

生物学データと化学データの 統合が創薬初期の意思決定を 支援する

目次

3

分断されている科学データがもたらす問題点

4

共通点と相違点：部門間のデータ統合が重要な理由

5

創薬データ統合がチームを成功に導く3つの方法

7

創薬データの統合：
組織の業務遅延の要因とは？

9

生物学データと化学データを
1つのプラットフォームに統合





断片化された科学データの 問題点

創薬は、生物学者、メディシナルケミスト、その他の科学者間の緊密な連携に支えられています。初期段階の創薬に関する意思決定を一致させ、プログラムを迅速に進めるためには、分野横断的なチームが共通の科学的基盤に基づいて作業する必要があります。しかし、生物学的データと化学的データが重複することはほとんどありません。むしろ、それぞれの部門が独自の科学的な視点や運用面の視点で情報を収集・整理するため、分断されたデータ環境が生まれがちです。

時間の経過とともに、データの断片化は複雑な創薬フレームワークの全体像を曖昧にし、チームが初期の知見を具体的な行動へ移すことを妨げます。不整合なデータ基盤の上では、コラボレーションが停滞し、意思決定に時間がかかり、創薬初期に重要な勢いが失われてしまいます。その結果、重要度の高いワークフローが損なわれ、パイプラインの効率や投資対効果が低下します。

データ統合は、部門間の断絶を解消し、研究者には共通の視点を、R&Dリーダーには一貫した分野横断的なエビデンスに基づく意思決定への自信をもたらします。スピードと精度が競争優位性を決定する状況下において、統合されたデータはチームがより適切な情報に基づいた意思決定を行い、最も有望なリードにリソースを集中させ、創薬後期で失敗するリスクを軽減します。



共通点と相違点： 部門間のデータ統合が重要な理由

創薬生物学者とメディシナルケミストは、問題へのアプローチ方法は異なりますが、基盤とするデータは共通です。疾患のメカニズム、薬理学的プロファイル、安全性情報などに基づき、各チームは専門領域における知見を導き出し、創薬初期の意思決定や、優先順位付け、選別を行っています（表1）。

表1.分野横断型の創薬ワークフローを支える基盤データ

基礎データ	創薬研究者による活用	メディシナルケミストによる活用
疾患メカニズム	生物学的パスウェイ、表現型、および臨床的関連性を理解し、 最適なターゲットを特定する 。	疾患メカニズムに基づき 化合物設計の方向性を定める 。
薬理学的プロファイル	疾患の状況における既知の薬理学的関連性に基づいて、 ターゲットの優先順位を決定する 。	期待される生物活性を示す 化合物を優先的に選定する 。
生物活性データ （IC50、EC50、Kiなど）	リガンド結合による活性変化に関する定量的な生物学的エビデンスを用いて、 ターゲットの妥当性を評価する 。	構造活性相関（SAR）と活性データの分析に基づき、 化合物の活性を最適化する 。
選択性／オフターゲットデータ	構造最適化によりバイオアベイラビリティを高め 、懸念事項や後期フェーズにおける失敗のリスクを軽減する。	既知のオフターゲット効果や有害作用に関連する構造の優先順位を下げ、 リスクを軽減する 。
ADME・毒性データ	ターゲット分子の調節に 伴う生物学的リスク （毒性や曝露不足など）を 予測する 。	構造最適化によりバイオアベイラビリティを高め 、懸念事項や後期フェーズにおける失敗のリスクを軽減する。

分野横断型ワークフローにおいて、生物学的・化学的な意思決定を行う上で極めて重要な役割を果たす、基盤データの概要。

生物学者と化学者は共通の基盤データを活用しているものの、同じデータソースを使うことはほとんどありません。各チームが異なるデータベースやデータ管理システムを使用していることが、部門間の知識共有を妨げています。基盤となるデータが繋がっていないと、初期段階での意思決定に一貫性が欠け、業務の遅延や開発後期における失敗のリスクが高まります。

創薬データ統合がチームを成功に導く 3つの方法

生物学的データと化学的データを統合することは、単なる技術的な改善にとどまりません。これは戦略的な転換であり、チームに治療領域の全体像を提供することで、パイプライン全体に影響を及ぼす、より情報に基づいた意思決定を可能にするものです。創薬データ統合の価値は、次の3つの領域にまとめることができます。

1. 分野横断的な連携の強化

学術論文や特許から構造化データベース、社内システムに至るまで、創薬に関連するデータは、多種多様な情報源から得られます。しかし、これらの情報源は体系的に相互検証されておらず、形式も統一されていないため、チームが信頼できる統合エビデンスとして活用することは難しいのが現状です（表2）。

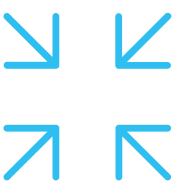
表2：一般的に利用される創薬データの代表的な情報源

データの種類	代表的な情報源
疾患のメカニズム	文献（例：PubMed）、ClinicalTrials.gov、Reactome、KEGG、DisGeNET
ターゲット・リガンド	PubChem、UniProt、ChEMBL、PDB、BindingDB、DrugBank、GtoPdb
薬理学的プロファイル	DrugBank、ChEMBL、PharmGKB、NCBI、規制機関のデータベース（例：Drugs@FDA）
生物活性データ	文献、ChEMBL、PubChem、BioAssay、自社のアッセイデータ
選択性／オフターゲットデータ	文献、PubChem BioAssay、SEA（Similarity Ensemble Approach）、GtoPdb
ADME／毒性データ	文献、eTOXデータベース、ADMETlab、PubChemのHSDB、DrugBank、Open TG-GATEs、NLMデータベース、社内の安全評価レポート

創薬初期段階の意思決定は、社内外の多様な情報源に存在する複数のデータタイプに基づいて行われ、それぞれのデータが生物学的・化学的な観点から重要な情報を提供しています。



情報を一元的に整理することで、疾患のメカニズム、アッセイ、薬理学的データといった重要な情報を研究者が解釈する際の齟齬を軽減できます。科学データが領域を超えて一貫した形で構造化されていれば、創薬チームが異なる角度からアプローチする場合でも、共通の視点を通じてデータを評価できるようになります。この戦略的転換は、解釈のズレを防ぐだけでなく、部門間の認識のズレや業務遅延を未然に防ぐことで、創薬パイプラインを加速させます。



研究の焦点を明確にし、有望でない候補化合物に費やす時間を削減します。



度重なる確認や誤解、引き継ぎの遅延、試行錯誤のサイクルを最小限に抑えます。



迅速かつ的確な「継続/中止」の判断により、後期段階で失敗するリスクを低減します。

2. 潜在的なつながりと新たな機会の発見

データの統合は、研究者間の連携を強化するとともに、これまで見過ごされがちだったエビデンスに基づく知見を明らかにします。生物学と化学のデータが連携することで、研究チームは、孤立した創薬環境では見落とされがちな、ターゲット、化合物、パスウェイ、疾患メカニズム間の関係性を明らかにできます。このレベルの統合により、チームは疾患において予期せぬ役割を果たすターゲットの同定や、見過ごされてきたオフターゲット相互作用の検出、新たな治療の観点からの化合物シリーズの再評価が可能になります。こうしたインサイトは、十分に解明されていない生物学的現象を探求する場合や、競争が激しい、あるいはリスクの高い疾患領域で意思決定を行う場合に極めて重要です。

科学的小および治療的な状況をより包括的に把握することで、創薬チームは、部門横断的な知見を適切な場面、かつ最適なタイミングで活用し、その価値を最大限に高めることができます。この統合的な視点により、プログラムのリスクを早期に軽減し、より迅速かつ的確な意思決定を可能にするとともに、従来のワークフローでは見過ごされていた新たな機会を明らかにします。



未開拓の生物学領域やケミカルスペースの**迅速な調査**



ターゲットの冗長性やオフターゲットリスクの**早期検出**



競争の激しい治療領域における**戦略的優位性の強化**

3. 重要な意思決定に、スピードと確信を

創薬における重大な決定は、直感に委ねるべきではありません。それには、生物学者や化学者が信頼できる、一貫性があり包括的なエビデンスに裏打ちされた判断基準が必要です。情報が断片化していると、意思決定を左右する重要な作用機序や安全上のリスクを見逃し、本来優先すべきでない候補化合物を選んでしまうリスクがあります。

データがつながることで断片的な解釈に頼る必要が減り、統合されたエビデンスに基づいて、トリアージやターゲット選定、プログラムの優先順位付けをスムーズに進められます。これにより、初期段階での実施可否判断に対する信頼性が向上し、リーダーは最も有望なプログラムに集中し、競争の激しい研究開発環境においてR&Dをより効果的に推進できるようになります。



一貫したデータに基づく、**確信をもった意思決定**。



限られたエビデンスしかない新規プログラムをトリアージする**能力の向上**。



科学的エビデンスと戦略的投資のより**強固な整合性**。

創薬データの統合： 組織の業務遅延の要因とは？

データ統合は手戻りや認識のズレを減らす効果があるものの、それを実現するのは容易ではありません。社内で創薬データを標準化・統合する取り組みは、しばしば研究者に本来の業務範囲を超えて負担をかけることになります。

科学データの性質上、データ統合は組織が想定するよりもはるかに複雑になります。
その理由はなぜでしょうか？





科学的なアウトプットは、連携して使われることを前提に設計されていない

生物学および化学的情報は、相互運用性を念頭に置いて構築されることは稀です。そのため、各データセットはそれぞれ独自のルール、構造、単位、慣例に従っており、結果として深刻な不整合が生じ、データ統合の妨げとなっています。特に影響の大きい不整合の要因は、次の3点です。

- **識別子の不一致**：タンパク質、遺伝子、化合物、および疾患は、データベース、出版物、ベンダー、および社内システムごとに異なる名称で扱われています。
- **フォーマットの非互換性**：化学構造、配列データ、アッセイ、パスウェイのアノテーション、および毒性試験の結果が、統一されていない形式で保存されています。
- **メタデータの整合性不足**：実験条件、プロトコル、エンドポイント、測定単位などが、分野、研究室、国をまたいで標準化されていません。

こうした一貫性の欠如は、整合の取れない、分断されたデータ環境をもたらします。

研究者には、重要なインサイトを見逃したり、間違っただけの候補を推したりすることがないように、できるだけ多くの情報へのアクセスが必要です。しかし、既存の科学情報を統合・標準化するには多大な労力を要し、貴重な時間を費やすだけでなく、創薬チームが開発候補化合物の選定や価値の高いプログラムを推進することの妨げとなります。

終わりのない科学データのキュレーション

創薬データは静的なものではありません。新しい論文、特許出願、データベースの更新が日々行われ、その追加が、ターゲット、化合物、あるいはパスウェイの解釈を変える可能性があります。新しい情報が絶えず創薬のランドスケープを塗り替えていく中で、最新の状態を維持することは、一度きりの作業ではなく、常に継続すべき重要な責務となります。

絶え間なく発表される論文、アッセイ、データベースの更新を、実用的なインサイトへと変換するには、専門的なキュレーションへの継続的な投資が必要です。これには、関連性の検証、創薬に関連する情報の更新、そして新たなエビデンスに基づく生物学的な解釈を行うための専門家の関与が含まれます。多くの創薬チームには、中核となる研究活動を止めることなく、またパイプライン全体の進行を遅らせることなく、継続的かつ高度なキュレーションを維持するためのリソースがありません。

データの統合には科学的な専門知識だけでは不十分です

創薬データを大規模に統合するには、科学的内容を理解するだけでは不十分であり、組織的な取り組みが求められます。これには、パイプラインの維持、識別子の管理、システムが新しい情報に遅れをとらないようにすることなどが含まれ、データ管理とITの両分野にわたる専門的なサポートが必要となります。さらに、専門のリソースを投入しても、科学分野における分類体系の標準化と索引付けは、膨大な作業です。適切な専門知識とインフラが整っていない場合、科学者は価値の高い研究と、資源を大量に消費するデータキュレーションに時間を割かざるを得なくなり、その結果、研究の進捗が遅れ、現場の疲弊が増大します。

生物学的データと化学的データを 1つのプラットフォームに統合

創薬データの統合にあたって、研究者がフォーマットの統一、情報の矛盾解消、あるいは文献やデータベースにおける絶え間ない更新情報の追跡に、貴重な研究時間を割く必要はありません。研究開発のリーダーが求めているのは、データ統合の負担を軽減し、チームが影響力の大きい創薬パイプラインの推進に専念できるソリューションです。CAS BioFinder®は、まさにこのニーズに応えるために設計されました。

創薬に特化して設計された**CAS BioFinder™**は、一流のライフサイエンス誌、世界の特許庁、そして信頼性の高い科学データベースから収集された、最新かつ構造化された専門家によるキュレーション済みデータを提供することで、研究の進捗を妨げるデータの分断を解消します。このプラットフォームの中核を成すのは、110年以上にわたり専門家がキュレートしてきた科学的知識を体系化した、世界最大級の科学情報資産であるCAS Content Collection™です。



ターゲット、リガンド、疾患、パスウェイ、バイオマーカー、アッセイにわたる生物学的・化学的・薬理学的情報を統合することで、創薬チームは、断片的なシステムの寄せ集めではなく、信頼できる共通の情報源に基づいて研究を進めることができるようになります。

データ統合がもたらすメリット

- 創薬情報の不整合や不備を照合する時間を短縮
- 部門横断的なチーム間で、より明確で一貫したデータ解釈が可能に
- 専門家が収集したエビデンスに基づいて意思決定を行うことで、初期段階のワークフローを推進
- 創薬研究者とメディシナルケミスト間のコミュニケーションと連携を向上

CAS BioFinderは、データの断片化を解消し、分野横断的なエビデンスを繋ぎます。これにより、創薬チームがよりの確な意思決定を行い、下流工程のリスクを軽減し、有望な候補化合物を自信を持って開発することができます。

「 詳細についてはお問い合わせください 」



CASは、世界中の科学の知見をつなげて、人々の生活を向上させる画期的な進歩を加速させています。弊社は、世界中のイノベーターが今日の複雑なデータ環境を効率よく探索し、イノベーションへの道のりの各段階で革新をもって判断できるよう支援します。科学知識管理を専門とするCASのチームは、人の手で収集・精選された世界最大の権威ある科学データのコレクションを構築しているほか、必要不可欠な情報ソリューションやサービス、そして専門知識を提供しています。さまざまな業界の科学者や特許専門家、そしてビジネスリーダーが、機会を発見し、リスクを軽減させ、そして共有知識を解き放つためにCASを活用しています。得られたインサイトがよりはやくイノベーションにつながりますCASは、米国化学会の情報部門です。

詳細はcas.org/jaをご覧ください。