knowledge required for effective data science and data engineering. In this paper we describe **the ideal digital-physical DMTA cycle that will equip drug discovery to leverage AI-augmented innovation**.

AI-Augmented Innovation & Enabling a Digital-Physical Virtuous Cycle

Innovators in various industries have recognized the value of implementing a digital-physical virtuous *cycle*. These innovators seek to implement this as a new foundation to support the discovery, development, and commercial launch of new products and technologies.

Specifically, this cycle refers to implementation of **a continuous and mutually reinforcing loop, where digital tools and methods enhance physical processes**; then, feedback from these improved physical processes informs further digital advancements.

The digital-physical virtuous cycle leads to accelerated innovation, improved efficiency, and the development of more sophisticated, data-driven solutions.

What is the DMTA Cycle?

To successfully discover new therapies, innovators must perform exploratory and confirmatory experimentation. A DMTA cycle must be completed for each experiment. DMTA involves the following distinct activities:

© 1994-2026 Revvity Signals Software, Inc. All rights reserved.

Design

Creating a conceptual framework for potential clinical drug candidates. Design includes brainstorming, ideation, and outlining the initial specifications and functionalities of experiments. In addition to this conceptual framework for output product and physical processes, methods of analysis are carefully selected based upon the output product's critical quality attributes and corresponding analytical target profile. Including analysis considerations as a part of the design assessment helps to ensure effective decision making.

Make

Conceptual designs are transformed into physical products. Make involves synthesizing compounds. Electronic capture of observations made during activities ensures that no latent variables are missed during decision-making.

Test

Experiments are subjected to pre-process, in-process, and post-process sampling (or probe-based continuous measurement). Analysis modalities include:

- **Targeted Analyses**—a specific sample feature or attribute is known, a suitable method to detect and measure the specific feature is available for use, and the results of the detection or measurement are unambiguous.
- **Non-Targeted Analyses**—a range of potential sample features or attributes is suspected, a suitable method to detect such potential features or attributes (within a specific attribute range of values) is available for use, and the results of detection or measurement is nondeterministic.

Analyze

Analysis of test results and corresponding observations made during the make step allow practitioners to derive insights and identify areas for improvement. It includes data analysis, interpretation of results, and making informed decisions for subsequent iterations.

The Vicious Cycle: Manual Data Transfer in DMTA Transitions

Each transition within the **non-digitalized DMTA cycle** requires significant human transposition and translation of information between systems to bridge the gap between different stages and domains of expertise.

From Design to Make

The transition from design to make requires translation of conceptual designs into logical and physical designs:

© 1994-2026 Revvity Signals Software, Inc. All rights reserved.

- **Conceptual Design**—the high-level idea and overall vision of the product. It includes abstract representations and doesn't delve into technical specifics.
- **Logical Design**—the conceptual design is detailed into logical structures and relationships. It involves creating diagrams, flowcharts, and outlines that define how different components will interact.
- **Physical Design**—translation of the logical designs into tangible components and assemblies. It includes detailed blueprints, specifications, and instructions for manufacturing the actual product. In chemical or bioprocesses, the translation from statistics-based and stoichiometry-based process factors must be translated to physical parameters, gravimetric, and volumetric units.

In this transition, scientists and engineers must also translate statistical terms used in conceptual design into experimental and chemical terms relevant to the manufacturing process. For example, statistical models predicting product behavior must be converted into precise chemical formulations and experimental layouts.

From Make to Test

The transition from make to test involves preparing for the appropriate methods of experimental observations. This requires translating the physical designs into testable entities and defining appropriate testing protocols. Scientists must bridge the gap between the experiment or process parameters, material attributes specifications, and appropriate testing criteria. Before, during, and after each process execution, scientists must assign the appropriate sample identifiers to physical samples to ensure analysis results are associated to the appropriate process, sampling times, etc.

From Test to Analyze

The transition from test to analyze requires translation of raw test data into meaningful insights. This involves interpreting test results, identifying patterns, and making data-driven decisions. Scientists must convert the empirical data from testing into actionable feedback that informs the next cycle of design and development.
Human translation is essential in each transition of the vicious DMTA cycle, ensuring seamless communication and integration across different phases and disciplines.

Consequences of the Vicious DMTA Cycle

Ineffective transition management in DMTA leads to inefficiencies and increased risk.

Productivity Loss

© 1994-2026 Revvity Signals Software, Inc. All rights reserved.



Translating from one domain to the next often requires practitioners to leverage subject matter experts to help with mapping of important, domain-specific terms from design systems, to execution systems, to analysis systems.

Transposition



Mapping often requires manual transposition and/or transcription of numbers and text into different software interfaces. Erroneous transposition can result in failed experiments, faulty analyses, or incorrect decisions.

The Virtuous DMTA Cycle in Drug Discovery

There are a variety of specific DMTA cycles required for successful candidate nomination. These involve cross-disciplinary project teams collaborating to, ultimately, promote a reasonable drug candidate to clinical testing. Disease pathology research, biological target and pathway analysis, molecular modality selection, in-vitro assay development, in-vivo study design, and test article scale-up research all require effective DMTA cycles to matriculate through to clinical studies. Fundamentally, successful lead optimization DMTA cycles will produce the "composition of matter" which organizations claim as intellectual property in support of their ultimate launch to broad patient populations. Here, we summarize the dual-purpose of DMTA cycles during lead optimization[1] in drug discovery.

AI-Enabled Drug Design—The Design Step

© 1994-2026 Revvity Signals Software, Inc. All rights reserved.

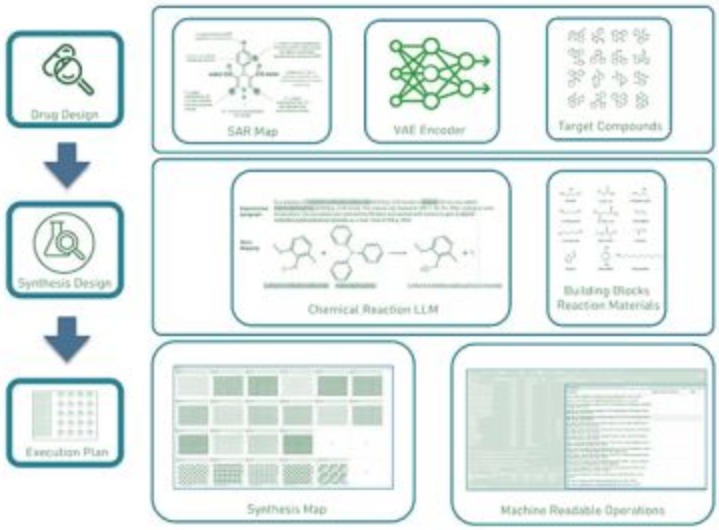


Figure 1. Machine-enabled virtuous design cycle in drug discovery.

*The **drug design** step involves the use of a SAR Map[2] and a variational autoencoder to generate a set of target compounds. The **synthesis design** step involves the use of chemical reaction LLMs,[3] retrosynthesis prediction tools, and material inventory web services to generate a set of reactions. Finally, the use of specialized planning applications to support the generation of a synthesis map with machine readable operations to support the transition to digitally supported and automated **execution** (make, in DMTA).*

The design phase of the DMTA Cycle in drug discovery addresses two key questions—what to make? and how to make it? Oftentimes innovators have solely focused on using AI to answer the first question. How AI designed test articles are synthesized, purified, and prepared for bioassay are left to manual and bespoke processes.

What to make?

Specifically, what composition of matter is most suitable for modulating the applicable drug target's biochemical function?

The virtuous approach involves the use of generative AI tools to produce a structure activity relationships *(SAR)* map. These maps, for a lead series, indicate which specific atomic and functional moieties impact biological target binding, ADME, and toxicological properties. Similarly, in the case of biomolecule modalities, SAR maps indicate how specific peptide/nucleotide/ saccharide sequences and related higher order structure (HOS) functionality influences therapeutic and ADME-T properties. Once sufficient SAR information has been obtained for a lead series, medicinal chemists attempt to optimize for potency, selectivity, and overall druggability. Modern drug discovery organizations have implemented a range of generative AI[4,5] that can generate an optimized set of target compounds.

How to make it?

To confirm the success of sequential rounds of generative AI, medicinal chemists must design efficient synthetic routes for each of the target compounds. They must also minimize the total number of reactions required to produce confirmatory test articles for all AI-recommended

© 1994-2026 Revvity Signals Software, Inc. All rights reserved.

target compounds.

Efficient approaches involve the use of well-trained retrosynthesis tools,[6] which perform the following design functions:

- Propose efficient, shared-path, synthetic routes
- Propose a set of reaction building blocks—available on-demand from either commercial suppliers or available within an organization's in-house inventory
- Summarize all input materials required to execute all reaction steps for all target compounds
- Summarize reaction parameters (temperature, time, pressure, etc.)

Medicinal chemists can then order these building blocks electronically.[7] Ideally, these requests are fulfilled in an appropriate source plate—a ready-to-react format suitable for robotics-based operations. Furthermore, medicinal chemists use the output from retrosynthesis tools to map all requisite reactions into parallel synthesis arrays. Each array unit represents an individual reaction. The contents of each source plate (or stock locations for other input materials) are mapped to appropriate locations in a reaction container. Finally, post-reaction operations (analysis sampling, assay sample prep, etc.) are also mapped from the reaction container. The complete map of the overall experiment is contained in a synthesis map representing the complete conceptual and physical design. Medicinal chemists also use applications[8] which generate an itemized, machinefriendly master procedure list that includes material dispenses, reaction operations, and sample preparation instructions. This procedure list is then disseminated to scientists to execute laboratory operations, and output as instruction files for automated robot operations.

AI-Enabled Synthesis Execution in Drug Discovery—The Make Step

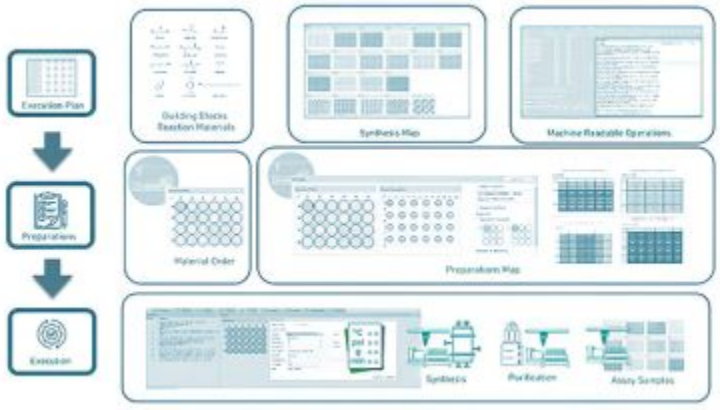


Figure 2. Machine-enabled virtuous "make" cycle in Drug Discovery.

*Leveraging the synthesis map and machine-readable instructions included in the **execution plan**, building blocks and other reaction materials are ordered and delivered in a reaction-friendly source plate format. The **preparations** map and corresponding operations are sent to synthesis, purification, and assay samples for **execution**. The digital representation of all assay samples (including applicable sample identifiers and related metadata) enables automated transition to the testing step.*

© 1994-2026 Revvity Signals Software, Inc. All rights reserved.

Upon completion of the synthesis design, the medicinal chemist can execute each unit operation— from building block dispensing, through reaction initiation, to final workup and assay sample preparation (Figure 2). Software applications relate the materials, procedures, samples, and container locations using relational identifiers. Furthermore, physical entities can be labeled with these relational identifiers, allowing for data generated before, during, and after experimentation to be associated to the appropriate digital entities defined during the design step. Due to the modularity of modern synthesis labs, the format of machine instructions must conform to the format requirements of each device supporting each unit operation. The master procedure list is segmented by device, then electronically submitted to each device's operating software interface. In accordance with ALCOA conformance standards[9] and upon completion of each operation, device log files are associated to the applicable procedure list item(s). Any variation between planned operations and executed operations is recorded.[10] Downstream operations use the logfile information to account for such variation.

The output from all executed procedures results in a set of physical output materials. These materials are electronically and physically assigned to designated containers, labeled with appropriate machine-and-human-legible material identifiers. Identifiers are related to information in appropriate software applications including output material identity, material metadata, container or vessel IDs (e.g., plate and well ID).

AI-Enabled Testing in Drug Discovery—The Test Step

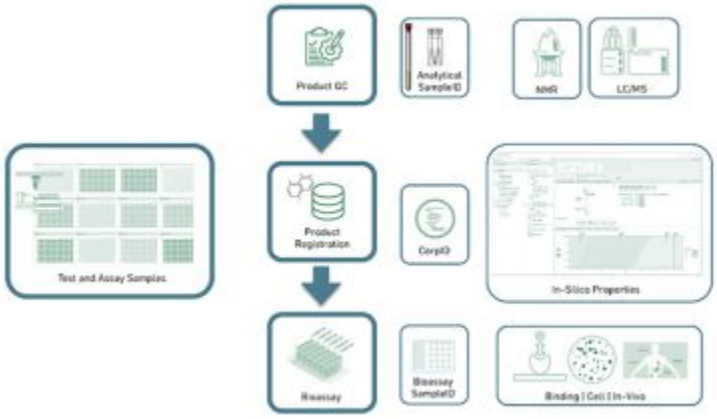


Figure 3. Machine-enabled virtuous "Test" cycle in Drug Discovery

For both product QC and bioassay, physical samples are labeled with human and machinereadable sample identifiers. Product registration generates a corporate identifier for each product structure. All assay results are related to these corporate identifiers (via relationship between sample identifiers and corresponding assay results.

Similarly to design, the testing of output materials (i.e., purified target compounds) serves a dual purpose.

1. Subjecting materials to a wide range of project-applicable bioassays to confirm the performance of the target compounds as possible (future) drugs.

© 1994-2026 Revvity Signals Software, Inc. All rights reserved.

2. Identity and quantitative compositional testing to ensure accurate SAR.[11]

As such, the modern DMTA cycle accounts for all applicable tests required to inform the overall lead optimization effort. Samples of output materials from reaction execution steps are prepared, then labeled with system-derived sample identifiers, including:

- Reaction quality control samples—type (preprocess, in-process, and post process) and preparation information.
- Product registration samples—including chemical identity, purity, salt stoichiometry, and any applicable physical form information that is required for assessing bioassay results for SAR map.
- Test article samples—including concentration, assay role (e.g., control, standard, blank, replicate number, etc.)

Software systems that support the overall DMTA cycle facilitate the capture and storage of sample information with identifiers. Additionally, these systems also generate instructions for scientists and robots for each sample preparation step (aspirate, dispense, dilute, etc.). This ensures sampleto-data integrity and the efficient preparation of all samples via automated unit operation execution.

The testing workflow starts with generation of the complete execution of test and assay sample preparation tasks during the make step. Software systems that support the test process generate analysis equipment sequence lists, which include sample identifiers arranged in the order that the physical samples are submitted to each device's operating queue. This allows the datasets generated by each test and assay device to be automatically associated to the appropriate material via applicable sample identifiers described above.

To ensure reliable and high-quality SAR, quality control (QC) analysis is performed before bioassay. Drug discovery organizations establish minimum quality thresholds as a prerequisite to material registration. The registration step (Figure 3) requires that purified reaction product information includes identity, composition, and related physical form information. The resultant characterization information can further inform the SAR. This is especially important in cases where unexpected results are obtained during testing. In cases where a target compound’s potency is significantly different from the expected SAR, having such physical form information may reveal the source of such unexpected results.[12]

Upon product registration, integrated prediction systems generate a set of orthogonal structure and physical form-related properties based on data included in the registration submission (chemical structure, salt form, stoichiometry, etc.).

Table 1. Physicochemical, ADME, and toxicological descriptors generated upon product registration.

Physicochemical	ADME	Toxicity

© 1994-2026 Revvity Signals Software, Inc. All rights reserved.

- Aqueous Solubility
- Boiling Point/Vapor Pressure
- Log*D*
- Log*P*
- pK_a
- Sigma

- Blood Brain Barrier Permeation
- Cytochrome P450 Inhibitors
- Cytochrome P450 Substrates
- Distribution
- Maximum Recommended Daily Dose
- Oral Bioavailability
- Passive Absorption
- P-gp Specificity
- PK Explorer
- Regioselectivity of Metabolism

- Acute Toxicity
- Aquatic Toxicity
- Endocrine System Disruption
- Mutagenicity
- Health Effects
- hERG Inhibition
- Irritation

Predicted properties are intended to complement the measured bioassay and QC results, enhancing the utility of the overall SAR Model. Upon receipt of confirmatory product QC results, these samples are electronically submitted to corporate registration systems. Corporate identifiers are proffered on samples upon registration approval. Bioassay samples are submitted for test execution and corresponding assays. This completes the final set of tests as test articles advance to the final analysis step.

AI-Enabled Analysis in Drug Discovery—The Analyze Step

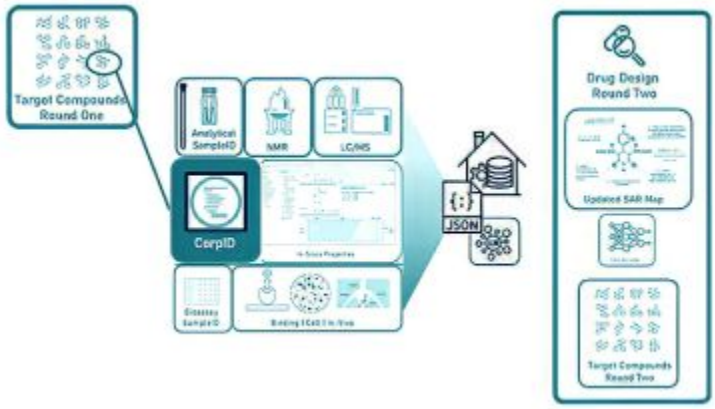


Figure 4. Upon completion of assay data analysis, results as JSON objects are used to re-train Generative AI Models, allowing for a virtuous next round of DMTA to start.

The first analysis activities support the processing and interpretation of product QC and bioassay data. Each device that is used in sample analysis produces datasets in specific (often proprietary) formats. (The variety of tools that support this primary analysis step has been described elsewhere.[13, 14, 15])

© 1994-2026 Revvity Signals Software, Inc. All rights reserved.

Modern drug discovery organizations intend to leverage the output of these analyses by:

1. Aggregating processed data into a data warehouse.
2. Implementing a rigorously enforced controlled vocabulary for all bioassay and product QC Results, as well as associated metadata structured as JSON Objects.
3. Relating the test results by applicable sample, corporate identifier, and in-silico properties (generated for product registration).

Scientists can then update applicable SAR maps based on the bioassay test results.

Value Realization of Digitalized DMTA

Large language model (LLM)-based prediction models can be automatically re-trained with JSON objects generated during analysis. Retraining of the machine learning (ML), LLM, and Variational Auto Encoder (VAE) systems can then be used to generate a new set of target compounds, for a new, virtuous round of DMTA. Without automatic retraining, data scientists spend up to 80% of their time on preparing training datasets. Data preparation time is effectively reduced to zero in the virtuous DMTA cycle.

Accelerating Clinical Candidate Nomination

Enabling the modern, virtuous DMTA cycle offers significant benefits to drug discovery productivity. By leveraging LLM-based prediction models and re-training them with JSON objects generated from data analysis, organizations can continuously improve the accuracy of their predictions. This approach allows for the generation of new target compounds and facilitates a seamless and iterative process. Consequently, the integration of structured data, controlled vocabulary, and advanced analysis tools not only enhances the efficiency of the drug discovery process but also accelerates the development of innovative therapeutics. Ultimately the virtuous DMTA cycle allows project teams to optimize lead series and arrive at molecules which meet clinical candidacy specifications, faster than ever before.

References

[1] Barcelos, M.P. et al. (2022). Lead Optimization in Drug Discovery. In: Taft, C.A., de Lazaro, S.R. (Eds), Research Topics in Bioactivity, Environment and Energy. Engineering Materials. Springer, Cham. DOI

[2] Frey, K. (2020). Structure activity relationship (SAR) maps. Curr. Pharm. Teach. Learn., 12(3), 339-346. DOI

[3] Vangala, S.R., et al. (2024). Suitability of large language models for extraction of high-quality chemical reaction dataset from patent literature. J. Cheminform., 16, 131. DOI

[4] S. Reddy N, et al. (2024, June 30). Leveraging Latent Evolutionary Optimization for Targeted Molecule Generation. 2024 IEEE Congress on Evolutionary Computation (CEC), Yokohama, Japan. Link

© 1994-2026 Revvity Signals Software, Inc. All rights reserved.

[5] Ochiai, T., et al. (2023). Variational autoencoder-based chemical latent space for large molecular structures with 3D complexity. Commun Chem, 6, 249. Link

[6] Discovery at Your Fingertips | SYNTHIA™ Retrosynthesis Software. (2025, April 16). Synthia Online. Link

[7] Modern inventory management systems allow for orders to be fulfilled from either in-house inventory or from external suppliers from integrated data sources. NRC (US). Prudent Practices in the Laboratory: Handling and Management of Chemical Hazards: Updated Ver. Washington (DC): NAP (US); 2011. Link

[8] Software to support HT experimentation workflows. Katalyst D2D, version 2024, Advanced Chemistry Development, Inc. (ACD/Labs), Toronto, ON, Canada. Link

[9] Durivage, M. (2019, Jan. 12). ALCOA Principles: Data Integrity for the FDA Regulated Industry. Quality Systems Compliance LLC. Link

[10] For laboratory operations executed by scientists, applicable software interfaces allow for contemporaneous observations and notes to be recorded by the scientist.

[11] Organizations will also subject output materials to such analyses in defense of composition of matter claims in IP Applications. Basics of claim drafting for utility patent applications. (Retrieved 2025, Apr. 22). USPTO. Link

[12] Bouchot, P., et al. (2020). Determination of the stoichiometry between a drug and its counter-ion by supercritical fluid chromatography using ultra-violet and evaporative light scattering detections: Application to ondansetron hydrochloride. Talanta, 218. DOI

[13] Mercer, K. (2023, Oct. 24). Enabling End-User HTE at GSK [Online presentation]. ACD/Labs Driving Efficiency with Spectrus Symposium. Link

[14] Standardization of Analytical Data: Best Practices. (2025, Jan. 6). ACD/Labs. Link

[15] FAIR Annotation of Bioassay Metadata. (Retrieved 2025, Apr. 22). Pistoia Alliance. Link

© 1994-2026 Revvity Signals Software, Inc. All rights reserved.

revvity signals

77 4th Avenue
Waltham, MA 02451 USA
(800) 762-4000
(+1) 203-925-4602
revvitysignals.com

Copyright © Revvity, Inc. All rights reserved. Revvity™ is a registered trademark of Revvity, Inc.
All other trademarks are the property of their respective owners.