



クリニカルデータワークベンチ：進化する製品カテゴリが、
質の高いデータへのアクセスを加速

veeva

サードパーティの患者データが増え続ける中で、クリニカルデータワークベンチの開発が進んでいます。このワークベンチは、データの質を確保し、アナリティクスの使い方を改善し、さらに手作業によるデータの集約、クリーニング、変換を削減するように設計されています。

どのような臨床試験も、それを裏付けるデータがなければ無効です。今日、ほとんどの臨床試験で使用される患者データの70%は、検査機関、ウェアラブル・デバイス、電子的患者報告アウトカム（ePRO）ソリューションなどの多様な外部ソースから得られるため、その有効性を確立しにくくなっています。データがCRFにきれいに収まらず、EDCで管理できないため、データ管理者は不整合を迅速に発見し、効率的に解決することが難しくなっています。

これまで、企業は大人数のデータチェッカーのチームを配備し、データの集約、変換、クリーニングを行ってきました。グローバルなチームが24時間体制で、手作業、統計コンピューティング環境、多数のスプレッドシート、数え切れないほどのメールのやり取りなどから業務をこなしています。

しかし、このようなコストのかかる取り組みを行っても、試験の遅れを防いだり、増え続ける外部患者データや複雑化する試験デザインに対応したりするのは困難です。例えば、今日のゲノミクス試験では、テラバイト単位のデータを主要エンドポイントにリンクさせることができるため、業界ではそ

れらすべてのデータをクリーニングし、アクセスする方法を改善する必要に迫られています。

さらに、手作業には一貫性があまりなく、文書化も不十分なことが多いため、データの質に懸念が生じ、結果的に規制当局から疑問の目を向けられる場合もあります。

ステークホルダーがあらゆるソースからの[信頼できるデータに迅速にアクセス](#)できるようにするために、クリニカルデータサイエンティストは、重要な分野ごとに取り組みを行っています。上流では、データの取り込み、クリーニング、変換に細心の注意を払う一方、下流では、臨床分析（可視化など）と業務分析（RBQMやダッシュボードなど）に集中しているのです。

これらの上流と下流の機能は、それぞれ新たな製品カテゴリの一部として、次の項目を実現します。

- ✓ あらゆるソースからの完全かつ同時に収集される患者データを
- ✓ データの整合性だけでなく、データの品質についても詳細なオーバーサイトを可能にし、アナリティクスの使い方を改善してリアルタイムでトレンドを可視化
- ✓ 手動プロセスでタスクを管理する代わりに、テクノロジーを活用して成果を管理できるようにクリニカルデータチームを
- ✓ ステークホルダーのグループ間で一貫性のあるワークフローを実現することで、コラボレーションを改善

さまざまなソースや領域にわたる信頼 できるデータのパイプラインを構築

過去2年間で、臨床データマネジメントから臨床データサイエンスへの転換が進んだ結果、質の高いデータに基づく迅速な意思決定の必要性が高まってきました。このデータを得るための要件が、進化しながら時に混乱も呼んでいるデータマネジメント・ソリューションのカテゴリを形作っています。臨床データワークベンチ（または臨床データハブ/プラットフォーム）は、手作業によるデータの集約、照合、変換、分析に由来するコスト、潜在的なミス、コンプライアンス違反を削減するように設計されています。

データワークベンチは、質の高い試験データのパイプラインへのさまざま部門からの迅速なアクセスを可能にし（図1）、調和の取れたクリーンなデータを可視化ツールやRBQMツールに供給することにより、部門横断的なコラボレーションを以前から改善し、試験のアジリティを向上させています。新しいワークベンチの中には、可視化と分析の機能を搭載しているものもあり、設計が改善されるにつれて、この傾向は今後も続くと予測されます。

外部からの患者データが大半を占める今日、誤ったツールに頼るとリスクや脆弱性が生じます。例えば、統計コンピューティング環境ではデータの集約と変換はできても、クリーニングはできません。臨床データリポジトリは、集約はできても変換はできず、可視化アプリケーションは、集約を効果的に処理できません。

図1
臨床データワークベンチを
使った望ましいコラボレーション



スプレッドシートでは目的に合ったツールを作れない

通常、データマネジメントチームは、EDCを使ってデータを収集し、クエリのワークフローを活用してデータの不整合を管理しています。しかし、EDCと他のデータソースの間の不整合を特定するためには、EDCからデータを抽出し、データリポジトリに入れ、リポジトリから統計コンピューティング環境にデータを読み込み、特殊なレポートをプログラムした後、スプレッドシートでレビューアーにメールで送る必要がありました。

次にレビューアーが、各不整合を見つけ、EDCにログインし直し、医療機関のスタッフやデータプロバイダーにメッセージを送って解決してもらわなければなりませんでした。

ある専門家は次のように語ります。「試験から1つのデータポイントを拾う場合、その試験期間中に平均で17回も、そのデータポイントをクリーニングしています。これは時間の使い方として正しいとは言えません。患者が新薬にアクセスするのを遅らせることにもつながります。」マルチソースの患者データを扱う環境では、より堅牢かつ合理的なアプローチを用いて、冗長なプロセスをなくし、データの全体像を把握することが求められます。

クリニカルデータサイエンスにおいては、クリニカルオペレーション、クリニカルデータ、生物統計、安全性、医学専門家などのさまざまなステークホルダーのグループ間で、より深い協力が必要になります。

クリニカルデータワークベンチは、各グループが独自の視点を持ち込むとともに、同じ同時並行的なデータに対し、生産性の高い方法で連携して作業できるようにサポートします(図2)。

図2

目標：さまざまな領域からアクセス可能な、一貫性のあるクリーンなデータ



価値とROIを生み出す

今日のデータワークベンチは、特に実用的なユースケースで使用した場合、すでにROIを生み出しており、価値を実証しています。データを1カ所で調和させることができるため、データの取り込みと集約を迅速化できた企業もあります。

変換においては、ワークベンチは手作業を減らし、カスタムプログラミングの必要性をなくしました。レビューにおいては、トラッカーを使うことなく、データ検証手順の約3分の1を自動化し、クエリ管理にかかる時間を週60時間以上も削減しています。ワークベンチはまた、企業が可視化および分析テクノロジーをうまく導入し、AI/MLユースケース用にデータを準備するお手伝いもしてきました。

このeBookでは、多様なマルチソースの患者データが増え続ける中で、質の高いデータへのアクセスを確保するために臨床データワークベンチが果たす役割について見ていきます。

臨床データワークベンチから得られるROIの例

- ✓ 人手を介さずクエリのクローズが可能
- ✓ スプレッドシートのトラッカーを排除
- ✓ リスト作成プログラミング要件への対応を不要化
- ✓ 外部の患者データの検証時に必要なステップの30%を自動化



テクノロジー環境の見通し

図3に示すように、多くの企業は、データマネジメント環境の各要素に別個のツールを使用しています。これは今後数年をかけて一元化されていくでしょう。

データワークベンチは、役割ベースのシステム機能と権限により、データライフサイクルの複数のステップやユーザーをサポートすることができます。この環境の中で、各人は最新のデータにアクセスし、レビュー、照合、可視化を行うことができます。

プロセスの次のステップは、これらのデータに基づき、臨床データとビジネスリーダーが解釈と意思決定を行うことです。これらの活動はリアルタイムのトレンド評価を目的としており、規制当局への提出に使われる臨床レポートとは異なります。

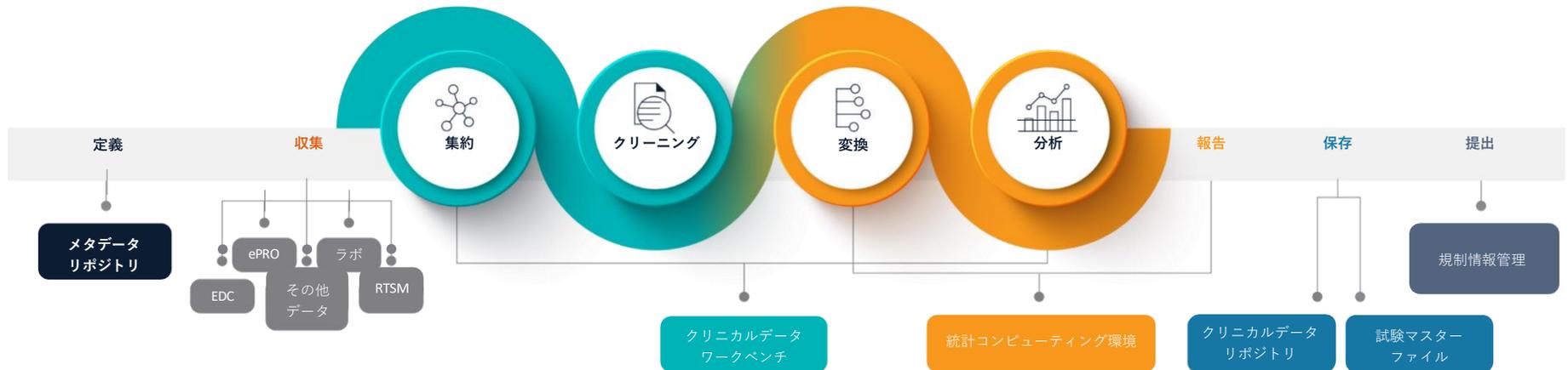
機械学習や人工知能の技術が臨床データワークフローに導入され、臨床上・業務上の意思決定に情報を提供するようになるにつれ、RBQMやRB-CDMのようなリスクベースのプロセスや、予測的・慣行的アプローチもますます使われる機会が増えています。これらの分析により、流れの中での意思決定が可能となり、「この施設では、高い問題発生率に対処するのにもっとトレーニングする必要があるか?」、「この施設を将来の臨床試験に使用すべきか?」、「新たな安全性シグナルが発生しているか?」のような質問にも答えられるようになります。

他方、臨床レポートでは、販売上の主張や規制当局への承認申請を裏付けるデータが精査されます。これらのレポートは、ある集団において薬剤のベネフィット/リスク比が許容できる水準にあるかどうか、あるいは対象薬剤がプラセボや比較対照薬と比較して有効性を示しているかどうか、といった疑問に答えるために作成されます。

分析の後に、臨床レポートは、規制当局への提出用のeCTDに収録されます。その後、すべての臨床試験データと書類は、通常、集約とクリーニングに使用されるものとは別のテクノロジープラットフォーム上に保存、アーカイブされます。

図3

クリニカルデータ製品環境



集約

さまざまなソースのデータを取り込み、統合し、調和させます。すべてのデータプロバイダー／データソースにわたり、データ取得方法を簡素化します。

クリーニング

すべての試験データについて、データのクリーニングとレビューを単一のソリューションで一元管理します。アクセスおよび盲検化を安全に管理します。データの問題の検出やクリーニングを自動化します。

変換

あらゆる下流システム向けに、相互運用可能なパッケージでデータを提供します。データの可用性を加速させ、ステークホルダーにとってのタイムラグを削減します。

分析／報告

知見の獲得と意思決定を迅速化します。レポートやダッシュボードを使用し、データ収集、施設の状況、ベンダーの活動、データマネジメント、および臨床結果におけるリアルタイムのトレンドを評価します。それとは別にクリニカルデータを分析し、規制当局への申請や販売上の主張を裏付けます。

定義

再利用可能なデータの収集、変換、および報告の定義を作成します。各試験データ要素を相互に、および定義済みの標準に関連付けます。再作業を減らし、再利用を増やします。

収集

プロトコル要件に従い、試験データを収集します。簡単で楽しめる施設体験や患者体験を創出します。

保存

データの保存や監査アーティファクトに対応でき、規制上のデータインテグリティ要件の確実な遵守を維持します。

提出

製品の市場投入までの時間を短縮します。患者に治療を提供します。試験コストを削減します。

千里の道も一歩より：業務の優先順位付け

新しいテクノロジーに関する意思決定やプロセスの改善に向き合うときは、一度に多くのことに取り組みたくなるものです。しかし、真新しいテクノロジーを組織に導入する場合、プロジェクトの成功とテクノロジーの採用に伴うリスクは相乗的に増大します。

したがって、評価時にはビジネス目標に優先順位を付け、業務の自然な順序を考慮するのがよいでしょう。

臨床データのライフサイクルは、その前の各ステップの上に成り立っています。下流のステップで求められる仕事は、上流から提供されるデータの質に左右されます。データの適切な集約、クリーニング、変換を行って初めて、可視化、リスクベースのツール、AI/MLアプローチなどの導入を成功させることができます。

マルチソースの患者データの課題に対処するためには、クリーンなデータ基盤がますます必要になります。質の低いデータにAI/MLアルゴリズムを適用すると、間違った結論が出たり、リスクが増大したりする可能性があるため、まずは質の高いデータを迅速かつ効率的に取得することに注力します。

このような環境の下では、プロセスを最適化するだけでなく、データフローの手前側から奥側までの投資先にそれぞれ優先順位を付けることが理にかなっています。

未来の状態を定義する

スポンサーやCROからすれば、クリニカルデータワークベンチの利用は、業務効率に大きな改善をもたらします。従来、クリニカルデータの管理は、まるでリレー競技のように、一人ひとりが次のチームメイトに少しずつデータを転送する必要がありました。骨の折れる手作業が求められ、誰もミスをしないことを祈るだけでした。

クリニカルデータワークベンチは、こうした状況を変革します。データ管理者をクリニカルデータサイエンスに取り込み、手作業によるタスク処理から、テクノロジーを利用した成果の管理へと移行させます。これは、クリニカルデータマネジメントからクリニカルデータサイエンスへの移行を実現するというSCDMのビジョンにも完全に合致しています。チームは今よりも早期にデータを分析し、試験の早い段階で問題を発見することが可能になります。

クリニカルデータワークベンチを効果的に導入すれば、申請資料提出までにかかる時間を短縮できます。クリニカルデータワークベンチは、質の高いデータがエンドツーエンドで流れ、すべての臨床試験関係者がアクセスすることができるデータパイプラインを提供します。

Veevaのクリニカルデータ
ワークベンチソリューションの[詳細はこちら](#)

参考資料

Alani, L., Bain, D. et al., "Functional Integration Map," Association for Clinical Data Management, 2023.