

食品安全委員会評価技術企画ワーキンググループ

(第3回) 議事録

1. 日時 平成28年12月7日(水) 14:00~17:02

2. 場所 食品安全委員会中会議室(赤坂パークビル22階)

3. 議事

(1) (Q) SAR及びRead acrossの実際

(2) その他

4. 出席者

出席専門委員

(評価技術企画ワーキンググループ専門委員)

川村座長、広瀬座長代理、赤堀専門委員、岡田専門委員、小坂専門委員、山田専門委員

欠席専門委員

小関専門委員

(食品安全委員会)

佐藤委員長、熊谷委員、山添委員、吉田委員

(事務局)

川島事務局長、東條事務局次長、関野評価第一課長、鋤柄評価第二課長、池田評価情報分析官、井上課長補佐、山原評価専門官、柳澤評価専門職、比留間技術参与

5. 配布資料

資料1 「(Q) SARとは? Read-acrossとは?」(非公表)

資料2 「複数の(Q) SARモデルによる予測精度の検証」(非公表)

資料3 「(Q) SAR及びRead-across解析の実際」(非公表)

6. 議事内容

〇〇〇 それでは、定刻になりましたので、ただいまから第3回「評価技術企画ワーキンググループ」を開催いたします。

先生方には、お忙しい中、師走の慌ただしい中、御出席いただきまして、本当にありが

とうございます。

本日は、評価技術企画ワーキンググループに所属する専門委員6名の方々に御出席いただいています。

食品安全委員会からは4名の委員の先生方に御出席いただきました。

本日の議題は「(Q) SARとRead acrossの実際」というテーマです。

開催通知等でお知らせしましたように、本日の会議につきましては非公開の勉強会の形で開きますので、よろしく願いいたします。

では、事務局より資料確認をお願いいたします。

○事務局 それでは、資料確認させていただきます。

お手元に議事次第、座席表、専門委員名簿のほか、

資料1 「(Q) SARとは？ Read-acrossとは？」

資料2 「複数の(Q) SARモデルによる予測精度の検証」

資料3 「(Q) SAR及びRead-across解析の実際」

以上、お手元にありますでしょうか。

ホームページで公開しない資料がございますので、取り扱いには御留意ください。

また、机上配付資料として、専門委員の先生方には、平成27年度食品安全確保総合調査「*in silico*評価方法等食品に係る新たなリスク評価方法の開発・実用化に関する国際的な状況の調査」の報告書を用意しております。

配付資料の不足等はありませんでしょうか。

それでは、〇〇〇、お願いいたします。

〇〇〇 続きまして、事務局から「食品安全委員会における調査審議方法等について」に基づきまして、必要となる専門委員の調査審議等への参加に関する事項について報告してください。

○事務局 それでは、本日の議事に関する専門委員等の調査審議等への参加に関する事項について報告します。本日の議事について、専門委員の先生方から御提出いただいた確認書を確認したところ、平成15年10月2日委員会決定に規定する調査審議等に参加しないこととなる事由に該当する専門委員はいらっしゃいません。

以上です。

〇〇〇 先生方、御提出いただいた確認書について、特に相違はございませんですね。

(「はい」と声あり)

〇〇〇 特に問題ないということでもあります。

それでは、本日のこのワーキンググループの開催目的について説明させていただきます。

評価技術企画ワーキンググループの重要な審議項目の一つとして、食品健康影響評価に(Q) SARあるいはRead across等の*in silico*の評価手法が活用できるか。また、活用できるのであれば、具体的にどのように活用していくのかを検討することが挙げられます。これらの検討のためには、まずこの手法の現状や課題について、このワーキンググループの

先生方で共通の理解に立つことが重要ですので、これまでに御専門の先生方から講演を伺うことを中心に審議を進めてまいりました。

前回、第2回のワーキンググループでは、反復投与毒性の予測について、国内外の現状を共有するべく、米国FDA、NCTRのWeida Tong博士、そして、山田専門委員の2名の先生に御講演いただきました。その際、実際にリスク評価にこの手法を活用する可能性があるものに適用範囲を絞りまして、(Q) SARやRead acrossでの試算を試行し、経験を積み重ねていくべきではないかとの御意見を、先生方からいただいたところです。試行し経験を積み重ねるとしても、どの化学物質を対象に、どのエンドポイントの予測を試みるのか、予測結果をどう取り扱うかなど、その戦略が必要になります。

本ワーキンググループでも、今後、試行について審議をしていくべきであろうと考えますが、食品安全委員会では(Q) SARやRead acrossの経験があるわけではありません。まずは、既存の予測ツールでケーススタディーをしてみて、関係者間でこの手法をどのように実感するか、試行のあり方をどうしていくべきということを検討する上でも、このケーススタディーは非常に大事であろうと思います。

本日は、最初に(Q) SARやRead acrossに関する基本的な内容を共有するために、〇〇〇から(Q) SARやRead acrossの基本的事項について御講演いただきます。続けて、〇〇〇から(Q) SARツールの予測精度の現状について御講演いただきます。最後に、〇〇〇から、予測ツールを使ったケーススタディーを幾つか御紹介いただき、本日参加している皆さんにコンピューターが予測した結果をエキスパートがどのように判断するか体験していただく予定です。よろしいでしょうか。そういう趣旨で、本日は行っていきたくと思います。

それでは、議事の(1)について、〇〇〇から資料1について御講演をいただきたいと思います。その後、質疑応答を行っていきたくと思います。

それでは、〇〇〇、よろしく申し上げます。

〇〇〇 〇〇〇と申します。

本日はこのような場をいただきまして、ありがとうございます。(Q) SARとRead acrossについて、簡単に説明させていただきたいと思います。できるだけゆっくりしゃべるようにしますので、速かったら言っていただければと思います。

(PP)

皆さん、ぱっとこの物質を見て毒性はわかりますでしょうか？私は自信があるという人はいらっしゃるでしょうか？通常、こういう物質を見たときに毒性情報を得るには、文献調査をして、それを読んで、どうかなということを細かく見ていくと思います。しかし、検索しても情報がないということもよくあるのではないのでしょうか？そのとき、皆さん、どうしますか？誰かがデータを出してくれるのを待つということが多いのではないかと思います。しかし、そういうこともなかなか難しいというときに、何とかして情報を得られないのかなと考えるのではないのでしょうか？そういったときに使う手法が、今日お話しする手法になります。

(PP)

今日お話しするのは、化学物質の構造、いわゆるベンゼン環、亀の甲とか、ああいうものですね。あの構造に基づいて、コンピューターが毒性を予測する方法についてのお話をします。毒性を予測する方法はいろいろあります。これがいいとかはありません。いろいろな方法があって、その中の主なものが (Q) SARとRead acrossというものになります。この基礎編では (Q) SARモデルとRead acrossがどのようなものか、まず雰囲気をつかんでいただきたい。あとは、データベースをつくろうというお話がありますので、データベースと (Q) SARソフトとRead acrossの関係について、簡単に理解いただければと思います。

(PP)

今のお話の中で、既に (Q) SARモデルとか (Q) SARソフトなどと、用語で混乱する部分もあるかと思うので、簡単なスライドをお示します。

まず、構造活性相関というのは、化学構造というものと毒性の間にどのような関係があるかなということ解析するのが狭義の (Q) SARの定義です。そのときに定量的に解析するのがQSAR、定性的に解析するのがSARになります。私はさっき (Q) SARモデルという単語を使ったのですが、この関係性を一般化した式やルールにしたものをモデルといいます。そして、このモデルをコンピュータープログラムにしてくれたものがソフトウェアという形で、一応使い分けがあるのです。ただ、このモデルとソフトウェアに関しては、特に市販ソフトに関しては、モデルとソフトがほぼイコールになっているということが実際使われている現状かと思っています。

(PP)

今日のお話のアウトラインですが、なぜ (Q) SAR、Read acrossを使う必要があるのかというところの国内外動向の簡単なもの。あとは、前提となる哲学というものがあるので、それについてお話をします。そして、(Q) SAR、Read acrossについて、それぞれお話をしていきたいと思います。

(PP)

まず、本日、先ほど〇〇〇のお話の中に *in silico* というというタームがあったのですが、毒性知見を得るための選択肢として、皆さん、よく使うのは、*in vivo*、*in vitro* というデータかと思っています。この並びでいく *in silico*。コンピューターはケイ素でできていることから *in silico* というタームが出てきています。この流れの中でいくと、*in vivo* というのはコストでいくと100万単位、期間でいったら月単位、動物を必ず使わなければいけないし、被験物質もグラム単位でということで、処理できる物質がすごく少ないです。*in vitro* になると、それがだんだん減弱していくのですが、*in silico* のすてきなところは、コストに関してはメンテナンス費用ぐらいでいい。期間に関しては、秒殺です。本当にびっ、ぱっと出てきます。それから、当然コンピューターでやるときには、動物は必要ないですし、被験物質も要らない。びっ、ぱんなので、大量の物質を処理できます。根本的な違いは、こちら2つ(補足：*in vivo*と*in vitro*)は実際に測定をしますが、こちらの *in*

*silico*に関しては、予測であり、実際には測定をしないというところにあります。

(PP)

こういった手法をどういったところに使えるのかというところですが、よくあるのがスクリーニングや優先順位づけ。沢山ある物質の中から優先度の高い物質を絞り込んだり、選定します。

また、試験ができない物質、例えば被験物質そのものが入手できない。不純物とかは、特にそうだと思うのです。そういうものについては、構造の情報だけを入手すれば予測ができるといった利点があります。

また、データギャップの穴埋めもあります。試験はできるのだけれども、誰か実験データを取得してくれないかというステータスの物質がきっとたくさんあると思いますが、そういったものに対して、予測した結果を適用することができます。

さらに証拠の重みづけ (Weight of Evidence) というアプローチが最近、国内外で使われるようになってきました。この際には、メカニズムに関する情報を提供してくれる要素もありますので、そういった材料としても有用な情報が得られます。

追加データを要求する際の根拠としての利用もあります。 *in silico*で予測して間違いなく黒っぽいのが、これで規制してもいいですかというときに、データを出す人がいたら、そのデータを採用する。そういった形のデータ要求の材料にもなります。

特に、化粧品において、動物試験が実施できなくて、適切な *in vitro*の試験がない。化粧品の業界ですと、できるだけ *in vivo*の試験をやらないということになっているのですが、それを代替する *in vitro*の試験もないので、 *in silico*が選ばれていっているところもあります。

繰り返しますが、最大の *in silico*の利点は、試験をしないで情報が得られるということにあります。

(PP)

国内外の活用状況に関してです。食品に関しては、欧米では既に一部の食品評価に行政が実際に使っている事例がございます。

医薬品に関しては、古くからドラッグデザイン等において使用されておりました。また、これはすごく大きなインパクトだったのですが、ICH M7による医薬品中の不純物の変異原性評価に関するガイドラインができました。このガイドラインにより、(Q) SARによる予測評価が初めて事業者に義務づけられました。これは非常にインパクトが大きく、ソフトが売れると思ったソフトメーカーは一気に真面目に開発を始め、できるだけみんなが使いやすいようなソフト開発が始まりました。そういう意味でも、これはすごくインパクトのある出来事だったと思います。

農薬の分野に関しましては、アメリカでは既に行政 (米国EPA) が審査等に活用していますし、将来的にもっと取り入れていこうということで、(Q) SARガイダンス等もつくっております。

さらに、化粧品におきましては、欧州で動物試験廃止という動きがありますので、ICCR（化粧品規制協力国際会議）等におきましても、(Q) SAR利用について、今、積極的な議論が進められているところです。

一般化学品につきましても、欧州の化学品規則REACHにおいて、できるだけ使いましょう、動物試験は最後の手段で勝手に試験しない、試験をするのだったら、まずは*in silico*だったり*in vitro*のデータを提供した上で所轄官庁の欧州化学品庁がオーケーをしてから初めて動物試験ができる、という法的な根拠がある中でかなり使われています。

また、米国ではTSCAという化学品規制があるのですが、ここでも(Q) SARを活用した審査が既に長く実施されてきています。実際に(Q) SARを使わない審査ができないという状況がありましたので、アメリカは自分たちで自らソフトウェアを開発して、それをみんなに使ってもらうために公開もしています。

日本の化審法におきまして、分解性や蓄積性の評価でも活用しているといったところが現状としてございます。

(PP)

開発のトレンドですが、古くから文献情報や公共データベースがありますので、その情報に基づいて(Q) SARモデルが開発されてきた経緯があります。特に、(Q) SAR開発に力を入れてきたのはアメリカです。特に、アメリカのEPA、これはレギュレーションの中で使わざるを得ないシチュエーションに追い込まれていたもので、ものすごく頑張ってたっていました。

FDAのほうでも、自分たちがもらえる、ヒト健康影響に関するデータを市販のソフトウェア会社に提供して、共同で(Q) SARソフトをつくるということをしてきています。

そういった経緯がある中で、2007年にOECDで(Q) SAR検証の原則というものができました。これは一般化学品の分野なのですが、産業界から(Q) SARをもっと利用したいという声が非常に高まってきたということと、欧州で動物愛護の観点で、できるだけ*in silico*化したいといった動きがある中で、どうやって*in silico*、(Q) SARツールというものをもっと活用できるのかということについて、(Q) SARの結果をより行政が利用するために何を考えたらいいのだということについて、OECDがまとめてくれました。これによって、予測結果の信頼性等に関する項目の共通認識ができました。これが非常に大きなインパクトがある。これがあることによって、以降のICH M7等もそうですが、(Q) SARを行政が利用しやすい環境ができてきたのではないかと思います。

下から2番目、先ほども言いましたけれども、ICH M7、(Q) SAR利用が義務化されたというものを契機に、(Q) SAR販売会社がかなりソフトウェアの開発に力を入れて、今までかなり専門家でない判断しにくかったものが、(Q) SARがよくわからない人も何となく判断しやすくなるような改良がかなり進められてきました。

しかし、(Q) SARでは予測できない化学物質もあるのです。そういったときにどうかというソリューションとして、Read acrossがあるのです。そのRead acrossという手法

をやるためのツールとして、OECD QSAR Application Toolboxも開発されてきました。

(PP)

さて、ここまでがざっくりとしたイントロでした。これからがRead acrossと (Q) SARをやるための、メソドロジーの位置づけについてです。

(PP)

先ほど、*in silico*という手法が予測手法ですと言った中に、この (Q) SARモデルとRead acrossという手法が入ります。Read acrossは、非常に広く見たら、コンピューターを使ったほうが作業が早いので*in silico*のツールに入るのかなというところはあるのですが、少し狭く使うときには、*in silico*イコール (Q) SARで使っているケースもかなり多くあります。つまり、*in silico* にread-acrossが含まれないことがある。

この (Q) SARの中には、先ほどもお話ししましたが、定性的な相関を見るSARと、定量相関を見る定量的構造活性相関 (QSAR) があります。

(PP)

それから、この*in silico*のやり方、(Q) SARに関しては、知識ベースやエキスパートシステムという表現を使ったり、統計ベースというやり方もあります。知識ベースやエキスパートシステムというものは、関係性を見出す際に、客観性だけではなくて、専門家の知識も既に織り込んだ形でシステム化しています。この方法のいいところは、既知見を最大限に取り込めるということです。例えば○○○の頭の中に入っている情報をできるだけコンピューターの中に入れてあれば、皆さん、すごく助かると思うのです。要するに、統計学的にやったものだけでは説得力に欠けるが、例えば○○○の専門的判断があれば、説得力が増す。そういった情報が知識ベースの中には含まれていると考えてください。

統計ベースに関しては、関係性を見出す際に統計学的アプローチだけを使っている。これは客観性があるって、すごくいいにはいいのです。基本的に知識ベースでカバーできない知見を入れられる。余りないと思うのですけれども、万が一、○○○が知らない情報があるかもしれない。そういうものに関しては、統計ベースでカバーができるという意味で、統計ベースというものが有用になってきます。こうした、特徴の違いを組み合わせると、お互いにいいところができるのではないのでしょうか。

(PP)

もう一つ、(Q) SARに関しては、別の言い方がありまして、Local (Q) SARモデルというものと、Global (Q) SARモデルというものがあります。Local (Q) SARモデルというのは、構造が似たもの同士を集めて、その中でモデリングをする方法のことを言います。メリットとしては、構造の類似性が高いため、類は友を呼ぶではないですが、似た構造同士は同じ性質を持っていて、非常に精度が高いモデルができます。

デメリットは、類似性がない物質には使えないため、適用性が非常に限定されてしまうことです。

一方、Global (Q) SARというのは、構造を限定せずにモデルを構築するため、メリット

としては、非常に幅広い物質に適用できるモデルを構築することができます。しかし、デメリットは、Local (Q) SARに比べると精度がどうしても落ちてしまうことです。

Local (Q) SARモデルでは、適用性が落ちるというデメリットに関して、様々な構造に関して、小さなモデルをたくさん作成し、それを組み合わせることで、適用できる物質の範囲を広げるやり方もあります。例えば、後ほど御紹介します、〇〇〇というソフトウェアは、そういうタイプのモデルになっています。

(PP)

Read acrossや(Q) SARを実施する際に前提となる哲学なのですが、Similarity Principle、類似性の原則といいます。類似の構造は類似の性質を持ちますということです。何でここにわざわざカエルを描いたのかというと、このカエルとこのカエルは似ていますか？似ている気がする。では、こっちのカエルとこっちのカエルが似ていますか？ある人は、似ていると言う。一方、違うカエルだという人もいる。このように、類似性をどう考えるのかによって、予測の精度や結果の解釈が違ってきます。できるだけ似ているもの同士を選べば選ぶほど、性質としては似ているものになってくるというものになります。

(PP)

類似性に関しては、カエルでは話にならないので、化学物質に落とし込もうとすると、基本的な考え方として、まず、共通の置換基、部分構造があるか？例えば、アルデヒド基かあるか、水酸基があるか、エポキシドがあるかといったことがあります。

それから、構造が、付加的な変化あるいは一定の変化、具体的に言うと、炭素鎖が伸びていくといった、炭素鎖が1個だったのか、2個なのか、3個なのか。これは少し伸びていくだけでだんだん水に溶けにくくなっていくという一定の変化をしていく。これも、類似性における一定の変化のカテゴリーに入ってきます。

それから、共通の前駆体または分解物が生成するような、酸、エステル、塩といった観点もあります。

さらに、機能的に類似の物質として、例えばアセチルコリンエステラーゼ活性阻害があるものといったもので、例えば有機リン系のものだったり、カルバメート系、構造的には全然違うが、最終的にもたらず影響が類似しているものに関しては、共通に扱うことができるという観点もあります。

基礎編ですので、ここで押さえてほしいのは、(Q) SARでもRead acrossにおいても、類似性というものは非常に大事だということ覚えていただければと思います。

(PP)

では、早速 (Q) SARについてです。まず、(Q) SARとデータベースの関係についてです。(Q) SARモデルを作成する際に、トレーニングデータセットというものが必要になります。このトレーニングデータセットは、基本的にはどこかの文献や試験報告書からの誰かがキュレートしたデータベースの中から、作成した (Q) SARにあったトレーニングデータセットを作成します。例えば急性毒性をやりたいのか、発がん性をやりたいのか、変異原性を

やりたいのか。その中でも、変異原性といっても菌株ごとにやりたいのか、テストガイドラインごとにやりたいのか。そういったコンセプトを決めた上で、データベースから必要な情報を抜き出してきて、予測に使うためのデータセットをつくります。

このデータセットに関しては、モデルの作成に用いるデータセットをトレーニングデータセットという言い方をします。

先ほど冒頭でもお話ししましたが、化学構造と実際の結果にどのような関係があるのかを見て、(Q) SARをつくっていくのですが、この関係性を見る方法は非常たくさんありますので、今日はどのような方法がありますということはお話ししません。この関係性を見る方法によって、でき上がる(Q) SARモデルは当然違ってきます。トレーニングデータセットが同じであったとしても、例えば〇〇〇が用いた関係性を見る手法と、〇〇〇が用いた関係性を見る手法が違えば、当然できあがるモデルが変わってきます。どちらの性能がいいかは、やってみてから見てみましょう、というものです。

また、用いるデータセットによっても、(Q) SARの性能や精度、予測できる物質の範囲も変わってきます。

(PP)

具体的には、まず、こういったトレーニングデータセットがありました。仮に Ames のモデルをつくらうかなというときに、化学物質に関して、この構造を入れてあげると、例えば脂溶性のパラメーターになります log Pow なども簡単に計算ができ、情報としても簡単に入手できます。また、この化学構造を数値化する記述子 B とか記述子 C というものも計算的に求めることができます。この記述子の種類や計算の方法もいろいろあるのですが、今回は割愛します。log Pow、記述子 B、記述子 C を用いて、データセットをつくったとします。ところが、このデータセットの中身で、例えば結果の解釈に誤りがある。これはポジティブと言っているけれども、本当はネガティブですとか、構造に誤りがあつたら、この記述子の計算も全部変わってきてしまいます。そうすると、先日 Tong 先生もおっしゃっていましたが、garbage-in, garbage-out ということで、ごみのデータを入れたらごみの結果しか返ってこないの、できるだけデータセット、データベースがまずしっかりしていることが大切です。トレーニングデータセットが最終アウトプットに影響してくる要素があることを覚えていただきたいと思います。

(PP)

もう一つ重要なところが、トレーニングデータセットのケミカルスペースです。ここにプロットしてありますのは、例えば log Pow、記述子 B、記述子 C というものに対して、それぞれの化学物質の特性をプロットしたものです。この線の引き方もいろいろあるのですが、仮にこういうような線を引くと、これがこのデータセットからつくったモデルの適用範囲ですという考え方になります。ここに例えばこの緑のポイントで記した化学物質群を予測したいと思っても今、囲った範囲内に入っていない。これがもし食品安全委員会がこれから予測したいというケミカルスペースがこの緑のドットだった場合、この

モデルで予測するのはやめておいたほうが良いという考え方になります。逆に、あの中に入るモデルであれば、使っても良いということになります。

このように、(Q) SARモデルが信頼性を持って予測できる範囲を、トレーニングデータセットが決めます。このような観点からも、トレーニングデータセットというものは非常に重要であるということを御理解いただきたいと思います。

(PP)

これからOECD (Q) SARモデルの検証の原則についてお話ししたいと思います。

(PP)

これから実際に(Q)SARを使っていきましょうとなったときや、実際に測定をしないもので結果を判断していくとなったときに、それが本当に正しいのかどうかということが非常に重要になってくると思います。このときの判断基準になる共通認識(OECD (Q)SARモデル検証の原則)をOECDが構築してくれました。これは、タイトルの中に「for regulatory purposes」ということも言っているので、行政利用する際には、できるだけこれを考えていただきたいということがOECDの考え方になっています。

この原則は、全部で5つの原則から成っています。

原則1が、定義されたエンドポイント。何を予測するのですかということ。

原則2が、明確なアルゴリズム。どのような方法で予測するのかということ。

原則3が、定義された適用範囲。どのような物質が予測できて、どのような物質が予測できないのかということ。

原則4が、適合度、頑健性、予測性。どの程度の予測性能があるのか。100%当たるのですか、60%なのですかといった類いのことです。

原則5が、可能な場合には、メカニズムの解説。どうして予測できるのかということについて説明しなさいといったものが原則になります。

(PP)

具体的に、少し中を掘り下げますと、原則1の定義されたエンドポイントについては、例えば単純に急性毒性値ですというのではなくて、ラットの経口のというところまでつきます。単位も当然重要ですが、例えば特定のテストガイドライン由来のデータセットからつくったモデルであれば、予測結果もそのテストガイドライン特異的な値を出すことになるため、どういった試験法から取得したデータなのかということも、一つ、エンドポイントを定義する情報として必要になってきます。

(PP)

次、原則2としては、明確なアルゴリズムです。このアルゴリズムは、先ほどあった化学構造と毒性の関係性を見るときに使っている手法のことを指します。このアルゴリズムというものは、予測値がどのように推定されたのかということを確認して、計算が再現できるように詳細に述べられるべきものということになっています。例えばQSARモデルなのか、エキスパートシステムなのかといったことです。そして、さらにQSARモデル

だったらどのような予測式なのか。予測式の中に使われている変数は何か？さっき言ったところのlog Powや記述子B、記述子Cといったものです。そういったものを含める必要があります。

(PP)

原則3の定義された適用範囲については、先ほど少しグラフでお示したあのお話です(スライド18)。先ほどのお話は、記述子ベースでのプロットをお見せしたのですが、適用範囲を考慮する際には、そのほかに、構造のフラグメントの範囲であったり、メカニズムの範囲であったり、代謝の範囲といったことも考慮する必要があります。特に、この原則3は、モデルによる予測が信頼できるのかどうかを担保する上で、非常に重要な考え方になっています。

(PP)

原則4の適合度、頑健性、予測性に関して、です。これらは(Q) SARモデルの性能評価に関する項目になっています。この(Q) SARモデルの性能評価は、内部検証といって、トレーニングデータセットと言う、モデルを構築する際に使用するデータセットによる検証と、トレーニングデータセットの中には含まれていない物質を使ったデータセット、通称テストデータセットといい、これを使った外部検証の2種類に分けられます。

このうち、内部検証には、適合度があります。前回の検討会で、確か○○○からグッドネス・オブ・フィット(goodness-of-fit)は何かという御質問があったと思うのですが、これが適合度です。トレーニングデータセットにおいて、どの程度良好に予測できるかといった尺度になっています。例えば重相関係数のRスクエア(R^2)というものになります。

それから、頑健性、ロバストネス(Robustness)というタームが使われるのですが、予測に使用するパラメーターの安定性です。これについてはすごく解りづらいので、後ほどこのような感じのものですということをお見せします。

外部検証で評価するのは、予測性です。皆さん予測したい物質というものは、当然、モデルに組み込まれていないことが前提になりますので、テストデータセットに含まれていない物質で、いかによく予測できるのかということが、一つ、性能については大きなパラメーターになってきます。

(PP)

後ほど、○○○のお話の中に、Cooper statisticsを使ったお話が出てきます。例えば感度や特異性というタームが出てきますので、参考資料として、後ほど眺めていただければと思います。

(PP)

最後、原則5が、可能な場合にメカニズムの解説をする、というものです。この「可能な場合」とついているのがなかなかみそでして、この原則がつくられるときに、必ずしも説明できないという話もあって、「if possible」という言葉を入れることで妥協されたということがあります。そうは言いながらも、非常に重要な原則です。基本的には、何で予

測できるのかということが理解できなかつたら、信頼性の担保がしづらいということもありますので、最近はこれを重要視してきているという傾向があります。例えばSARの場合、ある構造フラグメント（アラート構造）がヒットした場合に、Ames陽性という判断を例えばします。なぜ、そのフラグメントヒットが、Ames陽性を示唆するのかということについて、例えばたんぱく質と結合するからだ、DNAと結合するからだというロジックがあると、説得力があると思うのです。そういったものを説明しなさいというのが、この原則 5 ということになります。

(PP)

今、若干難しいお話をしたので、原則を理解しやすくするために、どうやって (Q) SARモデルがつけられているのかということを見ながら、(Q) SAR原則との関係性を見ていきたいと思います。

(PP)

皆さんの場合ですと、ヒト健康影響を普段やられているので、魚類だとイメージがわきにくいとは思いますが、魚類の毒性は比較的予測しやすいのです。ですから、一番わかりやすい例として、本日は魚類を使った事例を紹介させていただきたいと思います。

ある種類の魚が96時間でどれぐらい死亡するか、LC₅₀を達成するためのモデルをこれからつくってみようと思います。

この急性毒性のLC₅₀というものは、疎水性指標であるlog Powと相関することが知られているというところから、モデル構築を始めたいと思います。本来は、log Powと相関があるという知見を得るところから始めるのですが、今回はそこを飛ばしてやりたいと思います。

(PP)

(Q) SARモデルをつくる時の流れです。

まずは、先ほどお話ししましたが、(Q) SARをどういうときに使うのだということを実際には想定した上でモデルを構築していきます。それが決まった上で、トレーニングデータセットを準備して、実際にモデルを作成し、トレーニングデータセットを使った内部検証をして、テストデータを使った外部検証をする。それで、なかなかいいなと思ったら、ユーザーに性能を確認します。それで、だめと言われたら、またこのモデル構築の4段階のフローに戻っていく。各段階で自分たちが性能等に疑問を感じれば、1ステップ前あるいは2ステップ前に戻りながら、リモデリングを繰り返し、よりよいモデルをつくっていくのが本来の流れです。

(PP)

今回のトレーニングデータセットは、ここに書いてある全部で9物質です。モデルをつくらうというときには、まず実測値のデータが必要になってきます。また、相関づけをしようとしているlog Powのデータが必要になります。これらが今回のトレーニングデータセットになります。もう少し説明しますと、今回対象とする化合物群の、ある魚種における96時間のLC₅₀値で、OECDガイドライン〇〇〇で取得されたデータです。これらは、OECD検

証の原則における「定義されたエンドポイント」というものになっていきます。

(PP)

では、実際にモデルをつくってみたいと思います。非常にシンプルなモデルにしたいので、今回は単回帰によるモデル作成をやりたいと思います。さっきのデータ（補足：トレーニングデータセット）をプロットしてあげると、美しくほぼ直線上に乗っています。この回帰モデルとしては、ここに書いてあるようなYイコール何とかの予測式が出る。これがいわゆるモデルというものになります。

原則2の明確なアルゴリズムということに関しては、今、使った解析アルゴリズムやデータセットの話が明確なアルゴリズムというところに入っていきます。

さらに、モデル構築したときに使った変数、このYイコールのlog Pow、このlog Powの値さえ入れてあげれば、毒性値が出てくるというモデルを作成できました。このモデルをつくったときに使ったパラメーターのlog Powの範囲は、この-0.66から4.53ということになるので、定義された適用範囲ということに関しては、この範囲の中の予測であれば信用できますよという考え方になります。

また、今回対象としている化合物群であるという前提も適用範囲の一つの要件に入ります。

(PP)

続いて、今、作成した予測式の検証してみましよう。まず、トレーニングデータセットを用いた検証、内部検証です。皆さん、エクセルで相関係数をとると思うのです。それで「 $R^2=0.995$ 」、今回のモデルではこれがいわゆる適合度という値になります。

それから、内部検証としてよくやられる、リーブ・ワン・アウト (leave-one-out)、交差検証という方法です。どういう方法かというと、このトレーニングデータセットの中から1物質を取り出して、残りの8物質分でモデルをつくります。抜いた物質を入れて、合いますかという検証をします。これが、リーブ・ワン・アウトというやり方で、2番目を抜いて2を予測する、3番目を抜いて3を予測とやって、ずっとやっていき、9番目を除いて9を予測して、最終的にここに書いてある小難しい式に入れて、 Q^2 という値を出して、よく言われるのが、0.7以上だったらなかなかいいモデルではないと言われるのですが、これが頑健性です。

これらが、OECD (Q) SAR検証原則の4で言うような適合度と頑健性の2つパラメーターに関する説明になります。

(PP)

最後、外部検証を使った検証です。今回、データの的に2物質しかなかったので、2物質で予測したいと思います。

ここの青四角で書いてあるものが予測値になります。ところが、実測値をここにレイアウトしてみると、1つ目の物質については、いい感じに予測できているなというのがわかると思うのですが、もう一つに関しては、かなり実測との乖離があることがわかると思

ます。なぜだと思いますか?皆さん、薄々感じると思うのですが、ここはパラメーターの適用範囲外です。ここまでだったら信用するけれども、ここから外れたら信用できないよという要因がここに一つあります。魚の場合には特になのですが、脂溶性が上がれば上がるほど、水に溶けにくくなっていく。そうすると、体に入っていくにくい。そうすると、毒性も発現しにくいということで、溶解度の関係もあって、この辺になってくると、もう毒性が出てなくなるのです。そういう意味で、ここで予測の範囲を決めたというのは非常に重要です。少なくとも、ここからこの範囲においては、きちんと関係性が見出されている。ここから先、強くなるのかなと思ったら、今、申し上げたような要素によって、毒性はなくなってしまうということがあるので、適用範囲を考えるということは、予測結果の利用において、非常に重要になります。

外部検証に関しては、重要なことは、適用範囲の中で外部検証をしなければいけないということです。こういうもの(補足:今回のケースではlog Powが適用範囲外の物質)を検証に使ってはいけません。こういうものを使ってしまうと、モデルが悪いのではないかとなってしまったため、モデルの性能を判断するときは、適用範囲のパラメーターの中で予測するというのが本来の姿です。今回、悪い例の御紹介ということになります。

そして、この外部検証の結果が予測性という言い方になりまして、本来であれば、この小難しい式に入れて、こちらの Q^2_{ext} を算出します。今回の例だと、2物質しかなかったら、この Q^2_{ext} は算出しておりません。

(PP)

これで、OECDの(Q) SAR原則や、(Q) SARの性能を判断するための何となくのイメージを掴んでいただけたのかなと思います。

ここで紹介するのは、ヒト健康影響にかかわるソフトウェアの例です。かなりのエンドポイントに関して、いろいろなソフトウェアがあって、丸が入っているので、すごく予測できるのではないかと皆さん期待を持っていただいているのかなと思いますが。この部分は、皆さんにはお配りした資料と文言を変えさせていただきました。ここで紹介したモデルの多くは、実用面で若干の課題があります。性能に関しては、恐らく問題ないと思うのです。ただし、適用性の課題があります。ここで言っている適用性というのは、ある化合物群に関してはいいのだけれども、ここから外れている化合物に関しては予測性が低い。基本的に、この多くのモデルが適用性の範囲が非常に狭いモデルになってしまっているところが課題としてあるのかなと思います。それを、実際に結果をお見せできるのが、後ほどの実際の演習です。

(PP)

Read acrossについて。

(PP)

Read acrossはタームとしては皆さんには、もしかしたら新しいのかもしれないのですが、食品の分野では、私は知らなかったのですが、香料はグループで評価されていますね。

ある構造で、似た物質をグルーピングして、アセスメントするということを従来からやられていて、かなりそれに近い考え方だと御理解いただければと思います。一般化学物質の分野のほうで、Read acrossというタームが出てきたのかはわからないのですが、今はこういうようなタームを使って説明されています。

具体的に、定義上の話ですが、類似と考えられる物質のデータから、他の物質のデータを推定する方法。例えば、個々の構造を書きましたが、1個ずつ炭素が伸びていくというグループです。今、予測したい物質は、Chemical-3です。基本的に、この並びでいくと、陰性だと思いますか?やっぱり陽性だと思いますか?イメージとしては、こういうような考え方、これがRead acrossになります。

重要なのは、この構造を見て、直感的に類似性があるなと思ったと思うのですが、構造が複雑化すると何をもって類似、と言えるのか、ということです。これが一つの課題になると思います。それさえクリアできれば、Read acrossという手法はかなり広範に使える手法ではないかと思っています。

このデータのないChemical-3を予測したい。こういうデータがない物質のことをターゲット物質という言い方をして、このターゲット物質に対して、予測するためのデータを提供してくれる物質群のことを、ソース物質という言い方をします。

(PP)

もう一つ重要なこととして、内挿と外挿という考え方があります。基本的にデータがある物質で挟み込んで何かを判断するとき、ここが陽性と陽性（補足：挟み込んでいる2物質が陽性）だったらChemical-3も陽性という気がしますね。この考え方を内挿という言い方をします。

一方で、片方にしか、データがない場合です。そうすると、先程のLC₅₀のモデルのように、もしかしたらこれはひっくり返るかもしれないという可能性を考えつつも、ここも例えば陽性と判断します。こういう考え方のことを外挿という言い方をします。

一般的には、皆さん、こちら（内挿）が絶対に信頼性が高いという印象があるかと思います。この外挿を使うケースも、データがない場合にはあるのですが、内挿に比べると、予測の信頼性は落ちる傾向はあります。(PP)

Read acrossとデータベースの関係です。基本的には、ソース物質を見つける際に、何かしらのデータベースを使って、皆さん探すのではないかと。ソース物質はどういう物質かという、ターゲット物質と類似性がある、なおかつ信頼できる毒性データがある。先程、このような絵をお見せしましたが、この陽性が実は怪しい陽性では余り意味がないですね。できるだけ正しい陽性であってほしいという意味で、信頼性の高いデータがあって、構造の情報もとれるようなデータベースがいいということになります。ただし、世の中にはいろいろなデータベースがあるため、一つのデータベースだけ当たっても足りないことがあるので、いろいろなデータベースを当たって、それぞれでデータを抽出してきて、最終的にこういうマトリックスにして判断します。理想的には、こういうマトリックスをつ

くるときには、このデータベースが大きければ大きいほどいいし、毒性情報が充実しているものほどいいなというところがあって、OECD QSAR Application Toolboxが開発されてきています。

(PP)

留意点ですが、これから*in silico*を使おうというときに、やはりそれぞれの手法の限界性を知っておくことが非常に重要と思います。まず、正しい構造が必要だということ。混合物等、物質あるいは構造が特定できない場合には、単体ベースに関しては評価できるのですが、混合物としての評価はなかなか難しいところがあります。Read acrossは少しできる部分はあります。

あとは、予測できない物質があるということを知っておく必要があります。(Q) SARモデルの場合ですと、先ほど申し上げた適用範囲の関係があるので、どうしてもデータベースに依存した特性になってくるといふ点と、ほとんどのモデルで、金属だったり、無機物やポリマーというものは予測できない点、これらがまず大きな限界性として挙げられます。Read acrossに関しては、信頼できるデータを持っているソース物質がないと、予測には使えないといった側面があります。

予測精度に関しては、モデリングをするときに、何かしらの実験の結果を使ってモデリングするので、そもそも実験系のばらつきというものが予測精度に影響してきますし、(Q) SARモデルの場合ですと、その相関性を見るためのデータセットの数だったり、データベースの質というものが最終的な結果に影響を及ぼします。また、相関性を見るアルゴリズムというものもあります。

Read acrossの場合には、類似性の精度といったものというところがあります。

(PP)

いろいろ限界性があるのですが、これは動物試験も同じですね。そのような中で、皆さん、うまく使っている。思い出してください。*in silico*のいいところは、試験をしないで情報が得られるということです。今、お話ししたように、どうやって信頼性を担保するのかという考え方はある程度でき上がってきています。あとは、これをいかに運用していくかということを考え上げていけば、*in silico*というものはかなり有用なツールとして使っていけると考えます。

毒性予測法で必ず知っておいてほしいことは、唯一絶対という手法はない。どのようなモデルだったら絶対ですかという質問をよく受けるのですが、そのようなものはありません。まず、断言しておきます。これは、一人の専門家が絶対ではないのと同じなのです。皆さんが委員会をやっていろいろな意見を聞くということは、いろいろな考え方があるということを取り入れたいからだと思う。それはモデルに関しても同じだということです。

また、皆さんが必ず専門家を集めて会議をするのと同じで、複数の手法を合わせることで、それぞれの手法の弱点をカバーできます。これはぜひ覚えておいていただきたいこと

かと思えます。

(PP)

何で組み合わせるとよいのかということなのですが、そもそも(Q) SARモデルは予測できる物質が違っている。市販されているソフトは特にそうなのですが、使っているトレーニングデータセットが違うのです。そうすると、さきほど描いたようなスペースがみんな違うのです。それらのスペースの範囲をできるだけ使うことで、知らない範囲というものをできるだけなくしていくということが一つあります。

完璧な解析方法もないと申し上げましたが、異なる解析アプローチを使って、見落としを防いでいこうという考え方です。ナレッジベースであれば、例えば既知見は全て取り込んでいる。ただ、統計ベースは、既知見でもわかっていないことについても取り込んでいる。それを組み合わせることによって、メソッドの弱点を補っていくという考え方で、そういう組み合わせがいいという考え方になります。

(PP)

市販ソフトウェアとの付き合い方なのですが、よくある問題点としては、市販ソフトやモデル構築に使用した情報を全部は開示してくれなくて、ブラックボックスではないかと。この点が、行政利用する際になかなか難しいということがあるかと思えます。

なぜ開示してくれないのかというと、非公表を前提に企業等からデータをもらって、それをトレーニングデータセットに組み込んで、モデル構築をしているケースがあるからです。

一方、トレーニングデータセットを見せてくれるモデルもあります。

それから、全部は見せてはくれないのですが、一部を見せてくれるモデルもあります。それをいかに有用に活用していくのかということ、我々が考えて取り入れていけばいいのかなということがあるかと思えます。

更新されることで結果が変わるということもありますが、更新しなければいいという考え方もあります。後ほど説明する多くの市販ソフトウェアは、アルゴリズムはある程度皆さん、知っている、あるいは勉強したらわかることです。そして、そのアルゴリズムがある上に、自分たちのデータを乗っけて新しくモデルをつくることができます。要するに、アルゴリズムがわかっていて、トレーニングデータセットを自分たちが知っているもので作り上げたら、基本的にはトランスパレントな(補足:透明性がある)モデルがつくれるという考え方になり、それも一つの選択肢かと思えます。だから、この更新されることで結果が変わることがあるということに関しては、更新しないという選択肢と、自分たちでモデルをつくり、そのモデルに関しては、自分たちが管理する上においては変更されないといった利点があるのではないかと思えます。

(PP)

それから、これからデータベースをつくっていくのであれば、どういうデータベースがあるとよいかということ、モデル構築やRead acrossを行うためにはどういったこと

をしたらいいかということを書いています。

(PP)

これからどういったことを考えていけばいいかということについて、私見なのですが、
どういった目的で (Q) SARやRead acrossを使用したいのかということを確認した上で、
食安委が評価したいケミカルスペースはどこにあるのかということ特定し、どのツール、
市販ツールが使えるのか使えないのか、そして、オリジナル (Q) SARツールを構築する必
要があるのかないのかということについて考えた上で、オリジナル (Q) SARを構築するの
か。あるいは、Read acrossをするのだったら、その基礎となるデータベースをつくって
いく必要があるのではないのかなと考えています。

(PP)

データベースをつくる際には、何をしようとするのかということとはあらかじめ考えてお
かないと、後半になって大きな変更をすることはなかなか難しくなるので、これが最初と
いうことになります。

まず、これを明確にした上で、データベースの仕様を検討します。その仕様が決まっ
たら、粛々とデータを入れていくだけになるのですが、まずは優先度の高いコンテンツから
でいいから、とりあえず入れていく。物質当たりの量を増やすよりは、まず物質の数を増
やしていきましょう。最悪、構造だけでもいいですから、まずは入れてくださいと私は個
人的には思っているところです。

以上になります。

〇〇〇 どうもありがとうございました。

非常に根源的なところから説明していただいたので、御存知の方は復習できたと思いま
すし、初めて納得、という方もいらっしゃると思います。共通の理解に立った上で議論し
なくてはいけないので、私自身も大変勉強になりました。これから、議論の素地、土台が
できていくと思います。

まず、委員の先生方を中心に質問あるいはコメントをお出しただけであればと思いますが、
いかがでしょうか。

お願いします。

〇〇〇 わかりやすい説明、ありがとうございました。

7ページで証拠の重みづけ (Weight of Evidence) ということの説明していただいたの
ですが、ヒトのものだと、我々だとサンプルサイズが大きいか少ないかみたいところで
重みづけと、よく我々はやりがちなのですけれども、これは例えばその対象の動物が違
うとか、あるいは同じような実験がいっぱいあるのかとか、そういうところでやるのか
とか、その辺の重みづけのやり方はある程度決められたものがあるのか教えていただ
ければと思います。

〇〇〇 こちらの分野でのWeight of Evidenceだと、いろいろな証拠を積み重ねてとい
う意味合いになります。いろいろな証拠の中のひとつで、例えば発がん性をやりたいとい

きに、Amesの結果ではイクイボーカルだった、でも、染色体だったら何とかだった、*in vivo*の小核の結果だったら何とかだった、でも、いまいち決定打に欠けるといふときに、よくでき上がったモデルの中には、メカニズムに関してのインタープリテーションがすごくしっかりしているものがある。それで、例えばなぜ陽性なのかという、この官能基というものが、DNAとかなり強固に反応するのだということがあって、そうすると、ここのアンビギュアス（補足：曖昧）だった結果とか、染色体で、*in vitro*でポジだった、小核でと、この要素をかみ合わせると、多分まだ足りないと言われるかもしれないですが、だんだん証拠が重くなってきますね。それで最終的に発がん性はあると判断しようかというような考えがweight of evidenceであり、QSARやread acrossの結果はその判断材料の一つという考え方です。

〇〇〇 その重みづけの計算というのもある程度確立されているのですか。

〇〇〇 重みづけの重さのウェイトは、今のところ、そういうフローをつくろうという動きはあるのですが、これには係数の何かを掛けてと、そういうものではないのです。そこはもっとエキスパート判断的なことで今、運用されています。

〇〇〇 ありがとうございます。

他はいかがでしょうか。

先ほど、先生から（Q）SARといいますか、類似性からRead acrossがそういうところはあるのでしょうかけれども、類似性から攻めていくということでありましたけれども、形が似ていれば性格も似ているのだろうということなのですが、似て非なるという言葉もあるので、そういうことがどの程度起こり得るものなのか、事例などもありましたら教えていただきたいと思います。

〇〇〇 それはよく質問されることではあるのです。似て非なる、よくあるのです。基本的に似て非なるものに関して、エキスパートジャッジのシステムでは、似て非なるということを知っている、知っているということがあるので、その情報をシステムの中に取り入れてくれています。例えばメタノールとエタノールは全然毒性の出方が違いますね、ということはよく知られていることなので、そういった知見に関しては、既にインテグレートされている。ただ、まだ誰も知らない、未知の新しい「似て非なる」に関しては、知らないから知らないということが現状なのです。それが一個、こういうモデルのウイークポイントではあるのですが、それは誰もわからないというところでは同じなのかなと。

〇〇〇 それと、先ほど33枚目のスライドで、テストデータを用いた検証で、だんだん右へ行くほど下がってくるという、かなり直線に乗った関数を示されたのだけれども、その外に行くと、いきなり高いところに行くので、これは実データなのかどうかはよくわからないのですが。

〇〇〇 これは実データです。これは上に上がっているのは、本当は毒性はなかったのです。LC₅₀を求めようと実験を行ったが、溶解度まで最大やって、どろどろとまでは言わないですけども、溶けるだけ溶かして、その中に魚を暴露させたが、死ななかったのです。

そういう結果です。ガイドライン上だと最大濃度の考え方は決まっていますが、このプロットでいくと、このパターンで、疎水性が上がれば上がるほど毒性は強くなっていきます。ここだと0.01mM/Lだというレベルだと毒性は強くなっていくはずなのですが、疎水性とともに、水にも溶けにくくなっていく。ある段階で魚の中に入っていかなくなる、取り込めなくなるというところで、毒性がなくなってしまう。これはブレイクポイントというのですが、そういうものが存在しているのです。ですから、この事例だと、ここから外れたものは、今、言った説明で、上がるというよりも毒性がなくなってしまった。要するに、本来はプロットできないです。

〇〇〇 そういうものを普通は連続的と考えがちなのだけれども、この場合、スレッシュホールドのようなものがあって、ブレイクポイントでもいいですが、コンティニユアスなのか、スレッシュホールドライクなのかということもデータには入れられる。それも、知見があれば入る。

〇〇〇 そうです。その知見を得ようと思ったら、このデータが必要になってくるので、ここよりも外れたところがどうなのか。データがあれば、多分適用範囲もここからここまで広がるのです。今回、モデルをつくる時に、ここまでのデータでつくったから、ブレイクポイントがあると知らないの、知らなくても別にいいですが、もともとモデリングというのは、ここからこの範囲の中でどうかということが言える。ここからここまでのデータを使ったら、ここからその中の話を入れるというものです。

〇〇〇 今は、疎水性が強いと体にそもそも入らないという、多分、ドーズレスポンスとは別のメカニズムがもう一本存在する。つまり、デュアルに説明しないと行けない。そういう複合理論ももちろん可能と考えていいですか。

〇〇〇 そうです。

〇〇〇 いかがでしょうか。

どうぞ。

〇〇〇 QSARの予測が外れる場合には、どうして外れるのかという、そういう知見というか、ナレッジを整理して引き出せるようにしていけば、(Q) SARも使いやすくなっていくのかなとは思っています。

〇〇〇 ありがとうございます。

知識の総体ということなので、複数の理論を組み合わせてたりすることはもちろん、可能になるでしょうしということですね。

委員の先生方、事務局からも御発言していただいて結構です。

どうぞ。

〇〇〇 そのスライドの少し前の30枚目のスライドから始まった一つのケースの説明で、結果的に、最初、その30枚目だと化学物質でいうとある化合物群を題材にしていて、そこから得られたLC₅₀と疎水性の指標をもとに、その後、スライドが展開しているのですけれども、その次の31ページに行くと、私の頭が固いからなのでしょうけれども、縦軸と横軸

で、いきなり化合物の側鎖の長さというところが縦横の軸にならないで、横が疎水性で縦がLC₅₀になっているので、これは一見、構造活性相関というよりは、活性-活性相関に見えるのですが、当然間にそれぞれ側鎖の長さがあるって、それぞれの2つの指標が長さでリンクしているから初めてこの縦横の31ページのこの図になっているという理解でいいのでしょうか。

〇〇〇 すばらしい質問だと思います。

今、(Q) SARの定義、そもそもは構造活性相関なのですが、実際に運用されているのは、活性-活性相関に近いところがあります。本当に構造ベースで、例えば後で説明する〇〇〇だったり、〇〇〇だったり、〇〇〇といったものに関しては、いわゆる、本当に構造のフラグメントをとってきて式をつくっていますけれども、生態毒性の分野では、こういうようないきなりlog Powのモデル、log Powを使ってモデリングするものを(Q) SARと言っています。すごく厳密に言ったら、AAR、QAAR ((quantitative) activity-activity relationship) という言い方になります。今の内容を敢えて説明するのだったら、このlogPowというのはおっしゃった通りで、その疎水性は炭素の長さに応じて変化していくものだからという言い方はあります。本当にlog Powから入ってきてしまっても構わない。でも、実際、これは炭素をプロットしても、そこそこの相関は出ると思うのですが、ただ、実際に予測するときにはCの数よりもlog Powにしたほうが、精度は上がってくるのです。ですから、実際にモデルをつくる時には、こちらのパラメーターを使っているということでございます。

〇〇〇 よろしいですか。腑に落ちていない。

〇〇〇 多分、実際にこのモデルを使って評価というか、予測をするときには、当然、目の前にあるターゲットはコンパウンド、物質なのですね。そういうときは、この場合であれば側鎖の長さで一旦置きかえて見てという解析の仕方になるのでしょうか。

〇〇〇 多分、目の前にあるときには、その物質のlog Powの実測値の値を入れる。あるいは、目の前の物質のlog Powの実測がなかったら、log Powに関しては、計算値が簡単に出るようになっているのです。かなり、予測の精度も高いので。物質から予測のlog Powを出して、それで、その値をモデルに入れる。モデルを2回介するようなイメージになって予測されています。実際にそういうモデルはすごく広く使われています。

〇〇〇 いかがでしょうか。

どうぞ。

〇〇〇 今のお話で、記述子というか、説明変数は山ほど普通、ソフトを持っていて、一番いいのを選んでいっているのです。場合によっては、Hammettの σ とか、そのようなケースもあり得るでしょうし、log Powはもちろん、一番に入ってきますし、かえってアルキルの1、2、3、4を同定するほうが難しいかもわからないですね。人間が見たら当たり前ですが、ここでトレーニングデータセットを選んだのは、〇〇〇の思い込みなのか、明らかな説明があれば、みんな同意するでしょうけれども。

〇〇〇 トレーニングデータセットはまず、データがあるというところからスタートする。データセットをつくる時にはまずデータがないと。モデリングするときには実測値のデータがなければいけないので、今回、私はこのデータをつくり上げたわけではなくて、あるデータベースに行って、一番簡単にモデルをつくれるのはこういったシンプルな構造なので、このデータを引っ張ってきただけ。そうしたら、こうなった。それにlog Powの値をつけてあげて、トレーニングデータセット完成、のような感じです。

〇〇〇 多分、予測モデルをつくることというのは、基本的に素性は問わないというか、要は一番いい予測の結果になればいいということで、多分これに限らないですけども、私たちがやっている予測モデルも、別により根源的なものでなくても、とにかく現象を一番うまく説明するものであれば、論理的でなくてもいいと言うと極端だけども。

〇〇〇 そこはちょっと危険です。

〇〇〇 ちょっと危険なのですね。論理的ではなくていいというか、そういうことは余り気にせずに、一番効率よく予測するものを探すので、結果的には、もちろん根っこのところにつながりはあるのだけども、現象的に一番性能がいいものが最後に生き残るということだと思ふのです。だから、潜在的にそういうよりファンダメンタルな因子が入ってくるのだけども、表現されるのは一番結果がよくなるものだけが生き残ってくるということだと思ふます。

〇〇〇 この事例では、どの変数がいいかという解析を飛ばしたのです。でも、実はそれはすごく大事なのです。この後の(Q) SARモデルを構築するときのモデル作成というときに、log Powが本当にいいのか。変数はいっぱいあるのです。さっきの構造をばらばらにしてあげて、例えばモレキュラーウエートにしてあげてもいいし、バルキネス、体積で表現してあげてもいいしと、いろいろなパラメーターをつくって、それとLC₅₀との関係を見ていく。そうすると、どれが一番よくLC₅₀を説明してくれる変数なのかなと見出していったときには、やっぱりlog Powが一番いいかもしれない。今回は、毒性そのものが比較的わかりやすい魚の話だったので、これに関しては入っていきやすさということをよく説明されるので、理由も立つし、統計学的にもlog Powが一番よかったと。だからlog Powがいいよというのが、この事例の、最初ここを全部すっ飛ばしていったので、コンフューズさせてしまったのかと思ふます。

〇〇〇 ありがとうございます。

他はいかがでしょうか。

では、もう一つ伺います。44枚目のスライドに、〇〇〇の頭の例が出て説明されたこれなのですけれども、〇〇〇以外の方がいい説明をしていて、しかし、〇〇〇とその先生がずれていて、重なりがあるところにずれがある場合に、それを補正する仕組みを持っているのか、それとも、そのまま受け入れているのか、どうなのでしょう。

〇〇〇 基本的には、補正する仕組みのところはエキスパートジャッジメントになる要素になる。〇〇〇がおっしゃる通りです、〇〇〇がおっしゃることもその通りだと思ふます。

ですが、ちょっと引いて考えます。ちょっと他のデータベースに当たってみて、自分が今、予測したいという物質と関連している物質の実験データがどうなっているのかなということ眺めてみるのです。モデルの中ではこうやってきたのだけれども、実際のデータを見てみると、この物質に関しては、このモデルをまだ知らなかったのだろうと。だから、ここで例えばポジティブだと言われたとしても、この結果から見たらネガティブと言えると言って、結果をひっくり返すということがあります。ただ、それはシステムにはなっていません。これは、ここの部分がいわゆるマニュアルでやらなければいけないエキスパートジャッジの部分になっていきます。難しい部分もあるのですけれども、(Q) SARの結果に関する解釈だけと限定するとそんなに難しくないということを、後ほどの演習でお見せできればと思います。

〇〇〇 ありがとうございます。

どうぞ。

〇〇〇 すごくスペシフィックな話で、今の補正システムであるかという先生の御質問で、一般的には受け入れられていないけれども、試みはあります。それはなぜかという、それぞれの(Q) SARの信頼性と強度のようなものを変数に入れて組み合わせたときに、ウエイトをかけるのです。そうすると、この(Q) SARのほうが少し信頼性が高いので、そっちを重く置いて判定するというシステムもあります。

〇〇〇 ありがとうございます。

では、基礎的な説明はここで一旦仮に終わらせていただきまして、〇〇から、次の課題についてお話しいただきます。

予測精度の現状ということで、〇〇から御説明いただきます。資料2です。

〇〇〇 〇〇〇です。

(PP)

最後に組み合わせなりどういう使い方をするのかという観点から、少し行った事例を紹介させていただきたいと考えています。

(PP)

そうは言っても、今日のお話は、実はこれはもう10年前に国際学会で発表したスライドをそのまま持ってきて、今は、どのくらい進んだのかなと聞かれると、恥ずかしい話にはなるのですけれども、そのときは、使う方向としては、一般的な化学物質の評価です。

(PP)

目的としては世界中に何十万物質も存在すると想定される化学物質を評価しなければいけない。この頃から、実はWSSDの2020年目標というものがあって、2020年までに身近な化学物質をある程度、危険性や他の評価が終わらなければいけない目標のようなものがあるわけです。それに向かって、本当はもう10年前からこういう(Q) SARを使って、効率的にたくさんの物質を評価していこう。あるいは、優先順位をつけて試験をしていこうという観点から、こういう研究をやり始めました。そのときに、(Q) SARを使えばすごく効率化で

きるのではないかということの一つのツールとして取り上げたところです。ただ、この後説明しますけれども、そのときから〇〇〇とか〇〇〇とか〇〇〇といったソフトが既にあったので、使うのは、Ames試験は一般的に使われていたので、そこからやっ払いこうと。ただ、実際には、それほどいい予測精度、今でこそ90%ぐらいですが、その当時は、実は、この後紹介しますけれども、それほどいい精度ではなかったということがあります。ただ、組み合わせれば何とかなるのではないかという発想で始めました。

(PP)

組み合わせるのは先ほど〇〇〇から紹介があったように、いろいろなモデルを組み合わせている。同じモデルを組み合わせてもしょうがないので。そういう意味で取り上げたのが、3種のモデルです。

(PP)

これらを検証していくと同時に、開発もしていこうという研究のプロジェクトも始めていて、もう一つ、大事なことは、そのときにもデータの信頼性が大事だということで、非公表のGLPで実施された試験で検証と評価をやっていくことにしました。

(PP)

それで、これに対して、まず〇〇〇をやってみたわけです。実は、今でこそ多分〇%とか出るのでけれども、この当時は〇%ぐらいの一致度で、この感度、特異度の説明は、先ほどの〇〇〇のところを見ていただければいいのですけれども、陽性物質を陽性として判定する、陽性物質なのに〇個は外すという意味で、これが出ています。だから、〇個の陽性物質があったのに、〇個は陽性と判定できなかった。ですから、感度が〇%という、そういった予測性の指標です。これは陰性を陰性として扱う。もとのデータセットは陰性が多いので、どうしても陰性のスペシフィックが高くなるのですけれども、こういった状態でありました。この理由は、我々が用いた物質と言うのは、それまでに世の中で試験されていない物質を積極的というか、世の中でデータがなかったものを選んでつくっています。実は、〇〇〇も〇〇〇も、アメリカとかそれまでの文献情報、全部の知識を総合してつくったデータベースですけれども、そこから見ると、実はこれは外部データになります。外部データだったので、要するに、そういうものの知識がない、多分、ナレッジなり、そういう構造が、我々が試験した中には入っていたのだと思います。だから、その部分が抜け落ちて、こういう物質を外してしまうといったことにこのときはなったのです。

(PP)

こういうことをやっ払い、〇〇〇も同じように、そこはちょっと〇%と高いのですけれども、今はちょっと変わったのですけれども、アプリカビリティはケミカルスペースのドメインが、その当時は〇〇〇は小さくて、金属がある、塩があるとうまくいかないとか、そういったソフトウェア上の制限があったので、100%予測できなかったのです、そういった状態になっていました。

(PP)

〇〇〇、これは本当に純粹に〇〇ベースでやるソフトで、それはちょっと、その理由、完全に〇〇ベースだけでやるとどうしても悪い感度予測、感度はそんなに変わらないのですけれども、そのような特異度が少し下がってしまうといった傾向がある。

(PP)

では、この3つをどうやって使うのかということ考えたときに、いろいろな組み合わせがあるのですけれども、我々は2つほど提案してみました。

一つは、3つあるのだけれども、2つポジティブと言え、それをポジティブと判定しよう。2つネガティブと言え、それをネガティブと判定しようといった組み合わせで、ポジティブ、ネガティブを判定するといったやり方にしてはどうか。

もう一つは、全部ポジティブ、全部3つがネガティブといったやり方でやれば、要するに、逆に実はこの部分は判定できないものとして残ってしまうのですけれども、こここここの信頼性は上がるのではないかと考えた考え方で組み合わせ法を使ってみようといったことを考えました。

最初の一つのやり方は、2つが陰性、2が陽性。これは実際に戻って表を見ていただくといひのですけれども、〇〇〇とほとんど変わらない値になってしまったのです。これは結果論であって、実際はこれは何回かやり直していくところになってしまったので、最初にやったときには、この一致度が一番高かったのです。だから、こういうやり方をすると、少し一致度が上がるといった傾向は、数字として見せられませんでしたけれども、初期の結果にはありました。ただ、これはアプリカビリティが〇〇〇よりふえているので、要するに、〇〇〇が予測できないところまで予測しても、このくらい的一致度をつくることのできる利点があります。ただ、そんなに感度は実は上がらなかったということがあります。

(PP)

もう一つのやり方にすると、3つとも陰性、3つとも陽性というものを陰性、陽性と選んでいくと、感度とか特異度は、かなり上がるのがわかりました。これは当然といえば当然で、どのソフトも陽性と判定する物質は、陽性である予測確率を高められるということになります。ただ、そのかわり適用範囲は50%、要するに50%は判定不能といった分類にしてしまうので、この辺は使い方なのです。より信頼性の高い陽性を見つけていこうというやり方にすればこういうようになりますけれども、判定できない物質を増やしてしまうといった利点と欠点が出てくるということがあります。

(PP)

これを化学物質の評価の中でやっていこうといったときに、ここが *in silico* の話とちょっとずれるのですけれども、大きくは同じなのですが、先ほどの吸収性の問題もそうなのですけれども、分子量が大きくなると、実は毒性が出てこないということは Ames でも同じでして、分子量が3,000以上になると当然陽性になりません。

(PP)

化学物質の申請の中で、こういった高分子のものも出てきますので、そういったものも、あらかじめ (Q) SARをかける前に、ケミカルスペースのドメインという考え方からも取り除こうということでありました。分子量3,000以上だと、96.4%のうち、わずか実は陽性を示したのは7物質であった。このとき、実は7物質のうち4物質は実はポリマーの中に、こういうエポキシイドの基を持っている。これは正式な規則ではないのですが、今も化審法では、ポリマーでもエポキシイドの基を10%以上持っている場合は、スクリーニング毒性試験をやるといったスキームになっています。

(PP)

こういうものを全部組み合わせて判断樹をつくってみると、例えば、ある物質が来たときに、分子量が大きいかどうか。大きい場合には、エポキシイドがあるかどうかで陰性、陽性、これはかなり乱暴な組み合わせですけれども、こういった判定をしてしまって、3,000以下であれば、先ほど言った3つのシステムで判定して、2つ以上ポジティブならポジティブ、ネガティブならネガティブと、こういった大ききざっくり、感度は予測性が少し悪いですが、優先物質の優先づけぐらには使える手法になるかと。

(PP)

あるいは、さっき言った、3つともポジティブ、3つともネガティブというもので陰性、陽性を判定して、例えば、これも優先順位を使うのですが、こちらの試験は、実際には、仮にあった物質については実際の試験をやるように判断するといった優先順位づけにも今度は逆に使うことができるのではないかと考えて、こういった提案をしています。これはまだ提案だけでこれが実際に化審法の判断で運用されているというわけではありません。

(PP)

最後に、複数のモデルによると、高い一致度を得られるモデルをつくることはできたのですが、ただ、3つとも予測が一致する場合ということは、高い予測率を達成できますけれども、逆に判定できる物質は減る。あるいは2つ以上のモデルでやるという言い方ですが、予測率は高くはない。

今日はお示ししませんでしたでしたが、例えばスクリーニングの陽性だけを判定するといったシステムで使おうとすると、例えば、どれか1個でも陽性だったら陽性としてしまおうとくくってしまうと、かなりの部分を陽性として判定します。これは安全性の観点から考えると、3つとも陰性のものはとりあえず後回しで、こちらの物質から優先的にやっつけようといった使い方、要するに、一つでも陽性なら陽性といった判定の使い方も、(Q) SARの使い方としては行政判断としても使えるのではないかという考え方で、こういったことが組み合わせのやり方の応用として提案させていただいて、こういう事例もあるといったことで紹介させていただきました。

以上です。

〇〇〇 ありがとうございます。

もう10年前から、こういうやり方が既に検証されていたということではありますが、コメントはいかがでしょうか。

どうぞ。

〇〇〇 非常に素人的な質問で申しわけないのですが、これはスクリーニングをやるときに、感度を優先させるのか、特異度を優先させるのか、はたまた、この場合だと、ネガティブプレディクティブバリューが大事なのか。応用によって何を優先するのかは変わってくると思うのですが、今、こういう(Q) SARなどでこういうように使われるときのことを考えると、どこの数値を優先するのがいいのでしょうか。ネガティブなのが大切であればネガティブプレディクティブバリューですし、あるいは見逃してはいけないから、感度をとにかく上げるほうを優先させるのか。その辺によって、変わると思うのです。

〇〇〇 どっちが大事かというのは、運用するときに、どっちを選ぶのかで決まる話です。多分、厚生労働省は安全側で、なるべくスクリーニングで多くのものを判定したいということは、やっぱり感度を上げる。だから、要するにオールポジはふえても、オールネガは出さないといったものが、安全側の観点からはですね。ただ、評価の効率化を考えると、どんどん外したいということを考えると、それはコンフリクトした話になりますけれども、陰性と。

ただ、先ほど言いましたように、(Q) SARそのものは大体ポジティブのルールばかりでできているので、ネガティブのルールはないので、その辺が難しいですね。ネガティブのルールがないということで、3つともネガティブになったけれども、本当にネガティブかという話が問題としては残っていると思います。

〇〇〇 それはそうですね。ゴールドスタンダード自体が、ネガティブの証明ができるのかという問題もあります。〇〇〇、どうぞ。

〇〇〇 ちょっとだけ補足ですが、最近すごく(Q) SARの開発が進んできて、ネガティブな予測ができるようなもの、すごく最近なのですが、できてきました。〇〇〇も若干ネガティブに強くなってきましたと。

〇〇〇 では、それが10年の進歩なのですね。

〇〇〇 ケミカルスペースはやはり範囲が重要で、我々が始めたころは、〇〇〇、〇〇〇でさえ、多分、我々の新しく出てきたデータがもうアウトオブケミカルスペースだったわけです。よくなかったわけです。それは、我々のデータは実は提供しているので、そのせいもあって、今、上がってきています。

〇〇〇 ありがとうございます。

感度、特異度のところで、一つ気をつけなければいけないのは、表のツー・バイ・ツーのテーブルの描き方は、縦横入れかわる場合があって、というより、我々の世界はと言ったほうがいいのか、縦横が違うことはあるので、どっち向きに書いているかということ認識して、どっちがゴールドスタンダードのマーカーで、どちらがテスト手法かによって、2掛ける2の縦横が反対になるということが一つあります。

〇〇〇 そうですね。多分、それは〇〇〇のスライドの中でも。

〇〇〇 それは習慣の問題なのだろうとは思いますが。

〇〇〇 陽性予測をした物質の中で、どれだけ陽性物質が入っているのかに、縦横を変えらるとなります。

〇〇〇 少なくとも、先生が書いている部分に関しては、私のほうのスライドでいくと、資料の25ページ目のところに、Cooper statisticsのものを書いているのですが、これに応じたもの。

〇〇〇 違うと思ったのだけれども。

〇〇〇 今、多分、〇〇〇がおっしゃっているのは、左側のAmesのところを*in silico*と書いて、Amesを上を持ってくる解析の仕方を。

〇〇〇 描き方だけの問題で、両方あるので、固定した地理で覚えてしまうと、反対になっている表が読めなくなりますという注意が一つと、もう一つは、対象物質は、これは3つの手法を、原則同じもので確認しているから、その中でコンフリクトは起きないので、数字そのものはどういうサンプルを持ってくるのかによってすごく動きますので、この場合、例えば誰が見ても陽性に出るといふものと、誰が見ても陽性には出ないという異なるスペクトラムのものを合算してサンプルをつくと、感度と特異度がそれぞれ100%というものも論理的につくることができるものですから、数字を余り絶対視しないほうがよい。この感度、特異度の表現の一般的注意としてありますが、今回は原則同じテスト物質を3つの手法でやっているのだから、その間での比較性はもちろん、あると考えます。〇〇〇、どうぞ。

〇〇〇 本来、この検証をやるときには、使った114物質というのが、〇〇〇だったら〇〇〇の中、さっきの適用範囲の中に入っている物質が何物質あって、適用範囲の中に入っているもので予測精度はと見るのが本来あるべき外部検証なのですが、多分、これはアプローチとしては、このデータセットに対してどれくらい予測精度があるのかという考え方で、少し、そこは考え方が違うということは認識しておく必要があるかと思いました。

〇〇〇 よろしいでしょうか。

では、〇〇〇、どうもありがとうございました。

頭がかなり集中したので、休憩を5分ぐらいとりたいと思います。

その後、また〇〇〇をお願いします。

(休 憩)

〇〇〇 では、遅れぎみになってしまいましたけれども、続いて、〇〇〇に実際のところを御説明いただきたいと思います。

よろしく申し上げます。

〇〇〇 それでは、引き続きやらせていただきます。

(PP)

今日は、今からやる「解析の実際」で御紹介をするソフトウェアをまず紹介した上で、個別の(Q) SARソフトでの予測結果の解釈の演習、3題か4題ぐらいを用意してきたので、皆さんに実際に回答していただきたいと思います。ランダムに「はい」と当てますので、気を抜かないで聞いていただきたいと思います。

最後に、食品安全委員会で2012年までに評価した物質の中で、実際に予測モデルに入れられる物質を選んで、面白いかどうかあれなのですけれども、結果が出せるものについて、今日、持ってきたものをお見せします。

最後、Read acrossに関して、先ほどどのようなものかイメージはわかったと思うのですが、どうやる、がなかったと思うので、1例だけですけれども、やったものを今日はお見せしたいと思います。

(PP)

紹介するソフトウェアなのですが、(Q) SARソフトに関しては〇〇〇、先ほど〇〇〇が御説明した〇〇〇のものになります。

また〇〇〇や〇〇〇という統計ベースのモデルを御紹介します。

もう一つ、Read acrossに関しては、これはほぼゴールドスタンダードに近いソフトウェアになっていると思うのですが、OECD QSAR Application Toolbox、今は、Applicationはなくなったのかな。これについて御紹介したいと思います。(Q) SARに関しては、これがよいというわけではないです。

(PP)

まず、〇〇〇です。〇〇〇はすごく古くから使われているソフトで、〇〇〇ということで、私は今回、この発表をするに当たって初めて調べて、〇〇〇ってこんな名前だったのだと知りました。ありがとうございます。

〇〇〇なのですが、基本的には部分構造です。

〇〇〇はトレーニングデータセットに関してはごく一部しか公開しないというところが一つ。毒性とどう関連性があるのかということの説明してくれるという意味合いでは、すごくわかりやすいモデルです。ここをどう使っていくのかということが一つ重要な考え方になるかなと思います。さっきブレイクポイントの話がありましたが、いきなり毒性が変わってしまう、あのような情報に関しては、そういうことがあります、という情報が得られます。エキスパートジャッジメントをする上での強力なサポート情報を提供してくれる。ナレッジをデータベースにしてくれているというイメージになるかと思います。

また、モデルの改良等が可能なため、もし食品安全委員会で見出した知見があれば、それを組み込むことも可能です。

(PP)

エンドポイントに関しては、基本的にヒューマンエンドポイントで、ヒューマンヘルス

に関して見たいエンドポイントは、基本的には網羅しています。

字が小さいので、皆さん、お手元の資料を見てください。陰性予測はできるようになってきた。

もう一つ、〇〇〇、これは変異原性のモデルなのですが、〇〇〇がデータベースを提供しながらつくられた、統計ベースの (Q) SARになっています。この中では、トレーニングデータセットは全て公表データを使っているのもので、トレーニングデータはトランスパレントなもので、皆さんも見るできるようになっています。

結果に関しては、本当にこの予測の結果が正しいかということの判断材料を提供してくれ、非常にわかりやすいツールになっています。

この〇〇〇に関しても、一番下に書いていますが、自分のデータセットを入れれば、市販されている〇〇〇モデルと全く同じアルゴリズムでモデルを構築することができます。

(PP)

〇〇〇、これは〇〇〇のソフトで、統計ベースの (Q) SARモデルになっています。

このモデルもトレーニングデータセットをみることができるので、エキスパートジャッジをするときに、非常に強力なツールになってきています。

(PP)

〇〇〇の中では、予測できない物質、さっき言った適用範囲外の物質に関してアウトオブドメインを教えてくれるので、この物質は予測できないのだということがすぐわかる。予測結果の信頼性を考える上で、非常にいい機能かと思います。

予測の信頼性をサポートする機能として「〇〇〇」といって、トレーニングデータセットの中で、予測したい物質に類似している物質を見せてくれます。そうすると、本当に当たったのかというセカンドオピニオンをつくり出すことができるということがあります。

(PP)

続いて、〇〇〇です。

これは、〇〇〇社製の統計ベースの (Q) SARです。〇〇〇のいいところは、定量予測もしてくれるということです。例えばLOAELを出してくれるとか、LD₅₀値を出してくれるとかという意味では、いいモデルかと思います。

あともう一つ、いいところは、適用範囲内の予測かどうかを示してくれます。また、構造のフラグメントもこの〇〇〇のモデルをつくる時に知っている、見たことのあるフラグメントなのか、新しいフラグメントなのかということについての知見をくれるので、こちらのモデルに関しても適用範囲内かどうかということを見ることができます。

また、皆さんが持っているデータを入れたら、カスタマイズしたモデルをつくることはできます。モデルのアルゴリズムに関しては、皆さんが知っているようなものなので、アルゴリズムはトランスパレントです。皆さんのデータベースさえトランスパレントであれば、基本的にはトランスパレントなモデルを構築することができますというコンセプトになります。

(PP)

では、個別の (Q) SARソフトでの予測結果の演習をやってみたいと思います。まず、簡単なものから。

これは、物質Tという物質です。〇〇〇で予測した出力情報を読んで、結果をどう解釈するのかというところが、要するに、ナレッジを使うエキスパート判断というところになります。皆さんにこれを読んでいただきまして、この物質の判定をしていただきたいと思えます。2分ほどでよろしいでしょうか。ちょっと読んでいただいて、お願いできればと思います。つまらないと思わないで、頑張ってください。

あと30秒ほどで。誰か当てますね。

皆さん、本当に真剣に取り組んでいただきまして、ありがとうございます。では、〇〇、お願いします。

〇〇〇 化学と遺伝毒性は苦手なのですけれども、3行目に「〇〇〇」と書いてあるということと、下から3行目で「〇〇〇」ということで、私の解釈は、〇〇〇があったから、陽性と出てしまうのだけれども、化学物質そのものの結果ではないという判断で、ネガティブと考えました。

〇〇〇 完璧です。これは物すごく簡単な事例だったと思うのです。要するに、読んだらちゃんと書いてあるのです。これがエキスパートシステムのいいところなのです。要するに、こういう既知のナレッジをこうやってキュレートしておいてくれる。わざわざ文献調査をして、ここまで調べるのではなかなかたどり着かないと思うのです。でも、こういったナレッジをあらかじめ入れておいて、答えがそのまま書いてあるのです。これをやるのはどうかなと思ったのですけれども、ナレッジベースのよさをまず体感していただきたいということで、さすがです。〇〇〇。

ということで、これは〇〇〇の場合に陽性になりますということで、これに関しては、レファレンスもちゃんと原著のアーカイブまでの情報が全部入っているのです。だから、食品安全委員会で評価するときには、ソフトウェアが言っていることなんかと言わずに、もう一回文献までたどっていけるので、そういう意味では、すごくユーティリティーが高いのではないかと思います。

〇〇〇、ありがとうございます。

(PP)

次です。次は、物質Iという物質です。これについて、Amesの予測をしました。〇〇〇を使ったときには、ポジティブという結果になりました。統計ベースの〇〇〇というソフトを使ってみたら、ネガティブという結果が出ましたという事例です。〇〇〇の結果と、この〇〇〇という結果を見て、総合的にこの物質IのAmesに関して、どういう予測の結論にしますかということが演習問題になります。

ちょっと、3分くらいでいってみましょう。

では、回答に入りたいと思います。今回、これを演習するに当たって、本来は予測ツールの中の情報はもっとリッチなのです。ただ演習をするに当たって、それを全部載っけたら、載らないのです。ということで、ちょっと簡単な情報だけにさせていただきました。

まず、この〇〇〇の結果に関してなのですが、ここで何を見たいのかというと、要するに、既知見で、この物質と類似した物質は、実際にどういう結果が出ているのかというのは、当然見たいわけです。少なくとも、この〇〇〇の物質、本来はもっといっぱいあるので、もう少し全体的に見なければいけないのですが、この例えば4物質に関して。この物質は除きます。少なくとも、1番から3番までの物質というのは、基本的に〇〇〇ですね。〇〇〇が1つで、全部〇〇〇のものに関しては、基本的に全部ネガティブな結果が出ています。こっち、ポジティブになっているものに関しては、要するに〇〇〇が2つある。そういった意味では、活性がもっと高くなっている。これがポジティブになっているのは、いわゆるこれが一つあるからだけではなくて、恐らくこの〇〇〇そのものが実際にAmesに寄与していると考えます。それに関しては、実際にこの物質を〇〇〇でもう一回予測して見ればいいのです。もしくは〇〇〇で予測してみればいいのです。そうすると今、ここで〇〇〇として上がっている、この物質に関して、陽性予測をするほかの因子が見つかるのです。という考え方をするために、このデータを見ます。そうすると、この〇〇〇の結果に関しては、ここのポジティブに関しては、ほかの構造アラートがあるからポジティブになっている。こっちの物質、基本的には類似している物質については全部ネガティブだよねと。そういう意味では予測結果もネガティブだし、こういった判断をしてもネガティブ。

(PP)

もう一つ、〇〇〇のレポートの2行目です。日本語にさせていただいていますけれども、物質n-B。非変異原性ということが、まず、書いてあります。あとは、〇〇〇は〇〇〇や〇〇〇と比べて変異原性が低いところが書かれていますということで、基本的に、n-B から n-D に関しては、さっきの流れで行くと、一定の炭素鎖間は基本的に変異原性が陰性だということを言っているのです。Cは何個ありますか、陰性だと言っていいですねということで、オーバーオールとしては、〇〇〇でも陰性だったし、〇〇〇のここを見ても陰性だということで、陰性になります。ちなみに、〇〇〇は先ほどの予測結果だとポジティブと言っているのです。あれはいわゆるアルゴリズム的にはポジティブになってしまうのだが、アルゴリズムの中に組み入れることができないナレッジに関して、テキストでこうやって説明していく。そういうものを読み込んでいくと、オーバーオールでネガティブという判定が出ます。読んでいただくと、その結果が本当に正しいかどうかということに関して、エクンプラネーションがきちんとついていきます。信用できないようだったら、論文、原著まで当たって確認することもできますということの、そういった使い方ができますという事例になります。

(PP)

次は〇〇〇のケーススタディーをやってみたいと思います。全部Amesですみません。

〇〇〇に関しては、3つのモデルで予測を試みました。モデルとしてはこの3つのモデルがあります。これは〇〇〇のモデル、こっちは〇〇〇のモデル、この2つは統計モデルで、3番目のモデルは〇〇〇のモデルになっています。3モデルでやったのですが、全部アウトオブドメイン、予測対象外になっています。予測できない物質だと言われてしまったのです。予測できない物質だと言っている根拠としては、未知のフラグメントがあるためです。要するに、この構造の中の、この〇〇〇で示している部分、これがトレーニングデータセットになかったから、モデル上は予測できないのだということになってしまっています。

予測できませんというだけではなくて、〇〇〇は、さっきの〇〇〇と同じで、トレーニングデータ物質の中で、この予測した物質のデータを見せてくれます。皆さんのものは、ここを隠されているのですけれども、実際はこういうデータがあります。では、この物質についてどう判断されますかというケーススタディーをやりたいと思います。ちょっとやってみていただきたいと思います。結構簡単と思っているかもしれませんが、罍を仕掛けましたので、2分ほどでお願いできますか。

2分経っていないかもしれませんが、なかなか間がもちづらいので、それでは、〇〇〇、お願いします。

〇〇〇 先ほど、〇〇〇から罍を仕掛けたというコメントで少し心が揺らいでしまっているのですけれども、多数決で陽性的のように、非科学的ですけれどもそう見えてしまう。私はそう判断しました。

〇〇〇 ありがとうございます。

ケーススタディーですから、そういうのが大事なのです。

多数決ではないのですが、まず、構造の類似性というところでいくと、この物質はこれがないバージョンですね。この物質はこれがないバージョンで、完全にこの物質と一致しますね。そういう意味では、完全に構造類似。こっちはちょっとまだ〇〇〇がいっぱいあるので、Amesを予測するほかの要素がありますので、今回の解析からまず除く。そういう観点で、これも除くことができます。この2物質については、この物質のサブストラクチャーが入っているということで、類似構造の物質で、陽性なので、今まで、さっきのパターンでいくと陽性と思うのです。

それで、さっき〇〇〇の1問目、〇〇〇は、〇〇〇があるときのみポジティブだと〇〇がおっしゃっていました。これを加味すると、本当はこれは〇〇〇と〇〇〇の結果を並行してお見せすると、間違いなくさっきの〇〇〇の結果をリファーして、ここのポジティブはうそだということがわかるのです。そういう意味で、この結果は、〇〇〇の結果そのものでいくと、陽性。〇〇〇、正解です。ちょっと演習で意地悪にしたので、ここに〇〇〇の結果を織り込んだらどうなるのかというと、〇〇〇の結果から、これは〇〇〇だか

らポジティブになるのだということを考えると、オーバーオールではこの物質は陰性ということになります、という事例でした。意地悪をしてすみません。

(PP)

演習問題としてはこれが最後になるのですけれども、こちらもAmesのもので、〇〇〇の予測結果になります。3つのモデルで予測をしましたというときのこのいやらしいパターンです。〇〇〇の中にあるモデルの統計ベースのAmesのモデルではポジティブになりました。〇〇〇の統計ベースのモデルだと適用範囲外の予測になりました。〇〇〇のモデルでやると、ネガティブになりましたと、いろいろな全然違うモデルを使って予測したときに、全部結果が異なってしまいました。さっきの〇〇〇がご紹介されたようなやり方をした場合、これはポジティブととろうかということになると思うのですが、エキスパートジャッジメントを入れたらどうなるのかという事例を少しやってみたいと思います。

これはポジティブになっている根拠としては、ポジティブになるアラートというものが4つひっかかってきました。この物質に対して、ここで、〇〇〇で書いているところなのですが、この4つのアラートがひっかかってきたということで、ポジティブだと判定しています。

(PP)

この2番目のモデルに関しては、未知のフラグメントを含むのですが、ポジティブと判断できる要素も含んでいる。それはどこかということ、ここのフラグメントです、ということを行っています。

さっきもお話ししましたが、トレーニングデータセットで似ている物質を見せてくれると言っているのですが、今回、その物質に対してアラートA、B、C、D、Eと、全部で5つあったのです。そのうち、1つ目のアラートを持つ類似化合物について、ここでは見せてくれます。2番目のアラートを有する類似物質がここの結果です。ここから何を判断するのか。小さい字で書いてあるのですけれども、〇〇〇がポジティブで〇〇〇がネガティブな結果となったものになります。

(PP)

今回この予測に関しては、〇〇〇と言って、ネガティブ予測に資するアラートがヒットしました。だからここでの予測結果はネガティブなのです。いわゆる、ネガティブ予測ができるのです。

今回、この〇〇〇はどのようなものだったのかというと、〇〇〇の中で、この〇〇〇があった場合には、Amesはポジティブと本来、判定します。でも、必ずしもそうではない。ここの〇〇〇の周りに〇〇〇だったり、こういったようなフラグメントが来る場合には、活性を減弱させてくれる。ネガティブと見てもいいよと、要するに、組み合わせでネガティブと判定している。このフラグメントを持っている類似化合物はどのようなものかということ、こういうものでした。このようになりかなり複雑な情報になっているのですが、これでエキスパートジャッジメントをどうしましょうかということになります。

ちょっと考えてみたいですか。1分ぐらい考えてみていただければと思います。

まだ、皆さん、途中かもしれませんが、ちょっと進めたいと思います。これについては回答が複雑なので、私から回答したいと思います。

この物質に関して、5つのポジティブを示すアラートがヒットしていました。まず、1つ目のアラートに関してなのですが、まず、1つ目のアラートを持っている類似構造化合物がここにあります。この中で、ポジティブと言っている物質にまずは着目してみたいと思います。ここの3つの物質がポジティブと言っているのですが、これは実は、ここの〇〇〇とはこれは言っていないのですが、〇〇〇がある、1、2、3、4、かくかくという構造を持っている物質についてアラートがついているのです。実は、この〇〇〇というのは、〇〇〇単体で、まず、アラートとして成立しているのです。そういう意味では、この骨格があるからポジと言っているわけではない。そういう意味では、これは違う要素を持っているので、この物質のポジはなかったことにしよう、ということになります。

これについても、物質Bということで、同じ理由で違う要素が入っている物質であるということになります。こちらに関しても、〇〇が2つあるというアラートが実はついているのです。これは皆さんのお手元では見えないのですが、ソフトを使っているときは、こういうものが出てきたときに、すぐ右クリックか何かをしてやるとすぐ〇〇をもう一回予測させたらどういいう結果が出るか、どのアラートがヒットするのかすぐわかるのです。そうすると、今、言ったような判断をすることができます。こちらも似たような考え方で、必ずしもここで言っているアラートというもので、これがポジティブと言えるわけではない。この構造それぞれにあそこにあるアラートとは違うアラートを含んでいて、ポジティブになっている要素もあるということが言えるので、この予測対象で陽性となっているアラートが本当にこの物質でも陽性と言っているかという判断材料は、今は実はないというのがまず1つ目の結論になります。

全部で5個のアラートはあるのですがけれども、考え方は同じなので2つのアラートだけでも説明します。

(PP)

もう一つありました〇〇〇に関する物質、これも類似化合物を見せてくれているのです。これは今、予測したい物質の真ん中に5員環が接続されているので、基本的に平面構造をとっていますね。Amesのときに平面はすごく重要です。

〇〇〇として、このトレーニングデータセットに含まれていた物質を見てみると、ここの〇〇〇という要素があると、〇〇〇していいよと言っているのですが、このトレーニングデータセットに含まれているのは、全部その構造がフレキシブルに回転できる構造になっている。ちょっとだけエキスパートが入ってきているのですがけれども、そういう見方をして、そうなるくと、この〇〇〇と言っている要素としては、この物質については立体構造が異なっているということから、適用できないねというように考えたほうが良いと

ということになる。その〇〇〇構造というものにこの〇〇〇がついていると、本来は陰性と判断していいのだけれども、〇〇〇をつけてネガティブと言っているときの条件が構造的にはそろっていないということになり、今回は完全に陽性の要素をねぐることができないので、陽性という判断になります。

本来だったら、あと、これで〇〇〇だったり、〇〇〇だったり、他のモデルも走らせてみて、その結果を総合的に判断してやるのがいいのですけれども、ケーススタディーなので、これは〇〇〇の結果だけで、今のところ陽性の結果になりますという事例でした。

面白かったですか。ありがとうございます。

(PP)

続いて、食品安全委員会に関連している物質での予測というものを少しお見せしたいと思います。まず、言っておきます。期待しないでください。

(PP)

今回選定した物質については、この4つの物質です。〇〇〇評価をされているもの、〇〇〇評価のもの、〇〇〇評価のものというものを選んでいきます。これは意図的に選んではいません。ただ、意図は入っています。2012年までに評価されたものからまず選んで、その中で、いわゆる化学物質としてみなせるものですね。例えばカビ毒は、混合物だったりという難しい要素もありそうだったと思うので、単純に単体として評価できるものを選びました。基本的にバイアスは入れていません。

(PP)

まず、Amesの結果に関してです。ここの緑のところでは示しているのが、食品安全委員会の評価書の中でどうだったのかという結果です。全部陰性という結果でした。〇〇〇に関しては、全部ネガティブな結果でした。結果はお見せしませんけれども、私が見ても陰性だよという感じなんです。

〇〇〇に関しても陰性。特に、この3つに関してはトレーニングデータセットに既に含まれていて陰性だよという結果でした。

〇〇〇に関しても、2物質がトレーニングセットに含まれていたし、〇〇〇に関しても、1物質が含まれていました。基本的にAmesはもう完璧でした。これはもう皆さん知っているような結果かなと思います。

(PP)

次、急性毒性についてです。

急性毒性に関しては、物質Eに関してはデータがありませんでした。一応、〇〇〇に関しても急性毒性は予測できるということになっているのですが、限定的です。なぜか。トレーニングデータセットが少ないからです。いろいろ解析したくても、データセットが少ないのでわかる知見がちょっとしかない。そういう意味では、ここでアラートがマッチしなかったからといって、では、急性毒性がないのかといたら、これはわからないと判断したほうがいい。

〇〇〇に関しては、〇〇〇のLD₅₀値を算出することができます。実際の結果はこの値で、これをどう解釈するのかというところです。このモデルが使っているトレーニングデータセットは、〇〇〇ぐらい。さっきすごく結果のよかったAmesは5,000とか、場合によっては1万という物質でモデルがつくられているので、かなりトレーニングデータセットが広くて、包含できる予測スペースが非常に広くてということがありますが、それに対して、急性毒性は〇〇〇。しかも、〇〇〇。

(PP)

どういう結果が返されるのかというのは、私のほうのスライドだけなのですが、物質Bの例です。物質Bに関しては、プレディクションの値はグラムパーキログラムで結果を返してくれまして、さっきの〇〇〇や〇〇〇と同じで、この物質と類似した構造、トレーニングデータセットの中にある類似した構造を見せてくれます。小さくて見えないと思うのですが、ぱっと見、似ていませんね。この物質については、実際の実測値の値と、実際、このモデルで予測したときにどういう値だったのかということを見せてくれています。この値を見ると、大体このモデルがどれぐらいの精度で予測してくれているのかということの客観的判断になります。

ここは〇〇〇とって、この化合物とこの化合物がどれぐらい似ているのかということをお知らせして、〇〇〇になればなるほどそっくりさんで、〇〇〇とかだったら、ほとんど似ていませんというレベルなので、トレーニングデータセットの中には、物質Bに似ている物質はありませんでした。ちなみに、この実測値のデータに関しては、このレファレンス情報があるので、確認したいときに確認しに行けるというものになっています。

このモデルに関しては、適用範囲の中で予測できたのかということについてレポートしてくれます。これは完全にアウトオブドメインでした。これは〇〇〇の中では、〇〇〇というスペースの中に入って予測できたのだったらオーケーということなのですが、この物質に関してはアウトオブドメインということが書かれています。また、ここは見づらくて申しわけないのですが、構造フラグメントもトレーニングデータセットに含まれていないものでした。先ほどお見せしたこの値を信用していいのかということ、若干信頼性に欠けるということがあります。今、申し上げた、〇〇〇内の予測だったのか、構造的に適用範囲外だったのかということに関して、サマライズすると、最後、物質Tに関しては、ドメイン内にあるものでした。それ以外に関して、ストラクチャーが外れていたのも、信頼性には欠ける結果ではありますが、参考値としては、このような値が出ました。

(PP)

ちなみに、予測結果に関しては、〇〇〇は今まで御紹介したモデルよりももっと複雑です。構造フラグメントだけではなくて、物化性状に関するパラメーターだったり、本当にいろいろなパラメーター、モレキュラーウエイトからバルキネスから、いろいろなパラメーターを使って予測しています。その予測の中にどういうものを使って予測したのかということをお知らせしてくれないとわかりづらいのですが、このモデルでは、変数、例えば〇個

から○個使って予測していることもあるのです。その変数のうち、○○○だけを閲覧できる機能があります。余り興味はないかもしれませんが。

(PP)

続いて、一般毒性に関してです。一般毒性に関しては、2つの要素で見えています。標的臓器というものと、NOAELがどのような値だったのかということで見えます。食安委の評価では、こういうような標的臓器があったのですが、○○○では、肝毒性アラートがこの2物質で一切何もつかなかった。ただ、○○○ので、これも今のところは参考情報ではないかと思えます。

NOAEL、LOELに関しては、○○○の中では多分、一番○○○というものが近いのかと思えます。参考として、○○○の予測結果も示していますが、今から御説明するのは、○○○のものに関してのものになります。このモデルは、一応予測結果は返しているのですが、トレーニングデータセットが○しかない。これはなかなか信頼性、適用範囲を規定するときに、ものすごく限定された構造の中でしか予測できないということのインディケーションになるのかなと思えます。

物質P、これについて予測した結果なのですが、類似性、やっぱり似ていないですね。ここはぱっと見たときに、似ている気もするかもしれないのですが、ここの○○○を見ていただくと、○以上の値になっているので、余り似ている構造ではないということが言えます。

適用範囲には入っているのですが、構造フラグメントとして、トレーニングデータセットで知らないフラグメントが5つもある。このように、現時点では一般毒性の予測についてはまだ難しい部分があるということが言えるかと思えます。

(PP)

発生毒性、催奇形性ですけれども、こちらに示しているような結果で、4番目の物質についてはトレーニングデータセットの中にあつたと。○○○と予測しているのは、予測モデル上は陰性なのですが、トレーニングデータセットで陽性になっているから、さっきのようなエキスパートジャッジメントを行うならポジティブという判定をすることはできません。単純に予測結果だけを返すとこういう結果になってしまうというもので、催奇形性に関してもトレーニングデータセット○であり、Amesが数千、数万という単位でつくっているモデルに比べると、圧倒的に少ないということがわかるかと思えます。

(PP)

発がん性に関してもモデルはあるのですが、○○○、○○○、○○○のデータを使って、○でつくっていますが、現時点ではこういうような性能を示しているという状況です。

(PP)

今、お見せしたのは、いい悪いという説明ではなく、今、予測できる環境にあるのかどうかというところのざっくりした印象を持っていただければと思って紹介をさせていただきました。今からは、食品安全委員会で評価した物質で、Read acrossで評価したらどうな

るのかという事例を、Read acrossは何ということについて理解することの延長で御説明したいと思います。

(PP)

使うツールですけれども、OECD QSAR Toolboxを使ってやりたいと思います。このツールはどのようなものかということ、類似物質を探して、Read acrossによるデータギャップの穴埋めを支援する無償ツール。支援するのです。判断してくれません。さっきの(Q) SARモデルもそうなのですが、〇〇〇といった結果をそのままのみにするのではなくて、人が目で見て、テキストを読んで、オーバーオールで判定する。ああいうような要素が基本になります。Read acrossもその要素が非常に強くなってきます。このモデルは無償なので、誰でも使えるということで、ほとんどの私の知っている(Q) SARを使っている行政は導入しています。

開発しているのは、ブルガス大学なのですが、OECDや欧州化学品庁が出資をしながら、相談をしながらつくっています。

ツールボックスをつかう、一ついいところは、類似性に関する根拠情報を提供してくれるというところがあります。先ほどから類似性、類似性と言っているときに、何をもって類似性と言うのかということが非常に難しいので、その要素を少しでも提供してくれるという意味では、非常に有用なツールかと思います。

あとは、各国から大きなデータベースが提供されていて、この中に10万化学物質分のデータベースが蓄積されています。毒性に関するものも当然入っています。多分、今、世界中にある、(毒性に関する)最大のデータベースではないかと思います。

さらには、自分たちが持っているデータベースを統合して、インハウスで使用できるというアプリケーションもあります。

(Q) SARツールボックスで、何をしますかということの流れについてですが、(Q) SAR Toolboxで、まず、予測したい物質についての入力を行います。次にプロファイリング、この予測物質の特徴的な構造を特定していくということを行います。このプロファイリングを何のために行うのかといいますと、この後、類似物質を探すという工程のときに、このプロファイリング、特徴的な構造が似ている物質を探し出すということをやります。

3番目の工程としては、もしかしたら既存データがあるかもしれないということで、その中の既存データがないかどうかということを探します。

やっぱりないねというときに、では、Read acrossをしようかねと。では、まずは類似構造を探そうかということになります。ここで類似構造を見つけたら、では、実際にRead acrossを使って、データギャップの穴埋めをやってみましょうと。それで、終わったらそれをちゃんと文章に起こしてくれるというサービスまでついているかなりすてきなツールになっています。

ターゲット物質の入力に関しては、左上のボタンを押して上げると、ドローイングツールがあったり、CAS番号を入れてあげる、あるいは化学物質の名前を入れてあげたら、この

構造でいいですかと聞いてくるので、オーケーを押してあげると、すぐに構造を入力してくれるという、かなりユーザーフレンドリーなツールになっています。

(PP)

次に、プロファイリングを行います。このプロファイリングは基本的に入力した物質が持っている化学的反応性に基づいてプロファイリングをするというのが一つのコンセプトになっています。

プロファイリングにどういうことがあるのかというと、皆さん、お手元の資料を見ていただきたいのですが、基本的に、ここで書いてあるのは、経験則的に知られているもの、例えばケミカルエレメント、どういう元素が使われていますかということだったり、あとは官能基、どういうものが入っていますかということを見るプロファイリングのやり方もありますし、エンドポイント特異的な、例えば、Amesに関連するプロファイリングをする、発がんに関連したプロファイリングをするというようなこともできます。感作性に関するものであったり、ということができます。あと、物化性状に関するものであったりというような、いろいろな考え方のプロファイリングをすることができます。

これを使って、このプロファイルが共通しているから類似だという言い方をします。同じ官能基があるからという類似性もすごくわかりやすい説明の一つと思うのです。それを客観的に出してくれるというファンクションを持っています。

プロファイリングが終わったら、今度は既存情報があるかどうかの確認ということを行います。先ほど、物質Mについて、既存物質がないかなということを探したら、ありました。これは何についてあったのかというと、Amesだけだったので、急性毒性を予測したいときには、データはないという結果になります。

(PP)

これから急性毒性を予測したいなというときに、急性毒性を予測するために使えるプロファイルは、実はまだないのです。そうなってくると、構造に関する経験則のプロファイリングを使おうと。「似ていたら同じ性質を持っているはず」のあの事項に入っていきます。そこで、プロファイルの中の、経験則に基づくプロファイルを使うという予測をしたいと思います。

(PP)

この物質に関しては、プロファイルを見たときに、オルガニックファンクショナルグループのUS EPAがつくったやり方が、この物質の形をよく説明しているなというところで、このプロファイルを使って、類似構造を探していきたいと思います。ちなみに、この判断は主観です。これは絶対的なルールがないのです。ここがちょっとだけ (Q) SAR とか違って難しいところはあると思います。

(PP)

私が先ほどのUS EPAのオルガニックファンクショナルグループでプロファイルをする、(Q) SAR Toolboxの中に入っている10万の化合物の中から204物質だけを抽出することが

できました。実際、その204物質がデータを持っているのかどうかぼちっとやってくれたら検索をかけてくれて、データがあるかないかも全部見せてくれます。

構造を見てみると、ぱっと見、似ていないというものがまだあります。こういったときにどうするのかというと、サブカテゴリーということをやっていきます。それで、実際にもう一個別のオルガニックファンクショナルグループのプロファイリングがあるので、それを使って絞り込みをかけてみると、204あったものが7物質に絞り込まれていって、この中身を見てみると、大概似ているよねと。ここが主観なのです。ここが難しいところで、ここでオーケーとする根拠はなんだと言われたときに、若干、かなり専門的なエキスパート判断が入る、説明が必要になってくるというところ。これがRead acrossの難しいところで、ただ、それさえできてしまえば、ある程度予測を行うことができます。

(PP)

この物質に関しては、急性毒性をやりたいと思います。「2/13」とあるのは、2物質について13個の急性毒性に関するデータがヒットしています。見て見ると、この4番の物質と7番の物質について、急性毒性のデータがありますということがわかりました。このデータを使って、Read acrossをしたいと思います。

(PP)

ぼちっとボタンを押したら、もうぼんと出てきます。これはどういう結果なのかということの説明したいと思うのですが、基本的にToolboxの中でやっているRead acrossというのは、データを持っている最も似た物質の、たしか5物質を使って、その平均値をとって、実際の予測値にします（デフォルトが5物質というだけで、増やすことも減らすこともできる）。今回のケースだと2物質しかデータを持っていなかったんで、2物質の平均値をとって、この値が来ています。これがこのToolboxを使った場合のRead acrossのやり方となります。

Read acrossのやり方は絶対にこのやり方でやらなければいけないというわけではなく、かなり主観を持ってやるところがあるので、これをどう運営していくのか、そもそも使うのかどうかということのをこれから考えていかなければいけないのではないかと思います。

ただ、これのいいところは、(Q) SARで予測できない物質について、この方法でカバーができるというところがあります。ですから、必ずしも必要ない手段であるとは思っていませんというところで、終わりたいと思います。

〇〇〇 どうもありがとうございました。

クイズもやっていただいて、なかなか慣れていないと難しいなとは思いましたがけれども、いい刺激になりました。ありがとうございました。

それでは、今のデモに近いですね。デモと言っていいと思いますけれども、それにつきまして、何か御意見や御質問はありますか。

このRead acrossもないものを埋めてくれるという感じですけれども、どうしてもだめ

という結論もありなのですか。

〇〇〇 そのとおりです。データベースの中に、いわゆるソース物質、ある物質のデータを補ってくれる物質がなかった場合には、できませんという結果にはなってしまいます。

〇〇〇 いかがでしょうか。

割とユーザーフレンドリーで、ポチッと押すと、結構出てくる。その読み方にちょっとコツがあるというのはわかりましたけれども、割と答えに近いところまで出してくれている。

〇〇〇 かなり近いところに行ってくれて、よく知らない方が思うのは、出てきた結果が結果であるというようになるので、テキストを読んで判断してほしいとかという、単純に予測結果だけ見るのではなくて、そこの横に書かれていることにも気を配ってほしいということをお願いしたかったのです。

〇〇〇 それは非常に重要なことで、最後の判断はエキスパートとしての見識を持ってやらないといけないというのは確かですね。そのヒントになる近い線の情報ではこうですよということですね。

いかがでしょうか。きょうはAmesの割と答えが出ているものについてですが、これが最後に紹介していただいた、毒性の種類によっては、なかなかデータがそろっているわけではないとか、あるいは判断がそもそも難しいというものもあるのではないかと思いますので、特に発がん性、遺伝毒性系のものについてのほかに、急性毒性とかですか。キャラクターが全然違うものについて、同じように使えるものかどうか。使われた感触というものはいかがですか。

〇〇〇 基本的には、考え方は同じなのです。適用範囲の中に入っているか。それで、トレーニングデータセットと類似化合物での結果がどうなのかという見方をするのですが、いかんせん、そもそも適用範囲外なのです。トレーニングデータセットがどうしても小さいので、ケミカルスペースがすごく小さい。それで、今回みたいに食安委のものをやってみると、そこからどうしても外れてしまうので。では、この後、どうしようもないのかというと、そうではなくて、私が今日紹介したのは、あくまで〇〇〇の例なのです。〇〇〇では、たかだか〇程度のトレーニングデータセットでしかつくっていないのですが、例えば急性毒性に関しては、今、おそらく、1万以上のデータがそろっていて、そういったデータを使ってモデルをつくっているものもあるのです。すぐに名前は出てこないのですが、例えばそういうものを使うことはできる。使えばいいと思うのです。基本的には、トレーニングデータセットができるだけ大きければ大きいほど、適用性は上がります。

あとは、そもそも予測するスペースがかぶっているのかどうかというのも、あらかじめ知っておくと、余り外れないモデルを使えるのではないのかなと思います。あるいは、自分が持っているデータセットを使ってモデルをつくるというのも一つの選択肢で、アルゴリズムを明らかにするというのが非常に難しいのですが、今日紹介したツールは、ほとんどがアルゴリズムはそのまま使ってよくて、トレーニングデータセットがあれば、それ

でモデルをつくることができます。ですから、そういう意味で、アルゴリズムについてのエクストラネーションに労力がまずないのです。トレーニングデータセットは、公開することができるのであれば、そこである程度透明性の障壁はクリアできているのではないかと思います。検証も自分たちでやればいいので、そんなに難しい話ではないと思っています。

〇〇〇 ありがとうございます。〇〇〇、どうぞ。

〇〇〇 どのソフトがどれかややこしくなっているのですけれども、最後のApplication Toolboxだと、よく似たのを集めてくるということによくわかるのですが、前のほうは、アラートに一致したものだけが拾われる場合が多いのですか。まず、アラートを抽出して、それに持っているか持っていないかという判断で、具体例を出してくる。

〇〇〇 基本的には、トレーニングデータセットと似ている物質を持ってきている。

〇〇〇 単純にメカニズムとかアラートとか関係なしに、類似性でも出せるわけですか。

〇〇〇 基本的には、類似性で持ってきています。ですから、余りバイアスのかかったことはしていません。

〇〇〇 わかりました。

〇〇〇 よろしいでしょうか。

ありがとうございます。〇〇〇、大変多大な労力を使わせましたけれども、いい勉強になったかと思います。ありがとうございます。

それでは、残りの時間で今後のことを考えていきたいと思っています。(Q) SARとかRead acrossの概括的なこと、そして、実際にどうやって動かすかというようなこと、おおよそイメージできたかと思いますが、今後我々がどういう行動をとっていか。つまり、食品安全委員会の評価として、これをどう取り扱うかということに入りたいと思います。

今までの議論で試行して見るということ、今日はもうわかっているものについてデモがありましたけれども、我々が検討すべき物質などを使って試行を重ねていくということになりますし、それから、そもそもその前に、データベースをどうするのかといった問題もあろうかと思っています。

まず、今、食品安全委員会では独自のデータベースを持っているわけではありませんが、この(Q) SARとかRead across、今日見せていただくと、その利用までにどういう申請とかライセンスとかがありますのでしょうか。

〇〇〇 基本的に行政が使うとなると、手続以外は、本当にただ契約をすればいいだけになって、基本的に中身もある程度公開することは多分、可能だと思うのです。ただ、一般の契約と、多分違う契約の要素が必要になるのかなという気はしているのです。例えばどういった変更があったのかということ詳しく教えてほしいとか、そういったことは多分、組み入れる必要はあるのかなと。

先生の質問の趣旨と違う答えかもしれませんが。

〇〇〇 これは、つまり、要するにコマースリーアベイラブルなものです。〇〇〇というものを除いたら、バージョンはアップし続けられていく。

〇〇〇 そうですね。それで、全部日本のベンダーと契約することで使えるようになります。

〇〇〇 だから、一般のいろいろな医療機器とか、いろいろなソフトを利用するのとそう変わらない。要するに、利用権を買うということなのですね。思いのほか、さあどうぞという感じになっているので、食品安全委員会として、どうしたものかと、またちょっと実は原点に戻ってしまったところがあるのです。これをそもそも我々が独自に何かやっていくのか、こういうものをそのまま使うのか。

データベースについても、片方で食品安全委員会が今までいろいろなことについて、評価書もしっかり吟味して、この各専門調査会の吟味の結果というのは決して軽いものではなくて、相当な重さがあると私は考えていますけれども、その評価した、真剣に評価したものをきちんとした形で、将来の自分たちのために、それから、世界のために、きちんとデジタルアーカイブ化していくということは、それはそれで私は必要だと思っていて、それも一つの財産になるのですけれども、それとは別に、今、商業ベースであるものを食品安全委員会としてある程度使っていくのか、それとも、こういうものと距離を置くのかとか、その辺のところをひよっとしたら委員会の根本的な方針もあろうかと思うのですけれども、何か委員の先生方でお考えのことはありますでしょうか。

〇〇〇 勝手な個人の意見ですけれども、少なくとも、変異原性に関しては、ほとんど完成品で、データがバージョンアップされて、追加されてナレッジベースになっているので、もう自分たちでつくることはない。変異原性に関しては。

結局、毒性の側のほうは、かなりデータについてどれを使うかといったら、〇〇〇を自分たち用にモディファイして、自分たちのデータを入れていくというのが一番手っ取り早いし、利用価値も多いような気がするし、もう一つ考えておかなければいけないのは、我々が将来的に使った場合には、どういう根拠でどういう形でやりましたという、ある意味では、透明性を確保しなければいけないところがあるので、それに耐え得るような形で導入をしておかないと、企業側も突然だめですと言われると、堪えるでしょうし、その辺のところを、ある程度、それが透明性を確保していれば、そういうことがないのかなという気がします。

〇〇〇 ということは、もちろん、わかっている変異原性についてはいいのだけれども、そのほかの毒性評価はまだまだだから、自分たちのデータベースをもう一回構築していく。

〇〇〇 追加してできるような形で。

〇〇〇 追加して、公開されているものはそのまま利用して、それに上乘せして、多少そのスペクトラムが違ってもいいし、特に食品というもので、物質の特性もあろうかということで、そちらを中心に据えたデータベースをつくっていくということが基本線になるかということですかね。

そうすると、そういうことを、今までの商業的なシステムでは十分にはかなわないけれども、我々がいろいろ評価してきたものを、やはりきちんとデータベース化する作業を行って、それで今の既存のものではカバーしていないところは自分たちのツールをつくるということに。

〇〇〇 なると思います。

〇〇〇 あるいは、逆にデータベースをつくって、ソフトウェア自体は材料を提供すれば、あるいは契約を結んで利用できればいいのだけでも、そのデータベースに足していくということと理解すればいいですか。

〇〇〇 そうですね。それで、結局ためたのだけれども使わなかったら何の意味もないので、そこの何にまず使えるのかということのターゲットをある程度絞りつつ、それを増やしていくことになってしまうのだと思うのですけれどもね。

〇〇〇 今、現実、なくてやれているというか、やっている部分もあるのだけれども、これから化学物質も新たにどんどん開発されていくということも考えると、スクリーニングのツール、それから、本当にデータがないときのエスティメーションには必要性はあるだろうとも思いますし、国民の要求レベルも高くなってくると考えると、いろいろな質問に回答できなければいけないということになると、その評価の仕組みは自分たちできちんと押さえておいたほうがよいということになるのかなとも思います。この辺、ちょっと技術的な評価だけではなくて、ポリシーの問題も絡んでくるので、〇〇〇と本当に意見交換をしながら詰めていかないといけないなと思いますけれども、どうでしょうか。

〇〇〇 直接ではないのですけれども、今日、話を聞いていて、〇〇〇が教師なし学習によるということで、これはもうディープラーニングでAIの世界に入っているのではないかと期待しています。Watsonで、抗がん剤のデータを全部入れて個人情報を入れて、最適化した抗がん剤ができるように、今、AIとまでは言えないと思うのですけれども、いろいろな企業などが、共通した特徴を見つけ出してみたいなことがかなり進展しているようにも見えるのです。その辺の動きが速いのではないかと考えていて、これらの動きについて教えていただければ将来考えるときにも役に立つのではないかと考えるのです。

〇〇〇 いわゆる統計の中にディープラーニングもあるのです。さっき言った、物すごく簡単な単回帰モデルも統計モデルの一つ。それがだんだん複雑になると、人工知能というか、あれは実は人工の知能ではなくて、あれは単純に統計モデルの手法のネーミングでしかないのです。すごく複雑で、大量のデータを処理できるということです。データセットとして、Amesでも、せいぜいデータが1万とか1万5,000程度。そのレベルだと、例えばディープラーニングレベルの解析は必要ないのではないのでしょうか。あれが10万とかになってきましたと。そうなったときに、初めてディープラーニングレベルの、要するに情報がリッチになればなるほど複雑な解析をせざるを得ないのでというところがあるので、今はまだそこまでは行っていないのではないかと。ただ、エンドポイントによって非常に複雑、例えば一般毒性、ああいうものに関して、N数ではなくてポイント数、例えば血液

生化学もあって、尿検査のパラメーターも見られて、いろいろなパラメーターをマトリックスにしてやると、データポイント数はかなり大きくなる。そういうものに対して適用するという考え方はあると思います。私の知る限りは、まだやっているという話は聞いたことがないかと。

〇〇〇 〇〇〇、いかがですか。

〇〇〇 AIの話から離れてしまうのですけれども、先ほど〇〇〇からToolboxの話がちょっと出ました。Toolboxは入っているデータも多くて、無料で使え、カスタマイズもできる。ただ、Toolboxは入っているデータはほとんど工業化学物質のもので、食品安全委員会が取り扱っている物質と、ケミカルスペースが違っているということが一つの問題です。Toolboxを使っていくためには、地道だけれども、データベースをこつこつつくっていくことは必要なのかと。一度、構造や試験データを電子化しておいたらいろいろなところに使えそうな気がしています。例えば今、行われている評価において、似た物質だったらどうなのだろうという情報は参考にはなるときがあるのではないかと。今までの試験をやっていたものを予測で置きかえるとなるとハードルが高い場合が多いのですけれども、まず評価における追加のエビデンスとして使うというのも一つあるのかなということも思ったところです。

〇〇〇 今、ない領域、ターゲットがずれているから、独自のものをつくって、それをスクリーニング、予備的な検討、それから、疑問が起きたときの参照というような使い方があり得るといってお話でした。

何かありますか。

〇〇〇 例えばToolboxなのですが、EFSAなどの一般毒性に関するデータベースだったと思うのですが、そういうものをみんなに使えるように組み入れてもらえると、(Q) SARディベロッパーからすると、モデルをつくるためのデータセットを提供してくれたということで、開発の促進にもなるという側面もあるのかなと思うので、個人的にはぜひ進めたい。

同じ食品でも、日本とアメリカとヨーロッパで、少しはスペースが違うのかもしれませんが、基本的に構造ということで言うと、かなり近いスペースのところで行っているのかなと思う。データベースでも広ければ広いほど、その中のアプリケーションは広いのかなと思うので、EFSAなどでもかなり自分たちの評価したものをみんなに使ってもらうようにデータベース化してという活動をやられているようなので、そういうことと似たようなというか、例えばテンプレートを一緒にする。要するに、データの交換が簡単に、全然フォーマットが違って話にならないようだとすごく使いづらいと思うので、フォーマットだけでも合わせておくとか、そういうことがあるのは一ついいのかなという気がします。例えば急性毒性、私が持っているツールはあそこで限界だったのですが、もっといいモデルが最近できてきているのです。そういった意味では、最新のモデルの情報というものは、逐次収集していったほうがいいのかと。さっきの事例が全てではないので、そういうこ

とはあるかなと思いました。

〇〇〇 その点、事務局で3極の姿勢といいますか、調整、データベースの共有に向けたものというのは、話し合いはありましたでしょうか。ヨーロッパ、EFSAとか。

〇事務局 海外とのデータシェアリングについては、まだ具体的な話はしておりません。ただ、それぞれの機関との連携が可能であるかどうかから検討しております。

〇〇〇 ありがとうございます。

そうしますとやはり既存のいろいろなデータベースあるいはツールがあるけれども、我が国の食品安全ということを中心に据えたデータベースをこれから構築していくことを念頭に置いて試行をしていくということかと思えます。〇〇〇。

〇〇〇 多分どこで試行していくのかというと、さっきの構造が違うとかという話もある中で、試行をどうやってやるのかというのは、さっき言った、結局ただ単に結果を使うだけではなくて解釈も使っていくというのを、もうちょっと具体的なところを少し模索するのも将来的に考えてもいいのかなと。余り具体的に言うと、いろいろ差し障りがあるかもしれないけれども、化学構造から言うと、たとえば器具・容器から出てくるような化学物質については、この化学物質のケミカルスペースに近いかな。でも、そのかわりデータがないので、完全に今度は予測の世界に入っていく可能性がある。

〇〇〇 遠いところからになってしまうかな。

〇〇〇 そこは検証が難しいかもしれないけれども、でも、やってみたらどんな感触かという感触をつかむには、そういう試行もいいのかなとちょっと考える。あと、グループからだ香料なので、香料も大体同じ、先ほど言いましたけれども、JECFAは昔からグループ評価でやっているので、あれは完全なRead acrossなので、一連の香料は全部一緒に評価して、ADIもグループはADIにしているという、かなり定量的にやっています。ただ、*in silico*ではなくて、マニュアルということに注意が必要ですけどもね。そういうところも見つつ、可能ところで少し食品安全委員会がさわっていくことは必要なかと思っています。

〇〇〇 では、今、そういうデータベースを充実させるということ。それから、恐らく世界のヨーロッパやアメリカとの連携も将来的にはあり得ることも想定しておく。国内でも、厚労省とか経産省とか農水省、あるいは環境省ですね。そういうところとの連携というのもあり得るということで、しばらく試行を含めた模索を続けていくということになるかと思えます。

具体的に何をテーマにしてやるのかということはもうちょっと詰めないといけないとは思いますが、これは一定の間隔で会議体を開いていますが、事務局は常にオープンしているので、思いついたことがあって、これをやってみたらとか、こんな情報が入りましたということがありましたら、事務局に集約していただいて、随時検討していきたいと思えます。そんなに大勢のワーキングではありませんので、常時メールで相談しながら、新しい情報をお伝えしていければとも思っています。積極的にアイデアがありましたら、御提案いただきたいと思います。

あとは、〇〇〇、何かコメントはよろしいでしょうか。

〇〇〇 今日お話を伺って考えたのは、直接リスクアセスメントに使うまでは随分距離があるなという感じはするのですけれども、ただ、我々の考え方を少し柔軟にするためには、いろいろやってみるという手もあるのだろうなという感じなのです。

今日、かなりはっきりしたと思うのは、Amesは結構しっかりしたものが出てくると言うけれども、でも、それは逆に言えば、現実の世界でもAmesはちゃんとしたデータは出てくるし、データがなくても割合、簡単に出てくるのだと思うのです。その辺のところは、現実の世界でなかなか出ないものが出てくるところまではなかなか行っていないのだと思うのですけれども、ただ、頭を柔軟にするためにこういったものを使ってみた上で、もう一回これを見たらどうなるというような検討は非常にいいのだろうと思いました。

それから、データベースをつくるという話は、恐らくこれとも関係なしに重要なことだと思うのです。食安委が積み上げてきたものがあって、ただ、それはつくりやすいもの、それから、さっき〇〇〇がフォーマットのような話をしていたけれども、非常に最終形はどうかかわかりませんが、使いやすい形にしておかないと意味がないのだろうという感じでした。

〇〇〇 ありがとうございます。先生方、あるいは事務局から特別コメントはありますか。

〇〇〇 実を申しますと、国際機関における議論では、たとえば農薬の代謝物の毒性をどうするのかということで、Read acrossとか(Q) SARとかを使っていこうという形になっています。ただ、やはりそのときに一番大切なのは急性毒性と遺伝毒性ということになりますので、確かにそのときにすぐデータがあればいいのですけれども、予測をしていくということを喉元まで突きつけられている時代になってきたかなというように思っています。

あと、データベースについても、なかなか食安委は詳細データを評価書に記載していないことがあるものですから、つくりにくい事情はあるかもしれませんが、それなりの情報は評価書の中に入れ込んでいるものが多いと思いますので、データベースの形はぜひ必要だなと思っています。

〇〇〇 ありがとうございます。

それでは、データベースを含めて、方向性をやりながら出していくということになるのかと思います。

例えば、今後1年ぐらいかけてそういうものやっていく。1年といっても、1年間何もしないということではなしに、定期的集まって状況を確認して、軌道修正すべきか、あるいはどんな基準をつくるべきかというようなことは、きめ細かく打ち合わせしたいと思いますけれども、しばらくの間、この*in silico*のやり方について、少し実際にチャレンジしながら進めていってはどうかと考えます。

その間、作業を進めていく間、もちろん打ち合わせはしますけれども、ほかの課題もありまして、例えばベンチマークドーズ法などにつきましても、まだ検討しなければいけま

せんし、それ以外にもまだ課題はありますので、そちらの作業も並行して始めていきたいとも思いますが、よろしいでしょうか。

では、今日のところは結論を出すとかということではないのですが、方向性を出すためのイメージづくりということで、今日〇〇〇、〇〇〇には特にいろいろお示しいただいて、イメージが少しずつ湧いてまいりました。本当にどうもありがとうございます。

それでは、今日の審議というのはここで終わりとしまして、事務局にお返しいたします。
○事務局 本日いただきました御意見に基づき、次回の準備を進めたいと思います。

次回の議事及びお集まりいただく日時につきましては、決まり次第、お知らせします。

以上です。

〇〇〇 それでは、本日の審議は終了いたしました。第3回のワーキンググループを閉会いたします。

どうもありがとうございました。