

Visual Mining Studio、Visual R Platform、Deep Learner、Big Data ModuleはAlkanoとしてリニューアルしました。

細見 光一 教授 / 近畿大学 薬学部 医療薬学科 様

医療ビッグデータ草創期から 新たな価値をVMSで探索

新薬開発などで医療データの活用が当たり前になっているが、近畿大学 薬学部の細見教授はビッグデータ草創期の2010年から**Visual Mining Studio** (以下、**VMS**)によるデータマイニングを続けているパイオニアの1人。これまでの研究の道のりや、その成果の1つであるドラッグリポジショニングについて、またそれらの活動を長きにわたって支えている**VMS**について、お話を伺った。



薬学部 医療薬学科
博士(薬学)
細見 光一 教授

Interview

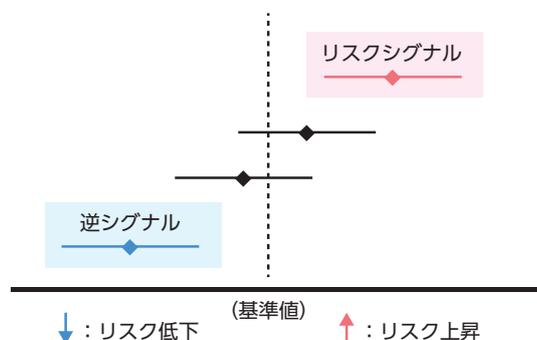
FDAデータベースの公開直後からVMSを利用

ご研究の概要を教えてください。

細見 この大学に赴任した2010年から、医薬品の安全性評価をテーマに医療データベース研究に着手しました。当時はビッグデータ草創期で、米国FDA(米国食品医薬品局)から有害事象(副作用)のデータが公開され、呼応してデータの解析手法も開発・公開され始めていました。私たちもさっそく、医薬品の副作用を割り出すシグナル検出という手法で解析にトライしましたが、データが大きくなりその項目も多岐にわたるためデータベース管理システムや表計算ソフトなどの一般的なツールでは困難でした。そこで**VMS**を使ったところ、狙った解析が容易にできたのです。そのときから**VMS**は私たちの標準ツールとなっています。このシグナル検出では、模式図のように統計的な基準値よりも大きい数値、すなわちリスクシグナルが出た医薬品と副作用の組み合わせを「関連あり」と判断します。

その後、日本でも副作用データベースの公開が始まり、またレセプトや処方箋データを研究用に取り扱う機会が生まれました。新しい解析手法も公開されたので、それらも取り入れて研究を進めました。そうした中で、複数の異なるデータベースを使い、異なる手法で解析しても、同じ結果が得られる項目があることが分かりました。これは、結果の信憑性を確認するための方法論の1つになると考えています。このように医療データベース解析の環境が拡充されるにつれ私たちの研究も進展し、現在はドラッグリポジショニングを視野に入れた研究にも取り組んでいます。

医薬品と副作用のシグナル値の模式図



ドラッグリポジショニングとは、どのような研究でしょうか。

細見 医薬品と副作用のリスクシグナルを探していたところ、基準値よりも逆の方向に大きく現れるシグナルの存在に気付きました。この逆シグナルは、ある特定の副作用や

PROFILE

細見 光一 教授

京都薬科大学卒業後、1993年、京都薬科大学大学院薬学研究科修了。日本チバガイギー株式会社(当時)、西神戸医療センター(現・神戸市民病院機構)を経て、2010年近畿大学薬学部 臨床薬情情報学分野(高田 充隆 元教授)の准教授として着任。2021年4月、医薬品情報学分野の教授に就任し、研究室で学生の指導にあっている。

臨床薬学部門・ 医薬品情報学分野 研究室

医薬品の安全性評価やドラッグリポジショニングを主な研究テーマとして、医療ビッグデータをマイニングにより多角的に解析。それにより得られた情報を迅速に臨床の分野にフィードバックするため、学会・論文発表を積極的に展開している。

[裏面に続く]

疾患を「抑える」ことを示唆しているのではないかと、この着想に至りました。データ解析により、その医薬品は本来の薬効のほかに、これまで気づかなかった効果を秘めている可能性が浮かび上がったのです。

これはとても大きな意味があります。すでに安全性が知られている既存の医薬品をヒントに、適応拡大や新薬開発への展開が期待できるからです。このことはド

ラッグリポジショニングと呼ばれ、近年の製薬業界で注目されている研究テーマのひとつです。これまで不可欠だった前臨床試験など膨大な基礎研究を短縮し、新たな効果を持った薬を低コストで迅速に開発して市場に供給できることから、私たちも製薬会社と共同で抗精神病薬のドラッグリポジショニングについて研究したほか、独自の研究も行っています。

ここまで優れたデータハンドリング機能を持つツールは、ほかに知らない

医療のビッグデータは取り扱いが難しいそうですね。

細見 有害事象データベースは、FDAのデータを例にすると、医薬品に関して市販されてからの有害事象が全世界的に把握できますが、その分母である医薬品使用の全体数がないため発生率は計れません。また7つのデータセットで構成されていますが、その構造を知らないとならば目的にあったデータを抽出できません。

また、レセプトデータは社会保険がベースなので、収録されているのは75歳未満のデータとなります。75歳以上を対象とするのであれば後期高齢者保険のレセプトデータを参照する必要があります。一方、処方箋データベースは薬局のデータですので、75歳以上も収録されていますが、患者さんが他の薬局も利用している場合、その処方を同じIDに紐付けることはできません。

厚生労働省が提供しているナショナルデータベースはほぼ全国民のデータが収録されていますが、個人情報保護の関係で使用条件や期間が強く制限されます。ですから、こうしたデータベースの特徴をきちんと理解した上で解析しないと、いくら良いツールを使っても適正な結果は導き出せないのです。

10年以上VMSにお使いになっての感想をお聞かせください。

細見 VMSによって医療データベース研究を始めることができ、現在もこうして自分の研究室を持って続けられていることをとてもありがたく感じています。当初は手応えのある解析ツールがなかったのですが、VMSを使った解析結果を発表したところ、高額な解析サービスを利用する製薬会社などから問い合わせが相次ぐなど注目を浴びました。FDAのデータは当初2,000万レコード程度でしたが、現在は1億レコードを超えたデータを扱うことがあります。PCの性能向上もあるものの、これだけデータ量が増大してもVMSなら対応できるというのは、ものすごい強みです。

機能面については、何と言ってもデータハンドリングのしやすさが挙げられます。FDAの有害事象データでは、1億レコードを超える中から狙ったデータを、プログラムを書かずともピンポイントに抽出できるのです。その抽出したデータを、マージ機能を使って紐付けていきます。例えば患者さんとその服用した薬、発症した副作用などの項目です。こうした一連の作業がここまでストレスなく実行できるツールは、私はVMS以外に知りません。

このほか、時刻処理の機能も助かっています。副作用のデータは、服薬を始めた日と副作用が発現した日を別々に持っています。それらをマージして時刻処理を行うと、服用してから何日後に副作用が出るか目安が分かるのです。このデータは医療現場で患者さんに服薬指導する際の貴重な情報となります。ほかにもメリットがありすぎて、とても挙げきれません(笑)。また、研究だけでなく指導している学生の標準ツールとしてもVMSを活用しています。

ご研究の評価や今後の展望をお聞かせください。

細見 医薬品に関して、価値ある情報を医療現場の方々にお届けしたい。私は製薬会社で研究開発や市中病院の薬剤師として勤務していたこともあり、そうした経験から常にその思いで研究に取り組んでいます。ドラッグリポジショニングでは実績ある薬に新たな効能という価値を探索し続けたいです。また、医療現場のニーズに応えるよう、医薬品の副作用など安全性の情報をさらに提供していきたいと考えています。

新たな取り組みとして、深層学習デザインツール **Deep Learner** によるAIを組み合わせた解析に興味を持っています。AIは私たちのような研究者の専門外だと思い込んでいましたが、これなら容易に使えそうです。これまでの研究データをさらに深掘りしたり、別の観点から新たな価値を導き出せたりするのはと期待しています。