

食品安全委員会評価技術企画ワーキンググループ

(第2回) 議事録

1. 日時 平成28年8月24日(水) 14:00~16:40

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) 米国FDAにおけるデータベース構築及びその活用について(米国FDA、The National Center for Toxicological Research (NCTR) Weida Tong博士による講演及び意見交換)

(2) 国内外の反復投与毒性に関するデータベース及び毒性予測ツールの現状について

(3) その他

4. 出席者

(評価技術企画ワーキンググループ専門委員)

川村座長、赤堀専門委員、岡田専門委員、小坂専門委員、小関専門委員、広瀬専門委員、山田専門委員

(添加物専門調査会・栄養成分関連添加物ワーキンググループ専門委員)

頭金専門委員

(汚染物質等専門調査会・加熱時に生じるアクリルアミドワーキンググループ専門委員)

吉成専門委員

(食品安全委員会)

佐藤委員長、山添委員、熊谷委員、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、橘評価技術企画推進室長、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 Regulatory Application of QSARs

資料2 国内外の反復投与毒性に関するデータベース・毒性予測ツールの現状について

6. 議事内容

○川村座長 それでは、時間になりましたので、ただいまから第2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、御出席いただきまして、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7名の先生方に御出席いただいております。

また、本日の議事に関しましては、御専門でいらっしゃる頭金正博先生、吉成浩一先生にも御出席いただいております。

そして、本日の議事の1つである毒性データベース構築に関しまして、多数のプロジェクトに携わっていらっしゃるWeida Tong先生を米国からお招きいたしております。

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いております。よろしく願います。

議事に入ります前に事務連絡がありますので、事務局からお願いいたします。

○井上課長補佐 事務局の人事異動について御報告いたします。6月17日付で事務局長の姫田が退職いたしまして、後任として川島が着任いたしております。川島は公務があるため、遅れて出席いたします。申し訳ありません。

前任の事務局長である姫田より、先生方への退職の挨拶文がございますので、机上に配布させていただいております。

また、7月4日付で評価技術企画推進室長の高崎が異動しまして、後任として橘が着任いたしております。橘より挨拶申し上げます。

○橘評評価技術企画推進室長 橘でございます。どうぞよろしくお願いいたします。

○井上課長補佐 なお、内閣府において、5月1日よりクールビズを実施しておりますので、御理解、御協力のほど、よろしくお願いいたします。

以上です。

○川村座長 それでは、議事を進めてまいります。

本日の議題は「(1)米国FDAにおけるデータベース構築及びその活用について」と「(2)国内外の反復投与毒性に関するデータベース及び毒性予測ツールの現状について」です。

開催通知等で御連絡いたしましたように、本日の会議につきましては公開で行いますので、よろしくお願いいたします。

事務局から資料確認をお願いいたします。

○井上課長補佐 お手元に議事次第、座席表、専門委員名簿のほか、

資料1「Regulatory Application of QSARs」。

資料2「国内外の反復投与毒性に関するデータベース・毒性予測ツールの現状について」。

こちらを御確認いただけますでしょうか。これらは近日中にホームページに掲載されます。

また、机上配付資料として、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。

配布資料の不足等はありませんでしょうか。過不足等がございましたら、事務局までお申し出いただけたらと思います。

○川村座長　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を行ってください。

○井上課長補佐　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いたします。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。

○川村座長　先生方、御提出いただいた確認書について相違はありませんでしょうか。よろしいですね。

（「はい」と声あり）

○川村座長　それでは、本日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1つとして、食品健康影響評価に*in silico*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ということを検討することが挙げられます。これらの現状と課題につきましては、第1回評価技術企画ワーキンググループにおいて先生方に御審議いただいたところです。

この評価技術企画ワーキンググループの専門委員の先生方は、それぞれ一人一人が専門家でいらっしゃいまして、それぞれの方面の専門家でいらっしゃいますけれども、その分野は少しずつ異なっておりまして、食品健康影響評価においてはさまざまな手法を比較したり、融合したりして、国としての判断というものをしていかななくてはいけないので、一人一人がそれぞれの専門家でとどまっていたらいけないわけで、総合的な評価を行うために、評価の手法について、あるいは評価の技術やその水準について、相互に理解を図っていくことが必要であります。

前回のこのワーキンググループでは、*in silico*の評価手法の使い方とか、あるいは限界について、ワーキンググループの中が共通の理解に立つために、まずはデータベースの構築や活用の御経験がある先生からお話をお聞きして、勉強会という形でスタートするということを御提案いただきました。食品健康影響評価に必要なデータベースの現状の把握、課題の抽出、評価するエンドポイントあるいはアウトカムを明確にすることは、今後のデータベースのあり方の審議に重要であるということを御提案いただいております。

今回は、今後の審議を円滑化するために、まず反復投与毒性の予測について国内外の現状を共有すべく、お二人の先生に御講演いただく予定です。まずはアメリカFDA、NCTRのWeida Tong先生から、これらの評価方法の導入が進んでいる米国での実態について御講演いただきます。そして、その後、山田隆志専門委員から国内外の反復投与毒性データベー

スの現状について御講演いただきます。そういう予定で今日は相互の理解というか、専門委員がきちんとした共通の理解の上に立って、今後の検討課題を進めていくための基礎をつくっていくという方向で議事を進めたいと思います。

それでは、まず議事(1)の「米国FDAにおけるデータベース構築及びその活用について」を始めたいと思います。Weida Tong先生、御講演の準備をお願いいたします。

ここに座ってしゃべるのですか。

○井上課長補佐 座って御発表いただきます。御発表の前に川村先生、お願いいたします。

○川村座長 今から紹介しますけれども、前に行かなくていいということであれば。

まず、Tong先生の御略歴を紹介させていただきます。Tong先生はFDAの研究機関であります国立毒性研究センター(NCTR)の生物情報学生物統計学部の部長でいらっしゃいます。欧米のさまざまな科学プロジェクトの委員や、米国や中国の大学で非常勤講師もいらっしゃいます。先生の研究グループでは、マイクロアレイの品質コントロール事業、肝毒性知識ベースの開発、希少疾病の治療促進のための*in silico*手法による薬効の再評価、FDAの生物情報システムであるアレイ・トラックの開発、内分泌かく乱物質知識ベースの開発など多数の*in silico*評価手法にかかわるプロジェクトを進めておられます。先生は現在まで230以上の科学論文や書籍を発表していらっしゃいます。

このような御経験をお持ちの先生でありまして、これからお話を伺いますが、御講演いただく前に少し訂正があるようですので、事務局からお願いいたします。

○井上課長補佐 資料の訂正があるので、御確認をお願いいたします。資料1の3ページをご覧ください。下に和訳資料を用意しましたが、英訳にあります「7. Office of Regulatory Affairs (ORA)」の訳が抜けておりました。和訳のほうの訂正をお願いいたします。「6. たばこ製品センター」の後の「(ORA)」を削除していただき、「7. 規制業務部 (ORA)」を追加ください。現在7番の「国立毒性研究センター(NCTR)」を8番としていただきたいと思います。大変申し訳ありません。よろしくお願いいたします。

以上です。

○川村座長 ありがとうございます。何回も読んだのですけれども、私も気がつかなくて見落としておりました。

それでは、Weida Tong先生から資料1に関して御講演いただきます。その後、質疑応答を行っていきたいと思います。それでは、Tong先生、よろしくお願いいたします。

○Dr. Weida Tong 座長、ありがとうございます。今回の公開ミーティングにおいて、お話できることを非常に光栄に思っております。井上先生、このたびはこの素晴らしい機関に御招待いただき、ありがとうございます。また、委員会のメンバーの皆さん、今回話す内容について、様々な御意見をいただき、ありがとうございます。皆さんに即した内容であるという形だけでなく、また、これからの討議を生産性の高いものにするために皆さんの非常に貴重な御意見をいただきました。

(PP)

本日、委員の皆様からの御提案に則して、このプレゼンテーション資料を作りました。まず最初に、私どもNCTRがどういった機関で、私どもが何をしているかということ。次に、よいQSARモデルの開発方法についてお話しし、その後はほとんど1つのプロジェクト、これはかなり長期間で行っているプロジェクトなのですが、これについてお話をしたいと思います。それが内分泌かく乱物質ナレッジベース、知識情報データベースになります。これについてお話しし、また、自前で作った場合の組織内で開発したツールと市販ツールの違いについて、特にその規制に応用した場合、適用した場合の違いについてお話ししたいと思います。最後に、化学物質規制のためのQSARのベストプラクティスについての私の意見で締めくくりたいと思います。

(PP)

まず最初に、私ども米国・食品医薬品庁（FDA）の紹介から始めたいと思います。FDAには8つのセンターがございます。下の日本語訳をごらんいただくとわかると思いますが、今、足していただいた、この7番の規制業務部、こちらは非常に重要な調査をする部門となっております。残りの7つにつきましては全て「Center for」という形で始まっておりますが、例えば1番ですと、医薬品を規制するセンター。2番目が生物学的製剤を規制するセンター。3番目は見ておわかりのとおり、食品を規制するところになりますが、私どものところだけ「National Center for」と始まっております、2つの部分でほかのセンターと違っております。

(PP)

まず1つ目の違いですが、場所です。FDAのほかのセンターは全てワシントンD.C.にあるにもかかわらず、私どものセンターはアーカンソー州のもっと南の暑くて蒸し蒸しするところがございます。

そして、2番目の違いですが、他のセンターは全て規制の責任も負っておりますが、NCTRだけが研究のみに特化しております。

(PP)

NCTRの研究部門ですが、6つ部門がございます。中でもかなり前から長期にある部もございまして、例えば、生物化学毒性部、神経毒性部、微生物部などは伝統的な毒性の管理部門ということになっておりますが、私どもの部門であります6番目、生物情報・生物統計部、こちらが最新の部となっております。

(PP)

そして、私どもの部門ですが、4年ちょっと前に設立されまして、現在は50人を超えるスタッフがいます。うち60%がサポートサービスに従事していきまして、残りの40%が研究に従事しています。

(PP)

本日は研究プロジェクトについてお話しいたしますが、私どもではいろいろな研究プロジェクトが走っております。そして、全て4つのカテゴリーに分類することができます。

高精度医療、ビッグデータ、希少疾患、そして、医薬品の安全性です。

まずは高精度医療ですが、私どもは高精度医療で使用される技術の評価を行っています。いろいろなコンソーシアムにおきまして、さまざまな文献を出しておりますが、30以上をもう既に出しております、うち3分の1をネイチャー・バイオ・テクノロジーで出しており、また、その論文、ペーパーに関しても、ネイチャー・バイオ・テクノロジーで2つ出しております。

特にアクティブに行っているプロジェクトが、この希少疾患に関するものになります。この会議の前に、日本における希少疾患ですとか各国の希少疾患について、その定義を話していたのですが、アメリカにおいては、特にこのFDAにおいては希少疾患を重視しております。既に7,000の希少疾患が特定されていますが、それに対して治療が提供できているのは500にすぎません。そして、FDAでは過去10年で3分の1の承認された医薬品のものが希少疾患対象のものとなっております。

次がビッグデータです。こちらはもう流行語になっていますね。ビッグデータと聞くと、皆さんは最初に思いつくのがデータの量だと思えます。かなりそのデータが増えてしまっている、データの移動ですとかバックアップが、今、問題になっています。ただ、私どもの部門ではビッグデータのサイズに関することに対処するよりも、ビッグデータからどういった知識を引き出し、それを規制にどのように応用していくかという手法を開発することを担当しております。

4つ目が医薬品の安全性になります。ここでも幾つかのプロジェクトが走っておりますが、最大のカテゴリーとして走っているプロジェクトとしては、薬物性の肝障害になります。

これについては詳しく説明することはいたしません。井上先生がラボの方にいらしていただいて、そこで色々話をいたしました、私どもが行っている肝毒性ナレッジベースプロジェクトについて、現在さまざまな努力をしております。次のページにありますネイチャーメソッドで出た記事ですが、こちらは時間のあるときにぜひ御一読ください。非常に公平な内容となっております、QSARのような予測モデルが規制面において、どうして使われるのか。また、どのように使われるのかということを目指している記事となっております。

(PP)

この肝毒性知識ベースですが、こちらは私どものグループで開発している1つの知識ベース、たくさんある中の1つになります。ほかにも肝腫瘍知識ベースですとか、たばこ成分知識ベース、内分泌かく乱知識ベースといったものも開発しております。この知識ベース、ナレッジベースの開発の仕方は非常にわかりやすいと思えます。まずはデータを収集し、収集したデータをベースにしてモデルを開発します。ほとんどの場合はQSARの手法を使っております。

(PP)

ご覧の通り、私どもの知識ベースは全てQSARのモデリング手法を使っております。つまりQSARのモデリング開発について、私どもは豊富な経験を有しているということです。次の2枚のスライドを使いまして、私どもが思う、よいQSARモデルの開発方法についてお話ししたいと思います。また、そのやり方と、特に規制応用面でどのように開発していくかということをお話ししたいと思います。

(PP)

最初に頑健なQSARモデルの管理方法ですが、やはりそのQSARにおいては大量のデータセットを使います。時にはあるもので我慢しなければいけない、選択肢がないことがあります。経験則によりますと、データセット量が多ければ多いほどモデルは頑健なものになります。

2番目ですが、生物学的活性データの質がとても重要になります。モデルを作るに当たって文献から情報をとるわけですが、この生物学的活性データの質がよくないと出てくるモデルも余り質のいいものにはなりません。ゴミ、がらくたを入れたら出てくるものはゴミであるというように、データ自体の質が悪かったら、モデルがそれ以上にいいクオリティーになるということは絶対にありません。

化学構造ですが、やはり多様性が重要です。時にこれは無視されてしまうのですが、ケミストリースペースとも言いますが、できるだけ多くのケミカルクラスをカバーすることで非常にいい予測モデルを出すことができます。つまり、いろいろなクラスの化学物質から予測をするということが重要で、1つのクラスでしかなければ、その予測自体も弱いものになってしまいます。

最後がモデリングのアプローチですが、モデリングのアプローチは本当に様々あります。すばらしいもの一つだけをつくれればいいというものではなく、とにかくたくさんモデルをつくり、それらを統合してコンセンサスの得たもので予測を行うことでクオリティーが高くなります。

(PP)

まず、よいモデルであることをどのように確認するかということですが、これらの全ての手順を踏む必要はありませんが、中には必要な検討すべき手順があります。まずはデータを2つに分類します。1つがトレーニングセット、もう一つがテストセットになります。

トレーニングセットをクロスバリデーションにかけます。このクロスバリデーションと言いますのは、少量のデータをとって、残りの化合物のデータに対してチャレンジするという形になります。このデータセットが内部的に一貫性があるかということ。これに基づいた予測モデルを立てていいかということを検証します。

2番目がパーミュテーションテスト、順列テストと訳されていますが、こちらはたまたまこのデータ分布でいい結果が出たのではないということを確認するということです。生物学的な活性でランダムにこのデータを分けまして、それを順列にしていって、本当のリアルなデータセットと同じパフォーマンスが出るかどうかを確認します。つまり偶発的

にいい結果が出たのではなく、きちんと順列にしたとしても、いい結果が出るようにするというので、それを検証いたします。常に独立的なバリデーションを行いまして、必要であれば文献に戻って、さらにデータを足します。ただ、全てのデータで全てのバリデーションができるわけではないということで、適応性の領域というところも非常に重要になります。

(PP)

では、これから1つのプロジェクトに特化して、お話を進めたいと思います。こちらが内分泌かく乱物質ナレッジベースに関するものです。私どもはこちらでもさまざまなQSARの手法を使って内分泌かく乱物質の予測を行いました。内容としては2つに分けておりまして、まずはどのようにモデルを開発し、なぜそのように開発したかというところと、2番目にはEPAが実際に入ってくるわけですが、EPAのほうでモデルの規制評価を行った、その内容について、お話をしたいと思います。

(PP)

まずは内分泌かく乱物質について、お話ししたいと思います。こちらはもう皆さんは御存じのとおり、昔から大きな問題となっております。20年前くらいでしょうか。1996年に米国議会で2つの法律が可決されました。こちらは食品添加物及び飲料水に認められる化学物質が人間の内分泌系をかく乱するということがわかりまして、この評価を要求する法律でした。その後、欧州及びアジアでも同様の規制が始まりました。EPAでは既に5万8,000の化学物質を同定しました。今この数が少し減っているようですが、これに対して実験的な試験をするということを規制しています。

ただ、1つの試験を行うのに大体24万ドルかかりますので、5万8,000となりますと、どれくらいのコストがかかるでしょうか。したがって、その試験の優先順位を決めることがとても重要です。どれをまず最優先で行って、どれについてはとりあえず後回しにしようということを決めることがとても重要となっております。この5万8,000の実験的な評価が必要なもののうち、化粧品原材料ですとか食品添加物、医薬品製品など、FDAが責任としているものもたくさん含まれております。

(PP)

まずEDKB、内分泌かく乱知識ベースの概要ですが、1996年にプロジェクトが発足いたしました。最初に行ったことはデータの収集です。しかし、すぐに文献データだけでは十分でないことがわかりました。先ほど私が御説明したQSARの基準に基づいた形での文献データでは、十分ではなかったということです。NCTRにおきましては、200を超える化学物質を4つのアッセイで試験を行いました。

まずはエストロゲン受容体結合アッセイ、アンドロゲン受容体結合アッセイ、 α -フェトプロテイン受容体結合アッセイ、SHBGの結合アッセイで、こちらの解析を行いました。これらは全てその内分泌かく乱に寄与するものになっておりますが、20年経ちましてデータベースの量もかなり膨らみ、現在での化学物質数は8,000を超えておりまして、また同時に

QSARモデル数も非常に膨大になっております。

(PP)

最近、私どもはこのQSARの全世界的な演習に参加することになりました。私どもはEPAの主催するエストロゲン受容体活性予測共同プロジェクトに参加したわけです。ご覧のとおり、21のチームが参加しておりますが、私どもは予測性能という意味では2位となっております。

(PP)

私どもはこの達成を非常に満足しております。私どものモデルの予測性能が非常にいい理由としては、やはりデータセットがカバーしている生物学的活性、ケミカルクラスが非常に幅広いからだと思います。

(PP)

こちらのスライドはご覧のとおり、エストロゲン受容体の結合につきまして、QSARモデルで使っているデータセットは非常に多様であるということがおわかりいただけると思います。ご覧のとおり、科学者でなくても、いろいろと異なるデータセットがあるということをおわかりいただけると思います。

(PP)

QSARのモデリング手法ですが、私のプレゼンテーションの中でもQSARについては本当に幅広く使っております。QSARというのは定量的構造活性相関ということで厳密な定義があるわけですが、したがって単に分類をする場合はSARと呼んでいます。便宜上このプレゼンテーションの中では一般的な用語としてQSARと使っております。

カテゴリーとしては4つございまして、QSARのモデリング開発では、最初にルールまたはフィルターづくりがあります。分子量が1,000を超えるものについては不活性であるといったようなもので、こちらは却下フィルターになります。

2番目が構造的特徴の有無になります。構造アラートとも呼びますが、構造特徴があるか、なしかで活性、不活性を判断するということで、そこに相関関係があるかを見ていきます。

3つ目が機械学習をするクラシフィケーションとなっております。モデリングについては非常にたくさんございしますが、こちらの観点で進めております。

次が分類モデルとなりますが、このグループの手法で化合物の予測を行い、活性か不活性かということを見ていくわけですが、こちらについては後ほど詳しく御紹介したいと思います。

最後がQSARになりますが、こちらでは定量的に予測を行うものとなっております。

(PP)

こちらは私どものほうで構築いたしました3Dの構造的アラートの一部となっております。こちらに出しましたのは、その中の2Dのものとなっております。特定の構造が含まれている場合には、化学物質が活性ありとみなすということになっております。こちらのほうに3

種類別々のものを出しておりますが、非常に私どものほうで広範に検証ができるということで、この3つのものを選んでおります。そして、このサブストラクチャーですけれども、こちらがER、エストロゲン結合部位とフィッティングするということで特定されています。(PP)

こちらのほうは2Dということですから、平面のレベルで化学物質の構造を見て、構造的アラートを出すというものになっておりますが、もちろん平面だけではなく立体的に構造を捉え、それに基づいてアラートを出す3Dの構造的アラートというものも可能であり、それがファーマコフォアと呼ばれるものとなっております。こちらが、私どものほうが構築した7つのクエリーとなっております。

(PP)

私どものほうで構築したモデル全てに対して、先ほど申し上げましたとおり、検証をかけております。まず、トレーニングセットにおけるモデルの性能について検証しておりますが、これには分類が含まれております。この分類テストを行うことによって全体としての予測性はどれくらいなのか、感度、特異性を検証しております。さらに交差検証も行うとともに外部検証も行っています。

(PP)

こちらのほうに示したのは、3DのQSARとなっております。すなわち定量的にどれくらいの活性があるのかという予測能を持ったモデルということになります。実験データの結果をこちらに示しておりますが、X軸のほうに実際のRBAが、Y軸のほうに結果として出てきた予測値が示されています。赤い点で示したのがトレーニングセットに基づいて出てきた結果、青い点で示されているのが検証用のデータを用いてモデルが解してきた結果をそれぞれ出しております。こちらのほうに出しておりますとおり、トレーニングデータセットには含まれていない検証用のテストデータセットを使ったとしても、モデルのほうが非常によい予測結果を介しているということがおわかりいただけるかと思えます。

(PP)

この結果、私どもが行った演習の教訓として、非常に重要な点として浮かび上がった点というのが、私どもが作り上げたモデル全てについて、どのモデルであったとしても一つだけをもって十分であるということと言えないということがわかったということです。これは冒頭申し上げたとおりで、万能薬的なモデルはないということです。したがって、ここで問題となってくるのは、私どもが構築したモデル全てを使って、どのような形で使えば最高の結果が出てくるのかということを考える使い方の問題だということになります。そのベストの結果を出すための手法というものがこちらのほうに出されており、最終的に4段階アプローチというものを構築することとなりました。この名前が示すとおり、4段階踏んでいくということになります。

最初の段階で却下フィルターというものを通し、ある化学物質について削除あるいは排除していくということになります。すなわち内分泌かく乱を引き起こすようなエストロゲ

ン受容体との結合能は一切、当該化学物質にはないということはこの第1段階で判断することになります。

そして、この第1段階を突破した化学物質につきましては、その次の第2段階に進むということになりますが、この第2段階で難しくなってくるのは、このプロセスを踏むためにどれくらいの数のモデルが必要なのかということを見極めるということになりますが、私どものほうでは、この第2段階で11のモデルを使用いたしました。内訳は構造アラートが3つ、フォーマコフォア、先ほど申し上げたものが7つ、デシジョンツリーモデルを1つ使っております。このやり方というのは、ここに挙げましたモデルのどれか1つでもある化学物質に活性ありとみなされた場合には、その化学物質については活性ありとみなすということです。全ての11のモデルに則って、ある化学物質が不活性であるということになりましたら、やはりそれは不活性であるという判断になりましたが、その理由につきましては後ほどさらに詳細を説明いたします。

この第2段階でも活性ありと判断された化学物質は、次の第3段階に進むことになり、これは3DのQSARを用いまして、定量的にどれくらいの活性なのかということ予測していく段階になります。最終的にはリスク評価ということになりますが、この段階になりますと、今まで出てきたデータだけでは足りず、例えば、ヒトに対する暴露レベルがどれくらいなのか、産生量はどれくらいなのかというほかの情報も考慮することが必要となり、それら全てを統合的に用いることによって、プラス生物学的活性に関するデータをもって最良のリスク評価上の判断ができるということになります。

(PP)

こちらを私どもでEPAのほうに提示したところ、ぜひぜひ、EPAが行うリスクアセスメント上、私どもNCTRのほうで構築したモデルがどんなふうに見えるのかどうかを見てみたいという高い関心が寄せられました。

(PP)

規制当局としてEPAがQSARモデルを用いて達成したい点というのは、こちらに挙げますとおり、2つありました。まず1点目に達成したいと考えていたのが、対象となる化学物質の数を減らすということです。実際にどの化学物質を検証すべきなのか、実験にかけるべき数を減らすということ。この工程を私どものほうではエンリッチメント、絞り込みと呼んでおります。絞り込みを行うことによって、後の段階で実際に検証、テストの対象とする化学物質の数を減らすというのが1つ目。

EPAがもう一つやりたかったのは、できるだけ偽陰性の確率を低くする、出てこないようにするというところにありました。陰性ではないのに偽陰性という結果が返ってしまうことによって、そのQSARモデルの結果、当該偽陰性と判断された化学物質が実際の実験というバリデーションの対象となることは二度とないからです。したがって、できるだけ偽陰性を低くするというのもう一つの要求事項でした。

(PP)

この2つのEPAの要求に対して、私どものほうで2つの戦略を立て、1つの目的に対して1つの戦略を適用するというのをやりました。まず1つ目の絞り込みに対して用いた戦略というのが、こちらに出しているスライドのとおりです。もちろん私どもが構築したモデルというものは完璧であるとは申し上げませんが、ただ、考慮すべき事項として見ていただければいいかなと思っております。

最初に、このエンリッチメント、絞り込みというこの目的のためにEPAのほうから私どものセンターのほうに対し、約6,600の化学物質の情報が提供されました。ただ、化学的な構造のみで一切の生物学的情報、データというものはありませんでした。この6,600のものを対象に私どものモデルを走らせて、その結果、内分泌かく乱能を持っているのではないかと疑われる活性があるだろうと可能性のあるものが319に絞り込まれました。

しかしながら、319とはいえ、これら全てに対し生物学的アッセイを行うだけのリソースはEPAにありません。そこでEPAのほうで、うち50を選択し、この50ほどの化学物質に対して実際の実験が行われました。その結果、この48のうち16については、実験でもって実際に内分泌かく乱能あり、活性ありということが確認されました。したがって、当初の6,600から16ということですから6分の1になったということになります。

(PP)

そして、2つ目のEPAの要求事項であり、偽陰性を低くしたいという要求に対する2つ目の戦略というものがこちらになっております。先ほどと同じ化学物質群を用いました。まず6,600ほどの化学物質のうち、EPAのほうでランダムに200を抽出してもらいました。この200ほどの化学物質に対し、商用で出回っているものなのかどうか、水溶性なのかどうかという検証を行った結果、189が残りしました。この189に対して、私どものほうでバイオロジカルアッセイを行い、実際の生物学的アッセイの結果とコンピュータによる理論値の結果を比較することによって陰性予測性というものを示しております。その結果がこちらで97.2%となり、私どものほうのモデルを用いると陰性予測性が97.2%であるということがわかりました。

(PP)

このプロセスの中においては、実は私どものほうでは内製開発ツールと市販ツールと双方を用いております。したがって、両方を用いたことから、この異なる2つのアプローチの両方の長所、短所を比較することができたわけですが、最終的にどうして内製開発ツールのほうを使うという決定に至ったのか。それについて、この後の2～3枚のスライドを用いて説明いたします。

(PP)

もうお気づきかと思えますけれども、この4段階モデルにおいて、私どものほうではさまざまな市販のソフトウェアを用いましたが、処々問題がありました。まずコストが非常に高額であるということ。また、規制当局の方が使うには非常に難しいという難点もありました。かつ、これは市販されているソフトウェアですから、常に変わり続けていて、単

に新しいバージョンが出ましたという知らせしかなく、それに対するバリデーションがかけられないという問題もあります。したがって、GLPのような実際に使えるものとしてメンテをしていくことが難しいという問題がありました。

こういった処々の問題を克服して、私どもが最終的に構築したいと思ったのは、もちろん一番重要である正確性と再現性があるということは言うまでもありませんが、それに加えることは、規制当局という必ずしもQSARについてはきちんと教育訓練を受けていない、専門家でない方であっても、規制当局として独立して運用できるものでなくてはならない。EPAの方が実際にこのモデルを使って作業をして、かつ、そのモデルが返してきた結果というものがEPAの方でも理解できるようなものでなくてはならないということがプラスアルファとして必要になってくるということがわかりました。

したがって、私どものほうでモデルあるいはメソドロジーとして、まずできなくてはならないことというのは、実際のスタディーの中では定義できるものではありません。その結果、私どものほうで新たにメソドロジーとしてデシジョンフォレストというものを編み出すこととなりましたが、この新たな手法でありますデシジョンフォレストにつきましては、この後のスライドで説明いたします。

(PP)

さて、デシジョンフォレスト、森ですから、そのもととなっているのは一本一本の木ということで、長年にわたって、ほとんど全ての分野において使われてきたデシジョンツリーというものがたたき台としてあります。こちらのほうは例えば、雑誌会社が使う場合であれば、最初はその雑誌を読む方は男性なのか女性なのか、男性のほうの方が雑誌を買う確率が女性よりも高いということが最初のデシジョンとしてあり、その次に、では、男性が買うのであれば、その方の年齢層はどれくらいなのかということでデシジョンツリーが踏まれていきますが、このデシジョンツリーをたたき台として私どもが構築したのがデシジョンフォレストというものになります。

(PP)

さて、このデシジョンツリーにはフィードバックループがないという、一方の方向にしか進めないという問題がありました。かつ、もしかしたらデシジョンツリーを用いる中でノイズが導入されるかもしれないという可能性もありますから、それを最低限に抑えるために1本のツリーではなく、複数のツリーを組み合わせたデシジョンフォレストを使うということになったのです。この前提といたしましては、異なる複数のツリーを組み合わせで使うということが重要になってきます。かつ、複数のツリーを使うというだけでは足りません。もし複数のツリーが全て同一のものであったらば、それは組み合わせてもよくないということになりますので、できるだけ性質の異なる、できるだけ合致、一致点がないツリーを複数組み合わせで使うということがデシジョンフォレストの仮説、前提となっています。

(PP)

こちらは皆さん、目隠しをした方と象の例えを御存じかどうかはわかりません。もしかしたら、この例え話というのは西洋だけで使われていて、日本ではないのかもしれませんが、よくあるのは、7人の目隠しをした方々が1頭の象のさまざまな異なるところを触った結果、7人7様の異なる結論が導き出されたという例え話があります。このデシジョンフォレストはこの例え話の反対側、同じことなのですけれども、それを反対側から見たものを前提、仮定として使っているということになります。

すなわち、通常の寓話ですと、目隠しをした方が同じ象を触りながら、違うところを触ったがゆえに違う結論になってしまったというものですが、もし7人の目隠しをした方が象の異なるところをそれぞれ触ったのに7人が異口同音に、これは象だという結論を導き出したならば、それは象に違いないという、それがデシジョンフォレストというメソドロジーを支える理論となっています。

しかしながら、このゴールを満たすためには、2つの基準を満たす必要があります。1点目の基準というのが、今の例え話ですと、7人の目隠しをした方がお互いにやり取りをとってはならない。コミュニケーションをしてはならないということです。Eメールのテキストメッセージなどを送って、これは象ではないかというようなことを言っては困るわけですね。

2つ目の基準というものが、7人の目隠しをした方の専門知識レベルが同じレベルであるということです。ある方は余りよく理解してないのに、ある方だけが非常に見識が高いということになりますと、これも余りよくないということになりますので、この2つの基本的な原則が満たされて、初めてデシジョンフォレストがワークするということになります。このメソドロジーに関するテクニカルな詳細というものは、ここでは割愛させていただきますが、今、申しあげました基本的なところが満たされて、初めてデシジョンフォレストというメソドロジーが成立するのだということを御理解ください。

(PP)

簡単ではございますが、こちらがデシジョンフォレストを使った結果となっており、このフォレストでは7つの木、7つのデシジョンツリーが使われています。デシジョンツリー1だけを使った場合には、17の化学物質について5分類が行われておりますが、プラス4つのデシジョンツリーを加えることによって、その5分類の率が半分ほどに減ることが示されており、もしこの1つのフォレストの中でデシジョンツリー7つを使った場合には、5つのみが5分類されるという結果になったということが示されています。

(PP)

こちらのほうにデシジョンフォレストのまとめがありますが、これは割愛させていただきます、次のスライドをごらんください。

(PP)

ここまでのところで、デシジョンフォレストを使えば、全ての問題が解決できる。まさ

にこれだという決め打ちの解決策のようなプレゼンをしましたが、実はそうではありません。プレゼンを通じて何度も何度も繰り返し申し上げたいのは、これだけが唯一の解決策ではないということです。すなわちQSARというモデルは、それ自体に制約があるということを知っておかなくてはならないということです。それをお示しするためにアメリカで非常に有名な映画『ダーティハリー』の言葉を用いております。人は自分の限界を知らなければならぬと、この映画の中で言われておりますが、まさにQSARというモデルも自らの限界を知らなくてはならないということです。

(PP)

こちらのほうですけれども、このQSARのレベルについて評価するために2つの要素を盛り込んでデシジョンフォレストを確認しております。それがこちらのほうに示されておりますが、X軸のほうに信頼レベルが示されており、左に行けば行くほど低く、右側に行けば行くほど高い。そして、それぞれの信頼レベルに応じて、どれくらいの正確性を持ってモデルが挙動するのかということがY軸のほうに示されております。当然のことながら、信頼レベルが高ければ高いほど、モデルの予測の精度も高くなるということが示されています。

(PP)

しかし、このモデルですけれども、ケミストリースペースという中で定義されており、このケミストリースペースの中にはトレーニングドメインの中に入るものと、その範疇に入らないもの。すなわち新たな化学物質が登場した際には、これまで既知のトレーニングドメインの中に収まらないものがある。それに対して、どうやってよい予測を出していくのかということが非常に重要になってきます。その端的なよい例をこちらに示しております。

下のほうの下線、こちらのほうは232だけの化学物質のデータからなる非常に小さなデータセットを用いております。この小さなデータセットの場合には、一旦このトレーニングドメイン外のものが入ってきますと、なかなか精度を高く予測するということができず、結果が悪いということになっております。しかしながら、上の例で示しておりますとおり、データセットとして、より大きなものを使いますと、化学物質の中でトレーニングドメインの範疇に落ちていないものがあっても、予測精度が30%を超えると比較的いい結果が返ってくるということが示されています。

(PP)

最後にこのプレゼンの中で訴えたかった重要なかぎとなる事項をまとめております。最初に市販ツール対内製開発ツールですけれども、まず手始めとしては、やはり市販に入手できるツールのほうから検証すべきであると思います。こうすることによって、より迅速なスタートが切れるからです。かつ、その市販ツールのベンダーさんのほうから、素晴らしいカスタムサポートを得られる場合には、そこから随分多くのことを学ぶことができます。とは言いながら、それと並行して、そのツールに対しては皆さんの側からのコントロ

ールが効かないということも認識しておいてしかるべきですし、また、ベンダーという会社のほうでやりたいような形でツールが変えられるということも意識しておく必要があります。

かつ、この市販ツールについて変更がなされた場合には、どこかの時点で行き詰ってしまうだろうということも意識しておかなくてはなりませんので、個人的にはスタートとして市販ツールはいいかもしれませんが、どこかの時点でやはり内製のツールに切りかえる必要があるのではないかと考えています。少なくとも規制というコンテキストの中で、規制に対してモデルを常に使っていくというのであれば、どこかの時点で内製のツールを使っていくべきではないかと思えます。そのような内製のツールを使うことによって、GLPのようなやり方がとれます。その上では、やはり内部で開発したツールというものの重要性というものが増してくるのではないかと考えております。この理解に立って私どものほうでは、MOLD2という組織内で開発したツールとデシジョンフォレストを組み合わせるというのをやるようになっております。

もう一つ、これももう既に何回か申し上げたところですが、QSARというモデルが全ての問題を解決する万能薬であると考えてはならないということです。したがって、万能薬ではないという制約がどこにあるのかということも理解しつつ、自分たちとしてはそのモデルをどんなふうに使いたいのか。何の目的のためにどのようなアプローチで使っていくのかということに目を向けて使っていかななくてはなりません。このモデルに限界があるということも理解しておくということは、QSARというモデルがどれくらい正確度が高いものかということも理解するよりも、むしろ大事な点ではないかなと考えております。

最後に申し上げたいのは、一つ一つのモデルを単独で使った場合と、異なるタイプのモデルを組み合わせる使った場合であれば、コンビネーションということで組み合わせる使った場合のほうが、どのようなモデルであれ、単独の使用と比べて、ずっといい結果が返ってくるという点です。

以上です。ありがとうございます。(拍手)

○川村座長 Tong先生、どうもありがとうございました。とてもわかりやすく論理的にお話しいただいたので、どんどん引き込まれてしまいました。

ただいまの御講演に対して、いろいろと聞きたいことがいっぱいあるのではないかと思います。この中には英語の達人もたくさんいらっしゃると思いますが、御質問は、通訳の方もいらっしゃるのです、あるいは必ずしも英語使いでない人もいますので、原則日本語で御質問をいただければと思いますけれども、いかがでしょうか。どうぞ。

○赤堀専門委員 スライドの12のところ、どのようにして、いいモデルをつくるのかということについて御説明をいただいたと思います。その際に5項目を提示いただいたのですが、必ずしもエッセンシャルではないということも言われていたので、この中でエッセンシャルなものは何なのかということも聞きたいと思えます。恐らくOECDのQSARモデルの検証原則を御存じだと思うのです。

委員の皆さんの資料の I の 9 のところに書いてある原則です。原則 1「エンドポイントの定義」、原則 2「曖昧ではないアルゴリズムの定義」までは多分普通に問題はないのです。その後の原則 3「適用範囲の定義」、原則 4「適合性、頑健性、予測性の適切な評価」と原則 5「可能ならば、メカニズムに関する解釈」という観点でやはり重要なポイントは何かということについて、お話をいただけたらと思います。

○Dr. Weida Tong ありがとうございます。こちらにつきましては、エゴバなどといったものにも参加しておりますが、QSARにおいては最終的にこういったドキュメントも統合していくということになっております。ただ、分野のほうでもさまざまな交流ですとか深化もされていまして、私の意見ですと、特にこの12ページの1番、クロスバリデーションはとても重要です。このクロスバリデーションを正しくやることによって、この結果と外部バリデーションの結果が一致することになります。

2番目、順列テストですが、時にデータセットがバランスを介していることがあります。例えば100の化学物質があって全て活性であり、非活性が1つもないといった場合は100%正しくなってしまいます。しかし、だからと言って、このモデルが正しいかどうかはわかりません。OECDのほうの原則にこの順列テストが入っていないのはなぜかと感じることもあります。

あとはバリデーションでエンドポイント、こちらもとても重要です。こちらはOECDも共通しています。ただ、中には予測が非常に難しいエンドポイントもございますので、ごみを入れたら出てくるものはごみというのとおり、データの品質がとても重要であるということです。

5つ目の適用性範囲ですが、こちらにつきましては2001年のドラフトにも記載されているようにとても重要で、多くの予測手法はこちらをベースにしております。

また、さらに申し上げたいのですが、OECDにおきまして、グッドネス・オブ・フィットという定義がありますが、クロスバリデーション及び外部バリデーションのところで実際に何が必要かとなりますと、やはりどれも必要かなと思ってしまいますのですが、最後の適用性範囲のところを見ますと、エンドポイントによっては全部やらなければいけないといったものもあります。皮膚の刺激性といった場合はいろいろな選択肢がありますが、肝毒性といった場合には選択肢が非常に狭いというところで、特定のエンドポイントに使ったモデルは完璧ではないというところです。したがって、グッドネス・オブ・フィットを知るよりも、そのモデルの限界を知るのほうが重要です。

○広瀬専門委員 グッドネス・オブ・フィットを少し説明していただければ。

○赤堀専門委員 グッドネス・オブ・フィットというのは、モデルをつくったときにどれくらい合っているかということ表現しているものになります。モデルをつくって、先ほど、Weida先生が見せてくれた直線回帰をしたときに、モデルの近くにいっぱい点があったら、もっともらしいモデルです。それを定量的に示しているのが決定係数等 R^2 で幾つという、それが例えば限りなく1に近ければ近いほど、定量的に予測できているというようなも

のになります。

ついでに質問をしてよろしいでしょうか。今の12ページの5つの項目は外部バリデーションに関して、アディショナル・バリデーションセットということで外部検証を2つ挙げているのですが、外部検証は場合によってはできないことがあるので、なかったら実施しなくてよいということでしょうか。それとも、原則としては欲しい。トレーニングセットをつくるときにできるだけExternal（外部データ）もつくるように頑張ろう、ということでしょうか。

○Dr. Weida Tong そうですね。できる限り集めていただくのが重要かと思います。例えば100の化学物質といったデータセットがあり、それに基づいてQSARをつくるとした場合、100のうち、これに対していろいろ行うことができるわけです。80をトレーニングセットとして、つまり20をとってしまうわけですが、この80に対して、いろいろと好きなことをやって、それに基づいて20を予測するという形をした場合、まずこの80と20に分けること自体バイアスがかかっていますので、80と20で分けたときの考察がその現実をあらわしていない可能性もあるわけです。したがって、1回だけではなく、統計学的に数百回、数千回を行わなければならないかもしれません。ただ、化学物質の数が1万だった場合は、80と20に分けた場合のバイアスは余り、先ほどよりはリスクはないと思われます。したがって、やはりここは1つのルールといったものではなく、一般則に準じた形で行うことになるかと思われます。

先ほどの100の化学物質に戻りますが、これが終わって、また文献に戻りまして、ほかにも何か物質があれば、それを見ていくということになりますが、例えば、薬剤性の薬物性の肝障害といったものがあつた場合、これが本当に医薬品による障害かどうかというところがなかなか難しく、ノイズがたくさんあり、完璧なものが難しいのですが、こういったものについてはQSARのトレーニングセットではなく、外部バリデーションのデータとして使う形になります。

○川村座長 ほかにいかがでしょうか。どうぞ。

○岡田専門委員 4フェーズ・アプローチで採用しておられる3種の警告構造と7種のファーマコフォアですけれども、これらはどうやって見つけられたのですか。それとも、これはその分野では常識なのでしょう。

○Dr. Weida Tong こちらにお示ししました7つのファーマコフォアについての質問かと思いますが、実はこれはもう市販用のソフトウェアがプログラム化されたものとして使える形になっております。私どもで検証したのは230の化学物質ですけれども、これらについては、どれがエストロゲン受容体と結合するのかわかるのか、するものとしらないものと、もう既にわかっておりました。これらの230の物質をソフトウェアのほうに入れて、そうするとソフトウェアのほうは3D、立体的にどう構造的な違いがあるのかということを出してきます。したがって、それを見れば、どれが統計的にロバストなものなのか、堅牢なものなのかということがわかります。

○岡田専門委員 もう一つ、デシジョンツリーをつくる時に使った記述子、ディスクリプターは普通の商用のソフトウェアの作成したものを利用されたのですか。

○Dr. Weida Tong 実は今、一番最後のスライドで出しましたところに記載されているMOLD2という、こちらは私どものNCTRの内部で開発されたツールとなっておりますが、これを使っております。したがって、私どもが用いた記述子というのは市販のものではなく、内製のものということになります。このMOLD2はファーマのウェブサイトのほうに行ってくださいますと公に公開されておりますので、ごらんいただけるかと思えます。

このMOLD2の中に、ラスベガスではありませんが、験を担いだわけではないのですが、777ということでラッキーセブン、スリーセブンのディスクリプターが入っております。というのも、私どものほうで毒性を検証していた際に、こと安全性に関する評価項目については3Dも2Dも同じほど、よいレベルであるということがわかりました。というのも安全性の評価項目についてはオフターゲット相互作用というものがないからです。しかしながら、有効性の評価項目ということになりますと、これは3Dのほうがずっといいということで、2Dは用いません。いずれにいたしましても、2Dの記述子として、私どものほうで内製で開発いたしました777の記述子を用いまして、安全性の評価に用いています。

○川村座長 ありがとうございます。

あつという間にどんどん時間が過ぎておりましたので、最後に山添先生、お願いします。

○山添委員 簡単な質問を2つなのですが、1つはエストロゲンの受容体のアフィニティーをNCTRで直接測っていらっしゃいますね。それはパブリケーションのリポートであると、親和性について非常に大きなばらつきがあるから、ベースのものについては自分のところできちんとやるという考え方ということでしょうか。

もう一つは、今回お話くださったERのようなもの場合には、1つの受容体に対する親和性として一応評価が可能なものですよ。ところが最初のところで説明をいただいた、例えば肝毒性のようなデータベースの場合には、標的が非常にたくさんの機序から起きてくる可能性がありますよね。そういう場合についても同じような考え方で行くのか。それとも、それなりのデバイスを考えなければいけないのか。その辺についてはいかがでしょうか。

○Dr. Weida Tong まず最初の御質問に関してですが、答えはイエスになります。1996年にプロジェクトをスタートした際に、やはりいろいろなラボから情報、データを取り寄せましたが、1カ所は10しかない、2カ所は20しかないということで本当にデータ量が少なかつたということで、こちらはゴールドスタンダードを使いました。EPAではセルラインを使っているようですが、こちらを使って私どものほうで親和性の測定を行いました。

肝毒性についての御質問ですが、こちらのナレッジベースについて私が説明をするとすると、次のスピーカーの方に怒られるぐらい長々と話してしまいたくなりますが、肝毒性に関しては同じ手法は使っていません。この場合は人の母集団のほうに吸収されたところ

から、またケミカルな予測に戻るといった形をとっています。エストロゲンの受容体の結合については非常にわかりやすく、また、モデルの品質も非常に高くなっていますが、肝毒性は本当に難しいです。おっしゃったとおり、機序がたくさんありまして、エンドポイントもたくさんあるということで、OECDの原則にあるとおり、あるエンドポイントに対してのモデルという形で、どうしても制限されてしまいます。したがって、化学物質の数も重要ですし、化学物質の数をふやせば、堅牢性も増しますが、それが正しいかどうかの測定もきちんとしなければいけません、やはり予測モデルの限界を知ることがとても重要です。

○川村座長 ありがとうございます。

○吉田委員 よろしいですか。済みません、ありがとうございました。EDCについてお伺いします。EDCの定義は内分泌機能をかく乱するものであって、必ずしもERに結びつくものだけではないと定義されていたと思いますが、ERに結びつかないもの以外のQSARについては何か新しい方法で解決法をお考えなのでしょうか。

○Dr. Weida Tong 確かにこの内分泌かく乱というものに対しては非常に複雑な機序が伴っております。したがって、アンドロゲン受容体結合だけではなく、エストロゲン受容体結合のほうにつきましても両方検証されてきましたが、当初はエストロゲン受容体結合のほうに有意であるということで、ERの中でもサブタイプである α と β に注力した研究というものがなされてきました。そのうちにそれだけではならず、 γ というサブタイプまで議論されるようになりましたが、こちらは γ のほうは最終的にそれほど重要性がないだろうということで落とされております。

いずれにしても、この内分泌かく乱に非常に複雑な機序がかかわっているということは事実です。それが理由となって、少なくとも1996年～2006年の10年にかけては、こちらのほうはこの10年においては、やはりこの機序においてはエストロゲン受容体の結合が有意であるという理由もあって、こちらのERの結合のみがドミナントのメカニズムとして検証されてきました。しかしながら、それ以降、たくさんの研究がなされ、プロゲステロン受容体結合を含めたAR結合のほうにつきましても検証がなされるようになっていますが、リスク評価という観点から見ますと、これら活性を持つものを全て含んで検証することが必要になっており、実際にNCTRのほうでも現在進めている研究の一つとして挙げられております。

随分時間をとりそうで恐縮ですけれども、この等式、方程式を考える最も重要な部分というのは、血清タンパク結合にあるのではないかと個人的には思っております。先ほど申し上げましたAFP及びSHBGも見ていくということが非常に重要です。というのも、一旦、血中に入りますと、AFP等の影響によって実際のこのERに対するセルラインの結合というのは低くなってしまふということが考えられるからです。したがって、*in vitro*と*in vivo*の結果の相関ということを考える上では、もちろんほかのレセプターバイディング、ほかの受容体結合もあるのですが、それよりも個人的には血清タンパク結合のほうはずっ

と重要ではないかなと考えております。

したがって、今までは有意な機序として捉えられていたエストロゲン受容体及びアンドロゲン受容体結合、これに加えてAFP、SHBGと一緒に見ていくということが必要だと考えられており、実際にそれに関する研究調査というものが私どものセンターのほうで進んでいます。

○川村座長 ありがとうございます。

それでは、Tong先生のすばらしいお話はここで一旦終わらせていただきまして、次のプレゼンテーションのほうに移りたいと思います。コンピュータのセッティングをつけかえますので、少し時間をいただきます。

それでは、山田隆志専門委員の御発表に移りますが、「国内外の反復投与毒性に関するデータベース・毒性予測ツールの現状について」というタイトルであります。それでは、山田先生、お願いいたします。

○山田専門委員 御紹介ありがとうございます。では、引き続き、「国内外の反復投与毒性に関するデータベース・毒性予測ツールの現状について」ということで紹介させていただきます。

まず、反復投与毒性試験ですが、これは動物に被験物質を一定期間、通常は4～13週間ですけれども、反復投与したときにあらわれる生体の機能及び形態の変化を観察することによって、被験物質の毒性を明らかにすることを目的にしています。

通常、ラットに複数の投与量で被験物質を投与しまして、種々の検査を実施して、毒性のプロファイルとNOAEL、無毒性量を求めます。1つの物質の試験をするのに、おおよそ700万円～1,000万円程度の費用を要すると言われています。

(PP)

私たちは、反復投与毒性の*in silico*評価を目指す取り組みを行ってきました。その中心となるHESSについて紹介いたします。その後、関連する海外のデータベース、ツールを紹介したいと思います。

(PP)

これは2005～2006年ごろから構想を始めまして、構造活性相関手法による有害性評価手法開発という事業名に決まり、2007年度から始まりました。プロジェクトリーダーはその当時、国立衛研にいらっしゃった林真先生でして、参加機関は国立衛研、製品評価技術基盤機構(NITE)、ブルガス大学、関西学院大学、富士通、東北大学でした。林先生と、当時、NITEで技術顧問をしておりました山田隼さんの2人が中心となりまして、目的を共有できる専門家グループをつくってスタートをしました。さらに、ここには名前を載せられなかったのですけれども、数多くの人々にも手伝っていただきました。

(PP)

開発方針ですけれども、まず専門家の判断をサポートするためのシステムを構築する。システムが判断を下すのではない。次に、毒性、病理、総合的なリスク評価を行う専門家

が主導的な役割を果たす。システムの専門家が主導ではないということ。そして、国際的に利用されるものを目指す。OECD QSAR Toolboxの開発者をメンバーに加えて、OECDと連携しながら研究開発を実施いたしました。

(PP)

これは開発システムの構成になります。試験報告書から毒性データベースをつくりました。これが基本となります。さらにその毒性の作用機序に関するデータベースをつくりまして、なぜその物質は毒性が出るのかということ考察できるようにしました。さらに毒性は化学物質が体内に吸収された後、分布して代謝されて、毒性が発現する場合があります。そしてその代謝には種差があります。そこで代謝のデータベースをつくりました。これらのデータをもとにして、カテゴリーアプローチ支援機能とベイジアンネットという2つの毒性予測モデルを開発しました。ユーザーが評価したい物質の情報を入力しますと、そのシステムは関連物質の毒性情報やメカニズム、代謝等の情報を提供します。これらのデータを用いてエキスパートジャッジを行って未試験物質の反復投与毒性を予測する、このような仕組みになっています。

(PP)

これは反復投与毒性を予測するためのリードアクロスと言われる類推のワークフローを示したものになります。まず対象物質がありまして、毒性データがない場合、試験済みの構造類似物質をデータベースから集めてきます。そして、さらに同じメカニズムを持つと考えられるであろう物質をグループ化してカテゴリーにします。そして適切な類似物質から対象物質の主要な毒性とNOELを推測していくということになります。HESSはこの類似物質の情報、代謝やメカニズム、信頼性の高い毒性データを提供することを目指しています。

(PP)

最初に紹介するのは毒性の知識情報データベースというものでして、一番重視をしたのはデータセットの質です。日本の化学物質審査規制法の既存化学物質点検『28日間反復投与毒性試験報告書』を中心に、GLP基準の試験報告書を採用しました。用量反応データを収載し、毒性・病理の専門家の意見、化学物質審議会の判定に基づくような有意差、フラグを付与しまして、毒性学的な注意喚起を図りました。

さらに、ここには載せませんが、試験に関連する情報としまして、試験物質の純度でありますとか、用量設定の根拠でありますとか、オリジナルの文献情報を付与しまして、必要があれば、もとの文献にまでたどれるようにしました。

(PP)

もう一つ、病理所見用語の統一を行いました。データベースをつくりに当たりましては、試験報告書の記載どおりに所見を抽出したのですが、このままではデータベースの検索はしにくい。そこで同義の病理所見をグループとして整理しました。これによって1つの所見を持つ物質をデータベースから探していただくことができます。

(PP)

次に、毒性作用機序のデータベースになります。試験報告書を精査しまして、変性、壊死など重篤な毒性のメカニズムの文献情報を収載しました。

(PP)

これは幾つかの毒性のパスウェイを整理したものになるのですが、このようなパスウェイに当てはまりそうな物質をいかに広げて見つけていくかということが、この予測においては重要になってくるだろうと考えています。

(PP)

さらに代謝マップのデータベースについてです。ラットの代謝マップデータベースは毒性試験が行われている物質、またはそれと構造が類似した物質の代謝マップを収載しています。現在のところは1,200マップが入ってまして、化学物質の代謝マップの数としてはかなり大きなものとなっています。さらにその代謝データをもとにして代謝シミュレータを備えています。

(PP)

もう一つデータベースがあります。このヒト/ラット代謝データベースにつきましては、吸収、分布、代謝、排泄の情報を収集しています。このデータベースのユニークなところは、主要な薬物代謝酵素でありますシトクロムP450の非常に精度のよい代謝予測モデルの結果を収載した点です。代謝データが全くないような物質につきましても、想定される代謝物の反応部位であるとか、その量的な順位を予測して、その結果を参照することができます。

(PP)

次のスライドに行きます。これとは別に物質間で毒性のプロファイルを比較できるような新たなデータベースをつくりました。各所見ごとに有意な変化があらわれた投与量をLOEL、その一つ下をNOELという形に整理しまして、それをデータベース化しました。

(PP)

そのデータベースの構造ですけれども、一般状態から病理学試験まで約500の所見から構成されています。

(PP)

続きまして、このMode of Action/AOPをベースにしました反復投与毒性のカテゴリーを構築しました。代謝、毒性作用機序、試験報告書のデータを用いまして化学構造と代謝毒性メカニズム、毒性試験で観察される所見のセットを関連づけてパスウェイを構築しました。

(PP)

これは1つの例になるのですが、このようにまず毒性パスウェイを整理し、続きまして、データベースからこのパスウェイを活性化するであろう物質を集めてきて、その構造領域を定義し、システムに導入しました。

(PP)

もう一つ、ベイジアンネット毒性予測モデルというものもつくりました。これはベシク・アクティブ・ストラクチャーという部分化学構造と毒性のさまざまな事象の因果関係を方向付きのグラフと確率であらわしたものになります。これは毒性のデータベースと、さらに専門家の暗黙知などを加えて構築いたしました。

(PP)

これが開発したシステムのワンショットになります。対象物質を入れますと、この場合、対象物質には毒性データはありませんでした。代わりにこのシステムは候補となる毒性カテゴリーを示します。そして、データベースから、その類似の物質を集めてきて、その試験データをここに表示することができます。このデータを用いて、対象物質の毒性を予測するということになります。この判断をするに当たって関連情報を参照することができます。

(PP)

代謝やメカニズムに関する情報を参照し、適切な類似物質を選んで、その試験データを用いて類推を行う。このデータの信頼性というのも確認をしながら類推を行うということになります。

(PP)

実際には、このシステムを用いながら、幾つかケーススタディーを行ったのですけれども、そのうちの1つを紹介します。これは香料ですけれども、この香料は体内で代謝を受けまして、肝臓において、ある一定の暴露量以上になりますと、肝毒性を発現するということがわかっています。ここには幾つかの類似した物質があります。そのうち、赤で囲った物質、これが毒性試験の結果がある物質で、それ以外のものは毒性データがなかったことを示しています。これらの代謝、メカニズム、毒性のデータを統合して、最終的にこれら一群の物質は肝毒性が発現するであろうと予測しました。その毒性のレベルは最も毒性の強い物質（ワーストケース・アナログ）を用いて、安全サイドに立ってNOELを推測いたしました。

こういったアプローチというのは、これまでも毒性の専門家ならば恐らく頭の中でやっていたのではないかと思うのですけれども、ここでは私たちの開発したツールと追加の文献検索を行いまして、科学的なエビデンスに基づいて、なぜこのような結論が得られたのか。さらにこの適用範囲はどこまでなのか。そして、構造の違いが違った毒性を引き起こす可能性がないか。こういったことを文章にあらわして、行政的に利用できるようにすることを目指しています。OECD IATA Case Studies Projectでは、こういったアプローチを活用促進するためのテクニカルなガイダンスを現在、開発中です。

(PP)

できたこのシステムにつきましても、データ更新と共有を行っていました。最初にOECDのQSAR Toolboxにデータカテゴリーを提供すると同時に、アメリカのUSEPAのToxRefデータベースと欧州のCOSMOSデータベースとデータ交換を行いました。また、厚労省の毒性試

験の結果、OECDの高生産量化学物質の毒性試験の結果等も追加するとともに、厚労省のToxicogenomics Project、経産省のToxOmicsといった研究開発プロジェクトで得られた毒性データも提供いただいてデータ更新を行っています。

(PP)

ここからは海外のデータベースについて紹介したいと思います。このRepDoseデータベースは、ドイツのフラウンホーファー研究所が開発したものでして、工業化学物質の毒性データが入っています。物質名とか動物種、雌雄、経路、器官、標的臓器、影響、NOEL、LOELが参照できるようになっています。継続的にデータ更新が行われていまして、ウェブ上で参照することができます。

(PP)

次にありますのはOECD QSAR Toolboxです。これはカテゴリーアプローチによりデータギャップ補完を支援するためのシステムです。種々のエンドポイントの実測データやカテゴリー化のための支援ツールが搭載されています。

(PP)

各国の行政機関や研究所などから化学物質の規制に関連したエンドポイントの試験データが提供されています。細菌の変異原性、発がん性、そのほかに生殖発生毒性等々、いろいろとあります。反復投与毒性に関してはRepDoseとHESSの2つのデータベースが入っています。このツールの1つの特徴は、1つの物質について複数の毒性エンドポイントを同時に解析することができるということになるかと思います。

(PP)

もう一つの特徴としまして、カテゴリープロファイラーを搭載していることが挙げられます。このクラスの物質はこういうメカニズムで、このような毒性を示すといったような既知見をこのようにリスト化していまして、評価したい物質がこのようなメカニズムで毒性を引き起こしそうかどうかを考える、いわゆるエキスパートジャッジをサポートするものになります。

(PP)

さらにデータのインポート、エクスポートができます。ユーザー独自の毒性のデータベースとか、あるいは毒性のカテゴリー、こういったものをユーザー自身が搭載することができます。そして、このシステムをカスタマイズすることができます。企業とか行政機関などで外に出せないインハウスのデータを既存のデータとあわせて解析することができるようになっています。さらに公開データにつきましては、このデータをエクスポートして他の用途に使うこともできるようになっています。

(PP)

もう一つはCOSMOSというものでして、これはSEURAT-1という欧州の反復投与毒性試験の代替法開発のプロジェクトになります。ここにCOSMeticsと書いてありますように、主に化粧品原材料などを対象にしたデータベースと*in silico*モデルを搭載しています。

(PP)

データベースとしましては、毒性の所見やNOEL、LOELといった情報、コンピューショナルモデルとしましてはTTC（毒性学的閾値）に関するモデルが入っております。

(PP)

もう一つ、eToxというシステムがあります。これは主に欧州の製薬企業や研究所がコンソーシアムをつくって開発したシステムでありまして、医薬品やその関連物質の毒性データベースと予測モデルを搭載するシステムになっています。

(PP)

各社の過去の遺産であります試験レポートを持ち寄ってデータベースの形に規格化しまして、データを共有している。そして、さらにこの集約したデータを用いて予測モデルをつくっています。プロジェクトの参加企業の中でデータと予測モデルが共有されていると聞いています。

(PP)

駆け足でしたけれども、このような形で、国内外でここ数年、いろいろなデータベースと毒性予測ツールが使えるようになってきたという状況をお話しました。

最後に、こういったものが今後どういうところで使えるのかということについて考えてみたいと思います。これは専門委員の先生方といろいろ話し合いながら、みんなで知恵を出していくべきものだと思うのですが、一つ考えられるのは、リスク評価の優先順位づけといったところ、やはり危なそうな物質から順に評価していく。危なそうな物質を漏らさないようにするということがあろうかと思います。もう一つとしましては、微量で多種類の化学物質を短期間で包括的に評価する必要があるような場合というのにも有用ではないか。毒性試験が困難な場合、不純物とか分解物とか代謝物とかデータがない場合の補完に使えないかということも挙げられます。さらにメカニズムに基づく人への外挿といった、リスク評価の高度化があると考えられます。

残念ながら、今のところ完璧なツールというのはありません。やはりこういったことをするには複数を組み合わせると同時に、追加でいろいろな情報を収集して、あわせて使っていく必要があるのではないかと考えています。さらに、国際的な動向を把握しながら、範囲を指定して活用経験を積んでいくということが重要ではないかと考えています。毒性データベースは更新していかなければいけないわけですが、では、どこまで更新をしていったらいいのかと。大きな目標を立てると、いつまでたってもデータが足りないの使えないというふうになってしまう。そうすると、予算を出す側も出し続けられないといったようなことがあろうかと思います。また、一体誰が評価の中に取り組んでいくのかといったような問題もあろうかと思います。

そういったことを考えますと、まず小さな問題からスタートして、その限定された範囲の中で使ってみる。検証し、足りないところがあれば補強する。それが終われば、また次のステップへ行く。そうすることによってツールの改良も進みますし、専門家もその活用

経験を積み重ねていくことができる。このようなアプローチによって、新しい手法をリスク評価に少しずつ取り入れていくことができないかと考えているところです。

以上になります。ありがとうございました。(拍手)

○川村座長 どうもありがとうございました。

ただいまの山田専門委員のプレゼンテーションについて、御質問あるいは自分の意見を率直にぶつけて、我々の知識の共有化をしたいと思えますけれども、いかがですか。

○赤堀専門委員 プレゼンテーションをありがとうございました。個人的には一般毒性を予測するには、多分リードアクロスしかないのかなと。QSARモデルをつくるよりもリードアクロスをしてやっていくことのほうが多分リーズナブルだろうし、今まで食品の評価においても、これに近いアプローチをされてきたので、これをやっていくことは適切かなと思っているのが、まず自分の感想です。

例えば、今、御紹介のあったデータベースの中に食品安全委員会で評価する対象の物質がどれだけ含まれているのか。それができないことにはリードアクロスそのものがないと思うので、食品安全委員会が評価しようとしている物質のケミカルスペースの調査をし、今、既存のデータベースでどの程度カバレッジがあるのかを見ていくことも興味深いかなという印象がありました。

さて、質問になります。先ほどOECDのIATAでリードアクロスをする際にどのように実施したらいいのかというようなガイダンスを作成しているというお話があったのですが、簡単にOECDのIATAのアクティビティーについてお話を伺いたいと思います。また今現在、HESSの中で一般毒性に関するAOPが61ほどあったと思うのですが、これからそれを拡張していく予定があるのでしょうか。この2点について、お話を伺えたらなと思います。

○山田専門委員 OECD IATA Case Studies Project、IATAというのは試験と評価のための統合的アプローチ、日本語で言うと多分そうなると思うのですが、種々のデータ、毒性試験データだけではなくて、代謝やメカニズムなどを組み合わせて毒性を評価していく。あるいは必要ならば、さらにその追加の試験を要求する。そういったアプローチです。いろいろな新しい技術を導入しながら毒性の評価をしていくことも想定しています。概念としてはあるのですが、具体的にどうしたらいいのかというのはなかなかわかりにくいということもあって、OECDではIATAのケーススタディーを参加国から募って、それを各国の専門家によってレビューをし、こういうところが足りないのではないかと、こういうことを説明すべきではないかと、といったことをみんなで考えを出し合っています。その結果に基づいて最終的にはガイダンスのようなものをつくって、こういうところには注意をしなければいけないのではないかと、こういう内容は入れなければいけないのではないかと、それをガイダンスという形で整理をしていこうとしている。そういう試みです。

○赤堀専門委員 何のガイダンスになるのでしょうか。IATAのつくり方、それとも例えば、毒性ごとのIATAのガイダンスでしょうか。

○山田専門委員 その辺はまだ細かいところは決まっていますが、基本的に各毒

性ではなくて、もうちょっと大きくやり方のようなもの、例えば類似物質の選び方や不確実性の評価など注意すべきところとかをまとめた、そういう指針になります。

○赤堀専門委員 AOPをつくられていますよね。今後も拡張していく予定がありますか。

○山田専門委員 それはそういうリソースがあればということになります。それから、新しいエビデンスが出てくるかどうかということがあろうかと思えます。やはりリソースが必要なのですけれども、OECDなどでは別にAOPプログラムとって毒性のメカニズムの情報をみんなで持ち寄って、それを使えるようにしていきましょうというプログラムも走っていますので、そういった形で、全部自前でやるのではなく、国際的に協力しながら、毒性のメカニズムの情報を整理し、使えるような段階に近づいてきていると思います。そういった仕組みをうまく使っていけば、少ないコストでAOPを拡張していける可能性はあると思っています。

○赤堀専門委員 ありがとうございます。

○川村座長 実際にこれはメンテナンスというか、システムの維持とデータベースの更新などは、どなたがどういうふうになら何のファンドでやっているのですか。

○山田専門委員 今これはNITEで所有しています。テクニカルなエキスパートにデータを入れていただいています。化審法の一つの運用支援の枠組みの中でデータの更新を行っています。

○川村座長 誰かマネージというか、ステアリングコミッティーみたいな組織としては、この開発者のグループがそのまま継続しているのか、あるいは新たな仕組みがつけられているのか。

○山田専門委員 基本的には所有者でありますNITEと、そのパートナーであります国立衛研とが協力してやっていくという覚書を結んで、一緒にやっています。

○川村座長 続けて質問で申しわけないです。例えば、この食品安全委員会なども一つのユーザーになり得ると思うのですが、そういう利用者から何か注文をするとか、あるいは逆にデータを提供するとか、そういう仕組みはあるのでしょうか。

○山田専門委員 それは赤堀専門委員もかかわっておられましたけれども、実際に毒性データを提供いただいた事例はございます。これまでは個人的な人のつながりや、学会発表等でできてきた人間関係、その中でデータをいただいたことはありました。

○川村座長 いかがでしょうか。御質問はよろしいですか。

○赤堀専門委員 多分リードアクロスで重要なことは、グループにしていいかどうかの判断だと思います。これが一番議論になるところだと思いますが、AOPだけでグルーピングをするわけではないと思うのです。どうやってグルーピングをするかというところについてガイダンス的なものの整備状況とか何かがわかれば教えてください。

○山田専門委員 それは一番難しいところで、今のところケース・バイ・ケースだろうと思います。ひょっとするとエンドポイントによっても違うかもしれないと思っています。これからその辺をどう整理していくかが議論されるだろうと思います。

○赤堀専門委員 ありがとうございます。

○川村座長 山田専門委員はメンバーですので、今後ワーキンググループの会議の中でもいろいろと意見を出していきたいと思います。きょうは時間がなかなかタイトでありましたので、議論の時間がやや詰まった感じはするのですけれども、きょう用意した内容については、ここで一区切りをつけたいと思います。

今後に向けて、このワーキンググループとして、次回以降、きょうは勉強会的なものがありますけれども、どういうふうな議論を進めていくか、あるいは審議の仕方をするかということについて、何か御意見はありますでしょうか。専門委員の立場として、あるいは親委員会からの御注文として何かありますでしょうか。

○広瀬専門委員 きょうの目的から次に進まないといけないと思うので、でも、Weida先生の最後のスライドもそうですし、山田先生のスライドもそうなのですが、やはり適用範囲をある程度限って、それを具体的に使っていくという、それをさらに食品安全委員会で使えるマターという感じで、多分、概論の話ばかりをやっても前に進まないと思うので、できるだけ具体的な、食品安全委員会の中でレギュレーションに使う、使わないは別にして、とりあえず何かでやってみるというのを試行するべきではないかと思います。

直近で思いつくのは、食品安全委員会ではまだやっていませんけれども、微量成分だと香料だったり、器具・容器の溶出物質だったりしますので、先ほど赤堀専門委員から食品安全委員会でやった物質マターのカバー率と言われると、多分、HESSは食品関連の化学物質のカバー率はちょっと悪いのですけれども、例えば、器具・容器だとほとんどが工業化学物質なのでカバー率という意味では高くなりそうなので、別に私はそれを推薦するわけではないのですけれども、一番手近で取りつきやすいところなのかなと思います。

○川村座長 わかりました。ありがとうございます。食品安全委員会は扱うものも非常に幅広いのですけれども、全てをカバーし得るものでもないですし、差し迫って答申といいますか、評価をしていかないといけない側面もありますので、実際に使う価値があるであろうものから、どんどん使ってみて、経験と試行を重ねていくような形で進めていくべきではないかという御意見をいただきました。

ほかにはよろしいでしょうか。まだまだ手探りのところがありますので、今後、随時また御意見をいただきますし、私も事務局としょっちゅう、いろいろと議論はしていきますので、その中で進め方、何か光が出てきたかどうかというようなことは進めてまいりたいと思います。時間が大変タイトの中でワーキンググループの審議を進めまして、手際が悪かった点はお詫びいたしますけれども、このように知識や認識の共有からスタートさせていただきました。

それでは、議事としてはここまでにさせていただきます、事務局にお返しします。

○井上課長補佐 本日御審議いただきました意見に基づき、次回の準備を進めたいと思います。ワーキンググループの最後に近いところで恐れ入りますが、公務で遅れておりました事務局長の川島から先生方に御挨拶申し上げます。

○川島事務局長 川島でございます。遅れて参りまして、大変申し訳ございませんでした。途中からですが、拝見させていただきまして、新しい評価技術がいろいろと国際的にも開発されてくるという中で、食品安全委員会としてもこういうものに取り組んでいく必要性、重要性というものを改めて認識したところでございます。ワーキンググループの先生方には、これからも大変お世話になると思いますけれども、よろしくお願ひします。

○井上課長補佐 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いたします。また、専門委員の先生方におかれましては事務連絡がありますので、このままお待ちいただきたいと思います。よろしくお願ひいたします。

以上です。

○川村座長 それでは、本日の議事は全て終了いたしました。

これをもちまして、第2回のワーキンググループのミーティングを終わります。どうもありがとうございました。