

タイトル：臨床バッチの品質データと患者アウトカムとの間の関係性の発見

著者： Valérie Vermylen, Jean-Etienne Fortier, Eric Rulier, Alain Bernard, Carl Jone,  
and Justin Neway  
(Pharmaceutical Engineering, 2017, Vol 37, No5, 60-66)

翻訳： 京都大学大学院医学研究科薬剤疫学分野 大学院生 名前：山下 将平(Shohei Yamashita)

### はじめに

バイオ医薬品の品質や製造、管理（CMC）がどのように安全性と効能に影響を与えるかを理解することは、製剤品質において大きな目標である。”ビッグデータ”を扱うことができる市販ソフトの数が増加してきていることにより、バイオ医薬品の品質特性の重要性を評価する能力は飛躍的に進歩している。これらの技術進歩は規制当局によって見過ごされてはおらず、患者の安全性や薬剤の効能に関する重要な品質特性のさらなる理解を求めている。

しかし、業界全体の技術的、組織的困難はしばしば CMC データと患者のアウトカムや製造プロセス、製品品質との相関を妨げる。なぜそうなるのかを理解することは重要である：

□バイオ医薬品企業は、部分的に規制ドライバに応じて、新たな製造技術や測定技術に加えて、クオリティ・バイ・デザインやプロセス解析工学、工程の特性化、継続的な工程の検証などの取り組みを通してデータの量の増加を引き起こす。

□医薬品企業は、患者中心の経営判断を支えることに役立つ相関のような、彼らのプロセスや品質に関するデータを統計調査や統計解析に用いるより良い方法を求めている。

□今日まで、いかなる規模の組織であっても、非常に多くのデータが依然として手動で取り込まれ、スプレッドシートに蓄積されている。加えて、構造化したデータがしばしば別々かつ互いに不完全なデータベースに保存され、集成を困難にしている。

結果、プロセスや製造作業に関する知識を改善し、この知識を維持、共有し、適切なプライバシーの水準を保証するために、データを集積し、体系化し、文脈化することが困難になっている。アメリカ食品医薬品局(The US Food and Drug Administration; FDA)は”プロセスバリデーション：一般原則と実践”という業界向けガイダンスの中で、その多くを認識している。

製造プロセスやその変化を理解することなく、適格性評価にのみ注目することは、品質の不十分な保証につながり得る。製造者は、製品をマーケットで流通させるにあたって

製造プロセスにおいて高い水準の保証を与えるにあたり、十分な理解を得ているかどうかを判断するべきである。

同ガイダンスは以下の数点を製造者たちに示している。

- 変動源を理解する
- 変動の有無、程度を見つける
- 変動のプロセス、究極的には製品特性への影響を理解する
- プロセスや製品が示すリスクと同一基準の方法における変化を調整する

Shashilov と Neway の上流のプロセスパラメータと下流での製品の品質アウトカムとの関連を示した報告では以下のことを示している：

複雑な製造プロセスにおいて容易に上流/下流の相関付けを実行することができるようになる重要な利点は、上流のプロセス条件と下流のプロセス結果との間の潜在的な因果関係を明らかにするための重大な障壁が取り除かれることである。このような関係は確認、展開、または機械論的知識や実験を拒否できる仮定の形成を引き起こす。それゆえ、上流のプロセス条件と下流のプロセス結果との関係に関して得られた情報はプロセスコントロールに用いられるプロセスモデルの重要な構成要素であり、加えて実時間適応制御(*real time adaptive control; RTAC*)を用いるための洗練されたプロセスモデルの発達に寄与している。

この研究の目的は、Shashilov と Neway の研究に影響を与えること、製造プロセスの変化の結果生じる製品の品質（特に不純度）と患者のアウトカムとの関係を見つけることである。特に、著者らは以下のことをより理解したいと考えている：

- 製品の品質のプロフィールの要因となるプロセスパラメータが臨床的必要性と見合っているかどうか
- 品質特性が患者の反応に影響を与えているのかどうか
- 免疫原性(安全性)が品質特性と関連があるのかどうか
- 患者へ承認されている不純物と関連する製品レベルがもっともらしく判断されているか

## 方法

本稿では過去の CMC や臨床データセットを用いた後ろ向き研究に関して報告する。私たちは以下の理由からこの方法を選んだ。

- デザイン研究に比べて比較的低予算であり、加えて臨床研究デザインを変えることなく、またデータ収集の必要なく既存のデータを用いることができる。
- 先達であり、前向き研究は臨床研究をデザインする以前に必要である。

## CMC データ

用いられたデータソースは、

□臨床試験に用いられた医薬品を製造した GMP パイロットスケールバッチ。紙のバッチ記録から、遊離試験や実行プロセスの情報を収集した。

□プロセス開発バッチ:スプレッドシートから、実験室内実験のデータの多くを収集した。

□CMC:内部および外部の受託製造機関(CMO)のデータ。以下を含む：

- 一般的なバッチのデータ。原料、細胞株、随伴の品質特性を含む。（重要物質特性）
- 重要なプロセスパラメーター
- 公開データと工程内管理（きわめて重要な品質特性）
- 安定性データ（例：純度）

□製剤が医療機関に輸送される間規格を維持していることを確かめるためにサプライチェーンデータを供給する。

—輸送中の温度逸脱

## 臨床試験データ

□臨床試験において用いられたキット一覧表（有効成分量の合計を満たすために、個々のキットはひとつまたはそれ以上の注射器を含んでおり、試験計画にて必要とされた）；それぞれのキットは製剤を1つまたは2つの製造バッチ及び/又はプラセボバッチから含む。

□予定された及び実際の個々の患者の治療法や用いたキットをリスト化した試験計画

□患者の特質（例、年齢、性別、肥満度指数）

□治療様式や詳細（来院日時、注入量など）

□有害事象（数、種類）

□個々の患者の治療への反応

□生理学データ（例、免疫グロブリンGレベル）

臨床チームは要求に応じて特定のデータを抽出し、本研究に組み入れた。この際、患者の秘密性、匿名性が維持され、臨床データセットが正しく解釈されたことを確認した。

## データセット系統の確立

私たちは商業上用いることができる十分に統合されたデータアクセス、収集、関連付け、解析、そして報告されているソフトウェアシステムを用いて、複数のソースから単一の整理された原理へとデータを揃えた（例えばプロセスバッチ）。これにより様々なデータの要素の発端（地理的位置、データソース、そして業務）に関する独立した有意義な比較に用いることができる単一のデータ構造が生み出された。

データの統合を平易にするために、私たちは内容というよりもフォーマットに従って統合する中間データ層をデザインした。（例えば、不連続性、複製、連続性、安定性、バツ

チ、系統性といったデータ群) これにより、論文の記録やスプレッドシートから得た際にも、元のデータソースに関わらず前後関係が失われないことを確実にした(図 1)。メタデータの数と種類は異なるデータソースから得られた。代表的な解析結果は、特定の統計手法、手法の構成成分、道具などと必要に応じて関連付けされている。原材料も供給者や等級などと必要に応じて関連付けされるであろう。平易なデータ集成を可能にするために、私たちは、全てのデータがメタデータを検索することで読み込まれ、回収できる構造を定義した。テーブルは常に製造や臨床の単位として用いられるもの(例えば、バッチナンバーまたは患者識別コード)を参照している。

異なる使用者による同時の検索を改良するために、データベースにおいて 5 つのテーブルが構成された。

1. **不連続性**：唯一かつ単一の測定。(例えば、親の年齢、バッチ製造日)
2. **複製**：複数回測定されるうちの単一の測定。(例えば、ある患者の注射日)
3. **連続性**：ある単一の製品バッチに関連した連続した測定。(例えば、バッチ製造プロセスにおける時間依存 pH 結果)
4. **安定性**：時間経過や状況変化でおこなう一連の測定のうち単一のもの。(例えば、安定性試験の間における、生物製剤の有効成分の集積レベルの変化)
5. **系統性**：工程段階の連続を経た被加工材の関連付けされたインプットとアウトプット(例えば、下流の製剤のバッチの一つに寄与する上流の原体ロット)

この方法は組織化原理を横切ってデータ値とメタデータの関連を保存し、使用者が医療機関で使われたロットによって、ワーキングセルバンクから個々のバイアルまで追跡することを可能にした。

工程の流れの系統を考慮することなしに、データから有意義な結論や相関は得られない。上記で用いたような商業的に可能なソフトウェアシステムによって供給される自動化された系図作成ツールを用いることにより、私たちは上流及び下流の重要なプロセスパラメータと、医薬品分割や混合が起こるプロセスの重要な製品品質特性をつないだ。

データセットはハードコピー、電子コピー両方の形をとった。過去のハードコピーの CMC データ(一般に CMO から)は、精度を確かめるために複写され二重点検され、上記の商業的に可能なソフトウェアシステムによって供給されるブラウザベースのデータ入力特性を用いて電子データベースに組み込まれた。

単一のデータレポジトリは元のデータソースやデータ処理アプリケーションからは切り離された。メタデータは、元のデータソースに影響を与えることなくセルフサービスのアクセスを通して抽出、保存、共有できるようデータ統合層に保存された。これは、自動でクエリーとプロセスのアルゴリズムを生み出すプラグアンドプレイシステムを生み出した。

適切なツールと方法論を用いて、CMC データや技術データの解析が臨床試験プロセス

とは独立に行われた。これらは切り離され、臨床データ処理に干渉しなかった。なぜなら、すべての解析はいかなる臨床データなく行われたからだ。

データのつながりを確認するため、臨床データセットは測定が報告された個々のパラメータを定義するための辞書を含んでもいる。私たちは工程条件や製品特性と臨床結果との相関を決定するためにプロセスモデル化ツールおよびデータ編成ツールを用いた。臨床データセットは以下のことを含んでいる：1)キットナンバーや使用情報などの用いる製品(完成品)に関する情報、2)識別コードや動員データのような個々の患者に関する情報

多くの企業にて、経験や予想、所在地、業務目標、重要業績評価指標が異なるため CMC/技術チームと臨床チームは独立して業務を行っている。私たちの方法論は2つのデータ群を結び付け、チームが共に働く手助けを行うために設計された。加えて、プロセスシステムを含んだ統合されたデータ解析も可能にし、早期の製剤の工程段階から臨床試験材料の個々のキットまで遡ることが可能になった。例を挙げると、単一の活性薬剤バッチは臨床使用される 1000 の製品キットを生み出し、個々の患者は複数の訪問を通して4つまでの異なる製品キットを受けることができるようになった。製品プロセス性能は一般に、歩溜まりや製品純度、サイクルタイムのようなアウトプットを測定することで評価される。にもかかわらず、異なるパラダイムの処理結果を解析するために、同じ数学的な統計概念の、情報科学技術のシステムやツールが用いられた。

図 2 は、原材料から患者反応までの生産プロセスの材料系図の複雑さを表しており、本研究で用いたデータモデル編成も同様である。エンドユーザーを活動基準の編成されたデータマップとして表しており、容易に用いることができるインターフェースを確保している。プロセスデータモデル構成はプロセス整備、製造工程操作、製造プロセス管理、材料系図、製品安定性、製品リリース、臨床観察、有害事象(AEs)や製品/患者の結合(系図として)を横切った解析を可能にした。

多段階の製造プロセスと臨床データを相関させることを可能にするため、工程段階の完全な追跡可能性が必要とされる。私たちのプラットフォームは、個々にそれぞれの材料取引を親子の対として解析できるように設計され、特定の親または子の区分に枝葉型にフィルタリングすることで素早いデータ検索、データ解析を可能にした。加えて、しばしば無限クエリループを生み出し長時間の検索時間を生み出していたリサイクルプロセスを取り除いた。

処理のそれぞれのタイプは特有の系図表を持っている。バッチメタデータ(工程、製品名、数)をフィルターすることで断続的な工程がつながる。

## CMC データと臨床データのつながりの理解

臨床集団は治療結果に沿ったグループに分けられた。

### 1. 治療の適応者

□Yes:治療に対する肯定反応

□No:治療に対する否定反応

2. 臨床試験の期間継続した患者

□Yes:臨床試験を完了した患者

□No:治療が中断された。(治療を完了しなかった患者が自動的に否定反応と考えられたことを留意)

3. 有害事象：異なる分類の有害事象の数

□なし

□限られた数(1-5)

□相当数(>6)

留意：ある特定の有害事象（例：皮膚障害）や臨床上の測定（例：C 反応性蛋白質）は本研究では報告されていないが確認した。

患者群における身体的特徴を相関させるために、以下のプロセスデータ解析を行うことで臨床観察や後の規格限界に影響を与える品質特性を測定した。

□パラメータの特徴づけと分布記述：基礎の記述統計量解析と形状解析を行う。

□一因子性の相関の確認：インプットパラメータがアウトプットパラメータ（例：分散分析）、相関行列、ノンパラメトリック検定、次元削減：最も影響のあるパラメータを選定する主成分解析、に影響を与えるかどうかを確認する。

□重回帰：段階的重回帰分析から選ばれたインプットパラメータのリストを用いる。影響のあるパラメータを保持し、パラメータの効果量を定めるために、段階の手順はパラメータを組み込む、または除外する。

## 結果

### 重要な品質特性

製品の品質のプロフィールを明らかにするため、私たちは製剤の発売日と薬剤承認間の品質特性の展開を推定し、その後品質のプロフィールのモデルと臨床結果を相関づけた。この方法によって治効に与える個々の品質特性のより現実的な評価が得られた。

患者へ投与されるまでの展望を予想するために、それぞれの品質特性に関する安定性モデルが用いられた。一定かつ適切な保存状態（5° C）が予想される価値を決定するために用いられた。

安定性試験は、製品試験間に評価された計測値と患者に投与される予定時間間に 3 種類の関係が確認された原体および薬剤（-70° C,+5° C,+25° C,+40° C）に行われた（表 B）：

**予測**：製品保管中の特定不純物の実時間の展開がプロセスモデルを発展させるために用いられ、そして製造から患者への投与までの様々な保管時点での個々のキットの患者への投与のデータの品質のプロフィールを予想するために用いられた。

品質特性は3つの判断基準の働きで評価された：

- 個々の患者の治療反応
- 研究期間が残っている患者
- 有害事象：異なるクラスの有害事象の数でスコア化

臨床反応間（例：有害事象、応答者、非応答者）の関係を調査するために、総患者集団、臨床試験を完了した集団、適用量、品質パラメータ値に注目した。図5は異なる状況下での特定のパラメータ値の変化を比較している。図は2つのグループに分けることができる：“治療に対する患者全体の反応”（AとC）と“臨床試験を完了した患者”（BとD）。医薬品有効成分量が100から1800ミリグラム(mg)の範囲内の全ての治療タイプ（AとB）に対して、および1200mg以上用量に該当するタイプ3の治療に対してバリエーション解析が行われた。これらのサブグループの観察は重要な変動要因を取り除いているだけでなく、研究の統計的優位性を減じている。

この相関を解析するために、以下のような複合的ツールを用いた。

- 箱ひげ図：グループ間の品質特性の分布の違いを評価する
- 回帰：臨床測定に影響を与える品質特性を評価する。それぞれの品質特性の変化範囲は応答者と非応答者間または治療を完了した患者と治療中の患者間に相関がないことを示した。

正しい統計手法を用いることで、タイプ3の治療（医薬品有効成分1200mg, P値=0.04）を完了した患者と治療中の患者間に統計的有意差があると結論付けた。しかし、サブグループ（タイプ3の治療を受けた患者と治療中の患者）のサイズは制限されており、観測される統計的有意差は大きくなかった。

### 品質のプロファイルの有害事象への影響

臨床結果は異なる方法で表すことができる：

- 量的：個々の患者の治療による有害事象の数
- 質的：有害事象が観察されると“Yes”、有害事象が観察されなければ“No”
- 半定量的：治療期間中に観察される有害事象の数（0,1-5または>6）

半定量的方法は数的相関よりもグループ効果をよりよく識別し、有害事象に注目しグループの同質性を確認するために推奨されている。

この相関を解析するために、私たちは統計的手法を用いた。

- 図7A：質的グループおよび半定量的グループ間の分布差（図5, 6）を評価するための品質特性（患者反応、中断患者、有害事象）への箱ひげ図とクラスター分析
- 図7B：品質特性の量的要因（有害事象の頻度、生物学的尺度）への影響を測定するため

## の品質特性および品質特性の組み合わせの多因子回帰主成分分析

どちらの解析も品質特性と臨床観察間のいかなる相関をも示さなかった。

私たちは品質特性を、有害事象または効能の指標いずれかに影響を与える臨床観察として分離させることが出来なかった。

### 結論

本パイロットスタディの目的は製品品質特性と臨床患者のアウトカム間の関係を理解する方法を発展させることであった。十分な総合データアクセス、集約、文脈化、解析、および報告が、臨床結果と製造プロセスデータの結び付けを可能とするために、丁寧にデザインされたデータ構成は商用ソフトウェアシステムと結び付けられた。

この方法に従うことで、私たちは質と臨床表指標（単一または複合）間の関係を、過去に用いられていた手動の方法と比較して、より容易に評価することが出来た。

治療効果、忍容性、または有害事象の観点において、製剤の製品品質特性と臨床結果間の相関はみられなかった。この結果の真価は（私たちが知る限りでは）このような論証の最初の公表された実例であることである。

本研究では、手動のデータ集約、文脈化方法の代わりに、ソフトウェアシステムを用い、人為的ミスの可能性を劇的に削減した。これは、10 から 1000 のバッチの系統的な解析をもたらした。得られた知識は容易に活用でき、他のデータセットと結合できる。

製造プロセスと製品品質データ、患者のアウトカム間につながりを作ることは最も重要な前進であった。なぜなら、患者のリスクが低下することで新たな製剤のコストがより低下し、市場までの時間が短くなるからだ。

本研究で用いたプロセスおよびツールは、製品の質と、重要な品質特性およびそれらと意味のある臨床仕様の特性との関係に関するデータに基づく意思決定を向上するような、進化した治療の安全性と有効性とをつなぐ、有用な手法を提供すると私たちは考えている。

臨床試験から上市および患者への安定供給までの製剤の経過には、膨大な量のデータの理解と解析が必要である。このように大きなデータセット（一般に”ビッグデータ”と呼ばれる）の解析はしばしば、製剤が期待される質、安全性、効能の基準を満たすことを証明するために複雑で困難な手順となる。

製剤の安定性、特性、製造能力に関する科学的理解が基となっているデータ主導の質の規格（製品およびプロセスの限界）の確立は合理的に複雑ではない。しかしながら、製品品質測定と安全性および有効性のデータを結び付けることは、一般に依然として容易な努力ではない。本研究で示したような”ビッグデータ”の扱いの発展は、科学に基づいた臨床仕様の特性の可能性を提供する。



本文以上

<図表の説明>

表 A：系統表

表 B：安定性結果をもとにした、品質特性展開モデル

図 1：データ統合層におけるシステム構成およびデータ構成

図 2：データモデルによる全行程の上書き

図 3：製剤貯蔵間の特定不純物の実時間変化

図 4：承認データにおける不純物レベルの予測

図 5：臨床パラメータ変化の比較

図 6：有害事象間の臨床パラメータ変化