

食品安全委員会評価技術企画ワーキンググループ

(第11回) 議事録

1. 日時 平成30年8月30日(木) 10:00～12:12

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) 食品健康影響評価技術研究について

(2) 食品健康影響評価におけるベンチマークドーズ法の更なる活用について

4. 出席者

(専門委員)

川村座長、赤堀専門委員、小坂専門委員、小関専門委員、
西浦専門委員、広瀬専門委員

(専門参考人)

井上専門参考人、岡田専門参考人

(食品安全委員会)

佐藤委員長、山本委員、川西委員、吉田緑委員

(事務局)

川島事務局長、小平事務局次長、中山評価第一課長、吉岡評価第二課長、
池田評価情報分析官、橘評価技術企画推進室長、窪崎課長補佐、青木評価専門官、
石神技術参与

5. 配布資料

資料1 食品健康影響評価技術研究(ベンチマークドーズ手法の健康影響評価に
おける適用条件の検討)

資料2 食品健康影響評価におけるベンチマークドーズ法の活用に関する指針
(骨子案、動物試験編)

参考資料 新たな時代に対応した評価技術の検討～BMD法の更なる活用に向けて～

6. 議事内容

〇〇〇 予定の時刻となりましたので、ただいまから第11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、暑い中、御出席いただきましてありがとうございます。

机上配布しました資料の一番上の議事次第をごらんください。

本日は、6名の専門委員の先生方が御出席でございます。また、専門参考人として、井上専門参考人、岡田専門参考人に御出席いただいております。

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。

議事に入る前に、事務局から御報告があるということですので、よろしく願いいたします。

○事務局 3件御報告がございます。

1件目です。7月1日付けで食品安全委員会の委員の改選があり、山本委員を除く6名の委員が就任されました。委員長には佐藤委員長が再任され、引き続き委員に就かっている山本委員が新たに委員長代理に就任しました。

そのほか、本日御出席されている吉田緑委員が再任となり、新たに川西徹委員が委員に就任されましたので、御報告いたします。

川西徹委員でございます。

○川西委員 川西でございます。新たに委員となりました。よろしく願いします。ちょっと直感派ですので、こういうことになかなか合わないかもしれませんが、よろしく願いします。

○事務局 また、本日御欠席でございますが、香西みどり委員と吉田充委員が新たに就任し、堀口委員が再任となりましたので、御報告いたします。

2件目としまして、本ワーキンググループで取りまとめていただきましたBMD法の指針作成に向けた議論の報告書「新しい時代に対応した評価技術の検討～BMD法の更なる活用に向けて～」は、7月10日開催の第704回食品安全委員会で報告され、内容が承認されましたので、御報告申し上げます。

改めまして、先生方の多大な御尽力に御礼申し上げます。誠にありがとうございました。

最後に、事務局の人事異動について御報告します。7月31日付で評価第一課の吉田が異動しまして、後任として中山が着任しております。

○中山評価第一課長 中山と申します。どうぞよろしく願いいたします。

○事務局 事務局からの報告は以上でございます。

○○○ ありがとうございます。

それでは、議事を進めます。

本日の議題は、「食品健康影響評価技術研究について」と「食品健康影響評価におけるベンチマークドーズ法の更なる活用について」です。

開催通知で御連絡しましたように、本日の会議は非公開で行いますので、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 お手元に議事次第、専門委員名簿、座席表のほか、

資料1 食品健康影響評価研究「ベンチマークドーズ手法の健康影響評価における適用条件の検討」。

机上配付資料食品健康影響評価におけるベンチマークドーズ法の活用に関する指針について。

資料2 食品健康影響評価におけるベンチマークドーズ法の活用に関する指針（骨子案、動物試験編）。

資料2別添という、2枚組みホチキス留めの資料。

右肩に「〇〇提出資料」と記載された2枚組みホチキス留めの資料。

ダブルクリップでとめられた右肩に「別添 補足資料」と記載のある頭紙の資料、をお配りしております。

ホームページで公開しない資料がありますので、取り扱いは御留意ください。

このほか、参考資料として、先ほど御報告いたしましたBMD法の報告書をお配りしております。

また、過去のワーキンググループでお配りした資料、海外機関等のガイダンス等について、ドッジファイルに綴じて御用意しております。

さらに、海外機関等のガイダンスをはじめ、BMD法の報告書において参考とした文献、本日の配布資料などがお手元のiPadに保存されております。審議内容に応じて適宜参照ください。

資料の不足等がございましたら、事務局までお申し出いただければと思います。

〇〇〇 よろしいでしょうか。

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、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。

〇〇〇 先生方、御提出していただいた確認書に相違はございませんね。よろしいですね。ありがとうございました。

それでは、本日の開催目的について説明いたします。

冒頭で事務局から報告がありましたように、本ワーキンググループでは、前回の会合で、ベンチマークドーズ法の活用に関する指針の策定に向けたこれまでの議論の結果を報告書として整理しました。

報告書では、今後、ワーキンググループは、報告書の内容をもとに、国内外の活用実績を踏まえつつ、指針の骨子を作成するとともに、集積した科学的知見を組み込んで指針を取りまとめるとしております。今回、事務局において、報告の内容をもとに指針の骨子案

を用意していただきました。

また、ワーキンググループの議論で明らかになった検討課題について科学的知見の集積を図るため、本年度、食品健康影響評価技術研究事業において、〇〇が主任研究者を務められ、〇〇も参画された研究班が設置されております。

指針作成に向けては、ワーキンググループが研究班と連携しながら科学的知見を集積することが極めて重要です。そのことは過去のワーキンググループの会議でも確認させていただいたとおりです。つきましては、本日は、最初に、〇〇から、研究班の研究計画・コンセプト等を御発表いただきまして、意見交換をしたいと思います。その後、事務局に用意いただいた指針骨子案について議論したいと思います。

よろしく願いいたします。

〇〇〇 おはようございます。〇〇です。専門委員に加えていただきまして、また、食品健康影響評価技術研究ということで、ベンチマークドーズ法の指針作成に向けた関連するシミュレーション研究を本年度からスタートさせていただきましたので、代表は〇〇なのですけれども、技術的な部分のシミュレーションなどは私が担当させていただいていますので、その中身に関して、間違っていたら〇〇にも正していただきつつ、説明させていただきます。

(食品健康影響評価技術研究の研究計画を発表の上、質疑応答及び意見交換)

〇〇〇 それでは、時間の関係もありますので、議事2に移らせていただきます。「食品健康影響評価におけるベンチマークドーズ法の更なる活用について」ですけれども、事務局に御用意いただいた指針の骨子案について御議論いただくに当たりまして、まずは指針の主旨について、ワーキンググループ内で共通認識を醸成させていただきたく、その案について事務局に机上配布資料として用意していただきましたので、そここのところの説明をお願いいたします。

○事務局 それでは、机上配布資料「食品健康影響評価におけるベンチマークドーズ法の活用に関する指針について(案)」を私のほうから御説明させていただきます。

先ほど〇〇からもお話がありましたとおり、この後、指針骨子案を御審議いただくに当たりまして、最初に指針の主旨などについて関係者間で認識を共有させていただきたく、事務局のほうでこちらの資料を整理させていただきました。

まず、主旨でございますが、先に取りまとめいただきました報告書にもございますとおり、食品安全委員会が、さらなる一貫性及び透明性を確保して化学物質の毒性評価でベンチマークドーズ法を活用するために、同法を活用する場合の基本的な考え方及び手順などを整理した資料とさせていただきます。

次に、指針の概要・構成でございますが、報告書の内容をもとに、BMD法の主要な作業ごとに基本的な考え方、手順、基準などを整理したものであり、そのうち、数理モデルの選択基準など、先ほど〇〇から御発表がございましたような一部の手順、基準等については、

その研究の成果も踏まえて改めてまた検討する、入れ込んでいくというようなこととさせていただきます。

そして、指針作成の流れでございますけれども、本ワーキンググループで御審議、取りまとめいただいた案を食品安全委員会に御報告していただいた後、食品安全委員会で公表している他の評価指針と同様にパブリックコメントを行った後、その出された意見も踏まえて食品安全委員会名で最終的に指針を決定、公表するという流れを想定しております。

また、本ワーキンググループの審議内容の外、スコープの外ではございますが、実際のBMD法の使用に当たりましては、先ほども少しお話しましたBMDSなどBMD法の支援ソフトを使用することを必要としますので、指針とは別に、ソフトごと、また適用するデータのタイプごとに、事務局内部用の作業手順書を別途整理したいと考えております。

2枚目に移っていただきまして、BMD法に関しまして、現時点で想定する今後のワーキンググループの審議スケジュールを記載しております。事務局といたしましては、先ほど御発表がありました研究班の成果を組み込む形で、来年度夏をめどに動物試験編の第1版のような指針を、さらに、再来年度の夏をめどに疫学試験編についても第1版を決定、公表するようなスケジュールを想定しております。

以上でございます。

〇〇〇 ありがとうございます。

指針をつくる際の基本的な骨格、時間軸、構成といったようなところについて御説明いただきましたが、よろしいでしょうか。何か御質問とか御提案はございませんでしょうか。

よろしいですね。見出しを並べたという感じのところですので、大きな異論はなかろうと思いますので、こういう形で進めさせていただければと思います。主旨についておおむね共通の認識ができてきていると考えますので、この考え方に基づいて指針をつくり、具体的な手順書へと進んでいくということにさせていただければと思います。事務局のほう、よろしく願いいたします。

それでは、資料2の審議に移りたいと思います。事務局から説明をお願いいたします。

○事務局 それでは、資料2について御説明いたします。お手元に資料2「食品健康影響評価におけるベンチマークドーズ法の活用に関する指針（骨子案、動物試験編）」を御用意いただければと思います。

最初に、一番上段にございます枠内の記載でございます。先ほども御説明しましたとおり、本骨子案の記載のうち、研究の成果をもとに改めて御審議いただきたい箇所については下線を付した上で斜体としております。

「第1 目的」でございます。先ほどの机上配布資料の主旨と同様の内容でございます。食品安全委員会がBMD法を活用する場合の基本的な考え方や手順などを整理することで、活用時のさらなる一貫性及び透明性の確保に資することとしております。

「第2 定義」でございます。本指針で使用するBMD法関連用語の定義を記載しております。（1）ベンチマークドーズ法に始まりまして、ベンチマークドーズ、BMR、BMDL、ペー

ジをおめぐりいただきまして、AICからリストラクションまで9用語を定義しております。基本的には、先におまとめいただいた報告書で整理いただきました用語解説の内容を引用させていただいております。

2ページ、23行目、「第3 食品安全委員会が行う毒性評価におけるBMD法の活用」でございます。報告書の記載内容を踏まえつつ、先ほど申し上げましたとおり、指針のたたき台となりますようにBMD法を用いる際の主要な作業ごとに基本的な考え方、手順、基準などを整理しております。

24行目以降でございます。まず最初に基本的な考え方としまして、化学物質の毒性評価でBMD法を活用する際は、本指針に基づいてBMDなどの算出を行う。また、BMD等の算出に必要な生物学的、統計学的な判断は、各分野の専門家の意見に従って行うこと。また、指針によらない考え方や手法をとる場合は、BMD等の算出結果を報告する際に、その内容ですとか理由をきちんと整理して併記するということが記載しております。

30行目、「1. BMD法を適用する用量反応データの収集、選択」です。

31行目からページをおめぐりいただいた5行目までにかけてまして、評価対象とする化学物質について、利用可能な全ての用量反応データを収集するとした上で、専門家の判断に基づいて試験なりエンドポイントの選択をとり得るとさせていただいております。

次に、3ページ、6行目ですけれども、毒性試験とエンドポイントの組み合わせから成る各データセットとさせていただいております。そのデータセットのうち、用量反応関係が成立しているものについて、BMD法の適用に当たって試験設計が適切であること、また、二値データなどのデータの種類ごとにBMD法を行うに当たって必要な情報を含むことを確認するとしております。

20行目、(3)ですけれども、データセットの統合について記載しております。より精緻な検討が可能となる場合があるとしつつも、事前に生物学的、統計学的な観点から慎重な検討が必要であるとして、幾つかの留意点を記載させていただいております。また、個体データまでさかのぼって活用することが望ましいとしております。

31行目から、グレーディングがあるデータを二値データとして利用する場合は慎重に検討すべきだといまして、具体的な注意点などは研究成果をもとに改めて御検討いただくという形にしております。

ページをおめぐりいただいて、「2. BMRの設定」作業でございます。

最初に二値データのBMRとしては、他の化学物質またはエンドポイントとの比較の観点、また、先ほどもお話がありましたNOAEL相当の用量で観察し得る過剰リスクも踏まえまして、過剰リスク10%を用いたBMD及びBMDLは基本的には算出しておくこととしております。

一方、2)ですけれども、この過剰リスク10%という値は、BMRの万能な基準値ではないとしまして、その後、幾つか例示も挙げつつ、各BMRは生物学的な意義づけのほか、用量反応データの特性等を考慮し設定するとさせていただいております。

次に、(2)連続値データのBMRでございます。この場合、用量反応データというのが多

様な分布を示すことが想定されるため、BMRはそれぞれのデータごとに専門家の関与のもとで設定することを基本とし、変化の最小レベルが生物学的に有意な影響とみなすことが可能な場合、また、同レベルが不明な場合に分けて、2)、3) でそれぞれBMRの考え方を記載しております。

21行目、4) で個体データ（原データ）が利用可能な場合は、情報量の減少を伴う点に留意する必要がありますが、カットオフ値を設定し、二値データに変換することが有用な場合があるとしております。

次に、24行目、「3. 用量反応モデリングと数理モデルの選択」作業になります。

まず、(1) 用量反応モデリングとし、26行目からの1) では、3 ページの6 行目以降に1 (2) で記載しておりました条件、試験設計が適切ですとかデータセットが必要な情報を含むという条件を満たしたデータセットについては、全てモデリングを検討するとして、28行目からページをおめくりいただいた2 行目にかけて、生物学的な根拠がある場合を除いて、各データセットに含まれる全ての用量反応データを用いてモデリングを行う。特定の用量群だけをカットするようなことは行わないということを記載しております。

次に、5 ページの3 行目、3) ですけれども、報告書でのワーキンググループの確認事項としても記載しております「『発現機序の本質を捉えた』数理モデル」があれば、そのモデルを用いてモデリングを行い、BMDなどを算出することが優先されるとして、そういった発現機序の本質を捉えた数理モデルがない場合とし、4) では、実際にはBMDS等のソフトウェアに収載されている数理モデルを用いてモデリングを行う。また、その際にはリストラクション、オンとオフ両方のモデルをモデリングの検討対象とするとしております。

次に、14行目、モデリング以降、PODの決定に至る部分でございますが、こちらについては後ほどに回しまして、先に「4. 結果の報告」と第4の部分を御説明いたします。

16行目、「4. 結果の報告」でございます。BMD等の算出結果は、過去の食品健康影響評価でも適宜文書化の上報告されておりますが、透明性の確保という指針の目的を踏まえまして、指針内にこのような結果の報告について規定してはどうかということで案を御用意いたしました。

BMD法を用いた結果につきましては、使用したBMD計算ソフトの名称、バージョン。BMD法を適用した各データセット、出典等も含めた情報。③とし、BMD等の算出に用いたBMRの値とその理由。ページをおめくりいただきまして、モデルのプロットなどを含むBMD等の計算結果。数理モデルの除外基準及び選択基準。PODの値及び当該PODを導いたデータセット、エンドポイントの場合もありますが、それを採用した理由を含む形で文書化し、評価の審議に当たって報告するというふうにさせていただいております。

最後の項目、7 行目、「第4 指針の見直し」として、国際的な動向、食品安全委員会での活用実績の蓄積等も踏まえ、要すれば指針を改訂することを規定してはどうかと

考え、事務局でその旨を記載しております。

それでは、先ほど後にお回ししました5ページ、14行目の(2)数理モデルの選択とPODの決定の部分にお戻りいただければと思います。

用量反応データへの数理モデルフィッティング以降、PODの決定に至るまでの手順でございますが、本ワーキンググループでも考え方が統一されておらず、さきの報告書でも異なる考え方が併記されているといった状態でございます。つきましては、過去のワーキンググループでの議論や報告書に記載された内容をもとに考え得る手順について、資料2の別添としてお配りしております資料に2つの案を事務局のほうで用意いたしました。本日は特にこの別添の内容を御議論いただきまして、可能であれば指針作成に向けてこの部分の作業について一定の方向性をお導きいただければと考えております。

右肩に「資料2 別添」とあります横表の資料をごらんいただければと思います。先ほど申し上げましたとおり、モデリング以降、PODの決定に至るまでの手順について2つの案を整理しております。

案1でございます。こちらは、あるエンドポイントに対して複数のデータセットがある場合、それらデータセットへのモデリング結果を総覧して、BMD等の算出に用いる数理モデルを設定する。すなわち、真のモデルを仮定するような作業を行う。仮に単一の数理モデルを選択し切れない場合はAIC値を指標、かつ重みとしてモデル平均化を行うとしています。

その後、設定したモデルに利用可能な用量反応データを当てはめていく。そして、適合度検定、プロットの目視などをもとに、真のモデルに当てはまらない用量反応データはBMD等の算出対象から除外する。そして、残った用量反応データの中からBMD値が小さいことを基準として、BMD及びBMDLを採用する用量反応データをエンドポイントごとに選択していくという手順でございます。

結果、エンドポイントごとに独立したBMD及びBMDLが算出されるということになります。

右側の案2でございます。こちらは、あるエンドポイントに対して複数のデータセットがある場合ですけれども、データセットごとにBMDなどを算出する数理モデルを選択するという手順でございます。各データセットに複数の数理モデルをフィッティングさせた後、まず用量反応データへの適合度、プロットの目視などをもとに、明らかに妥当でないモデルを除外する。その後、低用量領域における適合度合いを考慮しつつ、その先の選択基準は先ほどの研究成果も踏まえて改めて検討としておりますが、一定の選択基準のもとでデータセットごとにBMDなどを採用するモデルを選択、あるいはモデル平均化を行うとしております。実際に観察されたそれぞれの用量反応データに当てはまりのよいモデル、同データを基によい予測が得られるであろうモデルを選択する手順というふうに理解しております。

結果、同一エンドポイントでも、データセットが複数ある場合は、データセットごとに独立したBMD及びBMDLが算出されるものとなります。

ページをおめくりいただければと思います。次に、HBGVの算出に用いるPODの決定でござ

います。先ほど申し上げましたとおり、案1では、エンドポイントごとに独立したBMD及びBMDLが得られております。そこから各エンドポイント自体の意義、用いた用量反応データの特性、BMDとBMDLの値などをもとに、あるエンドポイントを選択し、そのエンドポイントにおけるBMDLをPODとする手順でございます。

案2は、データセットごとに独立したBMD及びBMDLが得られておりますので、POD決定の指標として挙げている項目は案1と同じでございますが、エンドポイントの選択、あるいは当該エンドポイントにおけるBMD及びBMDLの選択を含む形でデータセットを最終的に選択し、当該データセットにおけるBMDLをPODとする手順となります。案1に比べまして、この段階で少しまとまった判断を行う形を想定しております。

さらに、この案1及び案2に基づきますBMD等の算出手順、また、その作業をイメージしていただくため、補足の資料を用意いたしましたので、そちらの説明もあわせてさせていただきます。

お配りしている資料で、A4頭紙の右肩に「別添 補足資料」とある資料を御用意いただければと思います。本資料では、ある評価対象物質のあるエンドポイントに関しまして、利用可能な用量反応データが複数存在する場合を仮定しまして、実際にBMD算出ソフトウェアでありますBMSDを用いましてモデリング、フィッティング及びBMD等を算出した結果を整理しております。仮定した用量反応データは頭紙にございますAからCの3種類でございます。

その後、ダブルクリップを外していただきますと、ホチキスどめのものが4つ出てくると思います。開いていただくと「案1 補足資料」と真ん中ぐらいい書いてある資料がございます。こちらは先ほど御説明した案1の作業イメージを整理した資料でございます。紙ですと見づらいという場合もございます。iPadのほうに文献を整理しておりますけれども、その中に頭文字A-4という資料がございます。こちらを開いていただくと、こちらの紙の資料が電子化されたものが入っております。スライドは横方向にお願いできればと思います。こちらのほうがグラフ等をごらんいただく場合は拡大ができますので見やすいと思います。適宜ごらんいただければと思います。

案1では、先ほど申し上げましたとおり、全ての試験結果を総覧しまして、各用量反応データ全体を説明できるモデルを設定するという作業を行いますので、「案1 補足資料」の1枚目に3つのデータセット、A、B、CにおけるAICですとかBMDなどの算出結果を整理しております。2枚目以降、モデリングした結果のプロット一覧、縦にデータセットA、B、C、横に同一モデル、上からガンマのリストリクシオン・オン、リストリクシオン・オフ、ロジスティックといった形で、その後、5枚にわたってモデルを一覧にしておることとさせていただきます。案1ではこういった作業の整理を行った上で、これらを総覧して単一のモデルを仮定するという作業を行うこととなります。

次に、右肩に「案2 補足資料」とあります資料をごらんいただければと思います。こちらはA4横表とモデルのプロットが整理されたA3の資料を、データセットA、B及びCそ

れぞれについて整理しております。こちらにもiPadに同じ資料が入っております。頭文字はA-5になると思います。こちらを適宜ごらんいただければと思います。

案2では、データセットごとにBMDなどの算出結果とモデルのプロットをもとにしまして、妥当でないモデルの除外、代表モデルの選択を行い、結果、独立したBMDなどを得ることになります。ですので、こういった形でデータセットごとにプロットした結果、BMDなどを算出した結果を見て判断を行うことになろうかと思っております。

このため、先ほどの「案1 補足資料」をデータセットごとに分割したような資料をもとに判断を行うことになります。また、妥当でないモデルを除外するという作業も行いますので、除外に当たっての基準の候補と考えられますBMD分のBMDLといったような指標も、先ほどの案1の資料に追加して表のほうには加えております。こういったデータをもとに判断作業を行っていくというイメージをしていただければと考えております。

資料2についての事務局の説明は以上でございます。

なお、案1に関連しまして、〇〇より参考資料を御提出いただいております。

以上、繰り返しではございますが、この後の御審議では、特にこの別添の内容について改めてワーキンググループで御議論いただきまして、可能であれば一定の方向性をお導きいただければと考えております。

以上、重ねてよろしくお願いたします。ありがとうございます。

〇〇〇 どうもありがとうございました。

ただいま事務局から資料2の指針骨子案について説明がありました。その中で実際のBMD法の手順をまとめた第3の内容、中でも特に報告書で必ずしも統一されていない箇所を整理した別添資料の内容について、ワーキンググループの議論を深めたいという要望が出されております。

つきましては、会議時間も限られておりますので、今回、この後の時間の審議は、まず別添資料の内容について審議を行いまして、時間の許す範囲で資料2本体の内容のうち第3部分の内容と、その他の残りの部分の審議、この順番で審議できればと思っております。

先生方には事前に資料をお送りしてございまして、資料につきましては、御意見とかお気づきの点がいろいろあろうかと思っております。私自身も事前の資料で気づいた点もありますので、物の考え方のところ、これから議論するところ以外につきましては、終了後に口頭もしくはメール等で事務局のほうに御連絡いただくこととして、残りの時間を物の考え方の十分収束を見ていないところについて議論させていただければと思っておりますけれども、よろしいでしょうか。

それでは、資料2の別添について審議いたします。この内容につきましては、今までのワーキンググループとして一本化した考えには至らなかったところですが、今後の指針をつくるに当たっては必ず通らなければいけないところですので、考え方とか手順につきまして、案1とか案2と書かれたところを中心に御議論いただければと考えております。

余り〇〇が最初にしゃべるのはよくないとは思いつつも、どうしてもお伝えしておきたいことがあるのでしゃべらせていただきますが、配布された資料の中で色刷りの私が提出した参考資料というものをごらんいただければ幸いです。これはちょうど「資料2 別添」の本質にかかわる部分につきましてまとめたものです。今までのワーキンググループの審議で、当然皆さん共通の認識に立って、共通の物の考え方で進んでいると思っていたのですが、こうやって事務局にまとめていただいていると、必ずしも同じ土俵に乗っていないかもしれない。これはこのワーキンググループがというよりは、この食品安全委員会に、過去も含めて、それからひょっとするとEFSAも含めて、ちょっと物の考え方にずれがあるかもしれないなと思ったので、原点に戻るつもりで資料をまとめさせていただきました。

2枚組みになっていますが、最初は推測統計学、私たちが推計学と言っている考え方です。基本的に統計学にはベイズの考え方とネイマン・ピアソン流の考え方がありますが、推測をしていく場合には一応というか、我々の採用する統計学は基本的にネイマン・ピアソンの考え方によっております。この推測統計学では、まず初めに2つの重要な仮定があります。1つ目は、標本は母集団からランダムに抽出されるという仮定です。母集団はそもそも一つ。母集団というのは真実のことです。しかし、真実は見えない。見えない世界に真実があるということなのですが、それは一つしかない。ランダムに抽出された標本というのは、実際の実験の結果であったり疫学研究の結果であったりしますが、これは幾つでも存在し得るというものです。これは見える世界の話です。標本は母集団からランダムに抽出されるという一つの重要な仮定。

もう一つは、事象は数学モデルに従って発生するという第二の仮定があります。数学モデルというのは、真実を数式化したものでありまして、どういう用量反応関係を示すかとか、どういう時間経過を示すかとか、さまざまなモデルがありますけれども、いずれにしても真実の世界を数式化したものといえます。この数学モデルが論理的に決まれば容易なのですけれども、なかなか生命現象というのは複雑であって、簡単に理論化できないところがある。一部マルチステージのような、多段階モデルのような理論があることはあるのですけれども、普遍的かどうかもわからないので、なかなか論理的にエイヤッと決めるのは難しいところがあります。

かといって、実際に観察結果を見渡してみてもモデルを決められるかということですが、先ほど事務局がA3を折り畳んだ資料の中にグラフをつけていただいています。それをぱらぱらと案1のほうのグラフを見ていただくとよろしいかと思うのですが、いろいろなデータセットで当てはめてもらっていますが、いろいろなモデルが出ている。例えばガンマ分布であるとか、ロジットであるとかプロビットとかいろいろありますが、モデルが当たっているといえば当たっているのだけれども、何となくどんぴしゃという感じがなかなかしづらいと。確かにモデルは分布範囲の中を走ってはいるのだけれども、この線しかだめなのというわけでもないという感じがすると思うのです。標本を見渡して縦覧、横断

的に見て、モデルが仮定できればいいのだけれども、これも現実問題なかなか難しいところがあります。

サイエンスはどうしているかというと、ある理論、あるモデルを仮定して矛盾、すなわち逸脱がなければ正しいとするという科学の公理みたいなものがあるわけですね。例えば生存分析でコックス比例ハザードモデルなどを我々はよく使いますが、これは各時点の2群の曲線の傾きの比は一定、接線の傾きの比はどの時点ではかっても一定ですよという前提のもとにつくられたもので、形自体はあらゆるものを許容します。こういった比例ハザードモデルが使われているのだけれども、実際にこれ以外にモデルがどのくらいあるのかよくわからないし、もうこれはデフォルトのような形でたくさん使われていて、コックスモデルでやってみて実際余り困ることがないということから広く使われていて、厚生労働省の薬事承認もこれを使って行われています。

したがって、科学というのはある仮定をして、あるモデルをつかって、大体世の中のことが矛盾なくいける場合はそれを正しいとするというような仕組みになっています。光の速さはどんな状態でも秒速30万キロであるというのも仮定であって、光速に近いスピードで走っていったものにおいても光速は秒速30万キロと。そのようになると空間がゆがんでくるわけです。秒速に近い状態で物を見るとゆがむということになるのですが、そんなこと誰も確認したことはないのだけれども、そういう仮定を置くと世の中の事象がうまく説明できるので、光速はどこではかっても、どんな状態ではかっても一定であると。空間がゆがもうが何しようが関係ないというような仮定をして、それを正しいとしているわけです。

ですので、真実がわからない場合は観察結果に基づくのだけれども、これまたなかなか観察結果だけから導くのは難しいので、ある理論とか一つの仮定を提示して、矛盾がないことを積み重ねていって真実としているというところがあります。

そう考えると、BMD法におけるモデルというのは、合理的であればどれでなければいけないということはないとも言えるわけです。ちょっと乱暴な言い方をするのですけれども、例えばロジスティックですとかプロビット、あるいは発がんに関して言えばマルチステージとか、そういったものが何となく納得しやすいという印象があります。神様は簡単なものを好むということがあって、シンプルなもの望ましいということになります。

とはいえ、いろいろモデルがある中でエイヤツと決めてしまうことに不安がないわけではないので、その正当性の補強のためにAICとかその他の質、モデルの性能をあらわす指標を加味するということが現実的に行われるかと思えます。これがネイマン・ピアソンの考え方の重要な仮定と、それから科学における真実の推定というものの理屈であります。

ページをめくっていただいて、もう一つこれに関連して、考え出すと次から次へいろいろなことが出てきてしまうのですけれども、科学的な検討と行政上の措置というものをきちんと分けないといけないのではないかと思います。これはレギュラトリーサイエンスであって、国民の健康を守るためにやっているわけですが、国の規制は2つの異なるプ

プロセスから成り立っている。1つは科学的な閾値であって、これはサイエンスの手法に基づいて自動的にと言うと変ですけれども、論理的にといいますか、恣意が入らずに決まるという意味です。研究の質を評価するということで、ここでも先ほど事務局から説明があったように一定の条件を設けて、それを満たしているかどうかを判断するというので、ほかの世界でいくとコクランのシステマチックレビューであるとか、あるいはSTROBEと呼ばれる観察研究の必須の評価など。STROBEはここでも疫学研究に使えますけれども、そういった質の評価のための指標があります。

動物実験についてはよく承知しておりませんが、先ほど事務局から提示があったようなものが一定の質の評価の目安になるだろうと思います。

そして、最尤値（点推定値）を提示すると。点推定の確かさを信頼区間で示すものであって、最尤値と信頼区間は同格ではない。論文の表記も、例えばオッズ比を出す場合だと、オッズ比とその95%信頼区間。英語で言えば、Odds ratio and its 95% confidence interval、複数の場合が多いのでtheirが使われますけれども、いずれにしても、itsとかtheirとか日本語であらわれないけれども、英語だからわかる従属関係というのがよく見えます。こういうところが平均と標準偏差とか、あるいは中央値と四分位範囲というのとは全く異なる。平均と標準偏差は代表値とデータのばらつきを示すそれぞれ独立の指標ですので、そういったものとは異なって、信頼区間というのは従属的なものであるということが言えます。

こういったサイエンス的な閾値の問題と、もう一つは実際に国が行う規制値の設定というところですが、これは一定の目的があって、国民生活とか国民経済、あるいは国民感情に配慮して恣意的に決めるというところがあります。その一つが安全サイドに立つという考え方で、幾つかデータが出てきたら最も小さい値を採用するとか、あるいは信頼区間の下限値を採用するとか、安全係数を乗ずるとか、そういったところは施策として恣意的に行うものであるということになります。科学的に決められないものでも行政的に決めなければいけない場面が出てくるわけで、食品安全そのものではないですが、類似の例として、これは職場の環境、産業衛生の世界で許容濃度と管理濃度というのがあります。許容濃度は科学的に決められたものですが、発がん物質などは許容濃度を決められないので、許容濃度は決まらない物質もたくさんある。でも、だからといってなしというわけにはいかない。実社会で一定の管理をしないとイケないので、厚生労働省が管理濃度というのを決めている。これは行政措置としてのもので、許容濃度と一致する場合がありますし、全然一致してない場合もあるということです。

そういうふうに2つのプロセスがあるので、サイエンティフィックに論理的に決めるものと、使用目的に応じて意図を持って決めるという2つの段階があると思うのです。そう考えると、やはりそのプロセスはきちんと分けて記載したほうがよろしいと思うのです。EFSAの指針の書き方もそうなのですけれども、両者がかなりごちゃ混ぜになってというか、入り組んで書かれている感が拭えないので、我々としては、きちんとその2つ、2段階ということ意識して書くべきではないかと考えております。

最後に附録ですけれども、先ほど〇〇から御説明があったことなのですが、BMRも2種類に分かれる。〇〇から御意見が出たように2つの要素があって、生物学的、これは臨床的であったり毒性学的という表現もあろうかと思いますが、生物学的に意味のある最小の変化ということが議論できる場合、一般にnが大きいときです。その生物学的な影響も、個人レベルでの影響か集団レベルの影響かということを考えないといけない。例えば以前にも例に出しましたけれども、収縮血圧が2ミリ下がるということは、個人レベルではほとんど問題にならないのですけれども、国民全体を考えると万単位で虚血性心疾患とか脳卒中の影響が出てくるということになりますので、どういうスコープで影響ありとするのかということをはっきり意識しないとイケないと思います。

それに反して、nが小さいときはバイオロジカルな議論が多分ほとんどできないと思います。そうすると数に依存するのではないかと思います。この数式は私が頭の中のものをただ置きかえただけで、先ほど〇〇からそんなに変わらないよというお話でしたけれども、それは確かに10から100までの世界では、実質的にルートnなので多分余り変わらないのです。桁が違うということはないのですけれども、このように数に依存する部分もあるのではないかと考えています。

ということで、共通の考え方というか、そもそも統計学的な推測をするときに共通の認識があるかと思う。それに基づいて仮定が2つあるよとか、あるいは値を決定するときには2つのステージがある、段階があるということ。そして、BMRにも2種類あるだろうということを議論の素地として提案させていただきました。

かえってお邪魔になったかもしれませんが、こういう考え方に基づいて議論が進めばいいかなと考えております。批判も含めてこういう考え方をお出しいただきまして、別添の案1、案2について御議論いただければ幸いです。本当は座長が司会と意見を同時に言うのは問題があったかもしれませんが、済みませんが、そこはお許しいただきたいと思います。御意見をお出しいただければ幸いです。

〇〇〇 〇〇の資料の1枚目の一番左上の絵なのですけれども、多分、私自身は生物の実験なんてやったことがありませんので、どちらがいいとかいう話はできないのですが、案1と案2の違いで案2を推される先生方は、標本1、2、3の下に観測仮定が独立についているように考えておられるのではないですか。ですから、この3つのデータセットを全体一つのデータセットとして考えて、案2の先生方は、その変数として実験A、実験B、実験Cがあって、それで層別しなければ話は進まないとお考えなのではないかと思うのですけれども。

〇〇〇 そのことはよくわかります。我々はメタアナリシスとかをやりますと、ホモジェナイエティーとかヘテロジェナイエティー、同一性、異質性の問題にかなりぶち当たりますので、測定計とか対象者が本当にランダムでない選び方。ランダムで選んでいるというのはそもそもかなり理想論で、実際には手近なサンプルでやっているし、実験条件も実験環境も異なるので、そもそも同質性が担保できないということで別の系として考えるとい

うのは十分理解できます。ただ、どこかで最後でまとめないといけないので、実験条件とか動物の種は一緒でも生育条件とかが違った場合に、条件が違うから値が違って当然ですけれども、それを最後は一つの数字に置きかえないといけない。どの段階で統合的な考え方にしていかなければいけないのかなということで、今回の絵は、もともと同質であるという前提なので、明らかに異質性があれば、それは別の実験、別の観察、別のテーマであるということを想定しないといけなくなります。

ここで示したのは、あくまでも同質性があるという場合の話で、〇〇がおっしゃった意味は現実にはすごく大きな問題で、異質性があるのだったらそれはそもそも違うということで、別の系だと考えざるを得ないと思います。それをどうまとめていくかは難しいかなと思います。

ほかにいかがですか。

〇〇〇 〇〇のは別に反対も賛成もなく、私から見ていると、実は案1も案2も全然変わっていないのです。それはなぜかという、動物実験だと案1も案2もきっと同じことになるというのは、エンドポイント、イコール、データセットなのです。疫学データとエンドポイントは、例えば乳がんが出たときのデータセットはいろいろな疫学データがあって、そこに普遍的なモデルが存在するのですけれども、動物実験の場合は1個の実験の中にいろいろなエンドポイントがあって、その一個一個が疫学で言うデータセットで、しかも、エンドポイントにデータセットが1個しかないという状態になっているので、これを整理するのは多分、エンドポイントとかデータセットという言葉の定義をクリアにしないとごっちゃになる。

EFSAがごっちゃになっている理由は、動物実験を全体に置いてやってしまっていると、本来、多分最初に〇〇の質問にあったランダムに発生させるのはエンドポイントにとって実はランダムではない。例えばコレステロールの変化だったら分布は低いところとか高いところがある。だから、今、〇〇が、シミュレーションで二項分布と言われたのですけれども、あれはエンドポイントごとに分布がばらばらなので、シミュレーションはエンドポイントごとにやらなければいけないのですごく大変なことになるという話が根底にあるので、それをエンドポイントと言うか、データセットと言うかという、動物実験の中での定義をちゃんとすると、多分、1と2は同じところに入れられるのではないかと思います。

現実的に(2)の1)を、例えば、あるがんの発生はプロピットに近いだろうとか、別のがんはワイブルに近いだろうかというのを決められるのでしょうかけれども、多分それが決められないので、エンドポイント1個は1個のデータセットにイコールになって、そのデータセットのどれを選びましょうという議論で動物実験は暗黙にやってしまっているのです。それを説明なしに暗黙にやってしまっている部分が誤解を招いているだけなので、多分、案1でいいのです。案1でいいのですけれども、1)は現実的に不可能なので、エンドポイントはこのデータセットで代表しましょうという定義をした上で、そのデータセ

ットのモデルはどれがいいでしょうということを統計学的にセレクトしましょうと。理想論をちゃんと置いてから書いていけば、1と2は一緒に組み合わせられるのかなと思っています。

〇〇〇 ありがとうございます。悩む必要がなく、そもそも動物実験では、同一の物質、同一のエンドポイントの実験がそうたくさんあるわけではないと。

〇〇〇 あと、先ほど言った、多分、動物実験の場合は種が変わったりすると反応が違うというのを皆さん常識的に、暗黙のうちに言っているのです。だから、疫学と違って、動物実験が変わるとコンバインしないのですよ。それが発生しているものだから、エンドポイントごとに何かを決めるという操作がそもそも動物実験で評価するところの中に余り入っていないのです。暗黙的にやっていないので、スキップしているのです。本当はやらなければいけないところをスキップしているので、それでそごが出たように見えるのですけれども、私から見ると、案1のところ案2のモデルを組み込まなくてはならないのだと私は思いました。

〇〇〇 動物実験は数が限られるというのももちろん承知しておりまして、実際には同じ物質でほぼ同じ実験条件で、同じエンドポイントを持った研究というのは複数あることが余り多くない。ですので、複数ある場合はかもしれないですけども、その場合の考え方で、1つしかない。あとはとても同質性とは思われない実験、動物種が違うとか、投与ルートが違うとか、根本的に最初から異質だよねというのがあった場合は完全に別にしないといけない。だから、書かなくても自動的に決まってしまうのかもしれないので、複数ある場合はそういうことだし、ひょっとすると疫学研究のところではこれが大きな問題になってくるかもしれませんが、動物実験においては複数ある場合はこうする。そうでない場合は自動的に案2のほうに流れていってしまうということになるのだろうとは思っております。

ただ、その場合にも、今書いてあるのは、いろいろやってみてその結果から絞っていくというプロセスですね。サイエンスというのは、やってみた結果を見てから選ぶこと、ポストホックにやることをすごく嫌うので、アプライオリに決めていくのがサイエンスの正道なのですけれども、先ほど言ったようにこれは行政施策としてやるので、正道で出したものの中から社会施策としての選択をするという、ここをきちんと分けたいなとは思っているところであります。

〇〇〇 もう一つつけ加えたかったのは、このガイダンスの中で数理モデルと一つにくくっているところが実は2種類あると今、気づいて、生物学的数理モデルと統計学的数理モデルの2つあるのです。本当は理想的には生物学的数理モデル。例えば、酵素の反応系とかが全部わかって、たんぱくとのバイニングも全部わかって、それを全部数理的にあらわしてやるのが生物学的数理モデルで、それが本来あるべき姿で、薬などは薬理的にはみんなそういうモデルで、それも統計学といえれば統計学かもしれないけれども、PBPKのように本当に全部乗数も掛けて反応も全部導き出しているモデルはできないので、統計学モデル

で近似したモデルを探すという操作をするので行政的なところに少し落ちてくるのだと思っています。

先ほどの〇〇の2枚目の、閾値は科学的、自動的に決まるところで、疫学データだこの工程はできるのですけれども、実は動物実験の場合は、ここにも既に規制的なものが入ってきているのです。なぜかという、動物実験は幾つもあるうちの一番安全側のNOAELをとっている。その時点で実はもう最尤値を選んでいないのです。安全側のほうを既に選んでいるので、動物実験の評価をしているときのNOAELの選択の時点で既に実は規制的な側面の概念が入ってきているのです。疫学の場合は最尤値とか信頼区間を出しますけれども、動物実験の場合はこのときの信頼区間の下限値をNOAELに設定しているのです、ここで行政的なプロセスが前段階に入ってきているので、その辺がごっちゃになっている。サイエンスと規制とがごっちゃになっているところだと、私は〇〇の話を聞いて気づいたというか、そのように感じました。動物実験の場合と疫学の場合で、そもそものコンセプトとプロセスが結構違うなという気がしました。

〇〇〇 今の〇〇のお話が私はとてもずっと入ったのですけれども、例えば動物実験で、OECDのテストガイドラインもそのものがある意味では〇〇の2番目の規制値としての概念が入っているかもしれないわけですね。例えば、今回、アクリルアミドは飲水投与でしましたけれども、これがフィーディングスタディーの場合だと、投与量は年をとるたびにどんどん下がってきますから、最終的に一番低い値ということで安全サイドをとっていることになりまして、最近、バックグラウンドとしてヒストリカルコントロールデータをどうするかということになると、今回、データセットBでゼロが2つありますけれども、本当の集団というものから取り上げたときに、たまたまこの対照群、 $n = 4$ がこうだったけれども、本来の対照群はというのでヒストリカルコントロールでどう考えるかということもあるのです、コメントですけれども、〇〇がおっしゃったのは、確かに疫学というものとテストガイドラインで規制されたような実験データはかなり違いがあるということですね。

〇〇〇 NOAELの意味が違う。動物実験でふだんしている評価のプロセスを、〇〇が説明していただいた疫学のプロセスのどこに入っていくかを見ていくと、必ずしも下ではないのです。上のほうに既に入っているなと感じたので、そういうふうに思いました。

〇〇〇 どうぞ。

〇〇〇 疫学のデータと動物実験のデータでそんなに取り扱いが違うのかと言われると、自分が理解している限りではそんなに変わらないのではないかと信じていますということなのですが、幾つかデータセットがあったときに、動物実験だと投与方法とか実験条件で動物種が同じであると、同じデータとして扱っているのです。ほとんどデータ処理をせずに複数の研究成果をそのままガッチャンコしてデータ分析していることもあるみたいなのです。それぐらい同質であるということがある程度保証できるような実験条件をつくれるという点では違うのですけれども、ただ、複数のデータセット、複数の観察標本が何らかの設定であるときに、疫学データだったらばエンドポイント毎で比較をしないといけない。

見方、角度が完全に90度変わるかということ、システムチェックレビューとかの考え方は、標本はランダムに出てきているという想定のもとでガチャッとヒトの研究成果の疫学データを統合しているので、ほとんど同じように統合しているというのが、今、自分が理解している限りで、それよりもサンプル数が大きくなるということが、システムチェックレビューが求めているフィロソフィーですね。

もう一個だけ、リスク解析そのものが規制のリスク管理のためのサイエンスなので、これはサイエンスではなくてオペレーションズリサーチだと思って自分はやっています。どういうことかということ、サイエンスと規制の部分が分けられないのがリスク解析のモデルであって、BMD法というものの自体が、分位点をとって恣意的に何かを決めている時点でそこに客観性はないのです。なので、リスク管理のための恣意性のあるサイエンスでやっているから最小値をとるとかというような話もモデルの選択とかの中に入ってきてしまうというのが、今、自分が理解していることです。なので、ここを結構きれいに分けて議論できるかということ、まあまあ難しいかなという印象を持っています。

〇〇〇 ありがとうございます。

ほかにいかがでしょうか。

〇〇〇 今の議論を聞いていて、私はちょっと責任を感じる場所があるのです。というのは、実は、この報告書を書くときに急げと言った手前、議論が余り煮詰まらないままに報告書になってしまって、それが今に至ってこういう議論になったのかなと思うのです。逆に、主に〇〇と〇〇の話を伺っていると、物の見方が違うようだけれども、現実問題としては同じようなところを目差しているというのもわかって、一安心でもあると思います。

多分、今、〇〇もおっしゃったように、我々がやっていることは何であるのかというのを考えた上でこの指針はつくっていただきたいなと思うのです。一方で今のフィロソフィカルな議論というか、ファンダメンタルと言ったらいいのかよくわからないけれども、その辺の議論も非常に大事な気がするのです。だから、私の責任のとり方として、今すぐではなくもていいと思うのだけれども、今のような議論を報告書の補遺みたいな格好でつくられたらいいのかなと思いました。

指針のほうはあくまでもプラクティカルに、読んだ人が、これは何だ、どちらをやったらいのだと迷わないようにつくっていただきたいなと思っております。また注文してしまっただけで申しわけないのですけれども。

〇〇〇 ありがとうございます。〇〇の御下命でもありますので、補遺の形で、根底にある考え方を整理してメモとして残しておくということで理解いたしました。

実際に通るプロセスは、動物実験だとそもそも右から行こうが左から行こうが結局同じ道に入ってくるだろうから、しかも、動物種の問題とか、一つの実験ではエンドポイントがいろいろ違うので、どっちみち別々に扱うことになるので、余り実質的には問題はないということになることが大体確認できたということ。

もう一つは、結果の記述のところレギュレーションの視点から、最初からその視点が

入ってきてしまうので明瞭には分けられないということではあるのですが、最初から最も小さい値をとっているわけではなくて、並べた上で最小値をとりますし、信頼区間も最尤値を示した上で信頼区間の下限値を示すので、記述についてはある程度段階的に、まざるでしょうけれども、記述の仕方としては順番を決めることはできるかもしれないと思っております。

ほかにいかがでしょうか。

〇〇〇 1点質問してもいいですか。〇〇への質問になるかもしれないけれども、この指針にはどのソフトウェアを使うかというのは多分書いていなかったと思うのですが、BMDSを使うか、PROASTを使うかというのは何か根本的な議論があるのですか。それとも、イクイバレントだと考えていいのですか。

〇〇〇 詳細の点では結構異なります。コンピューターリテラシーがどれだけある方が使うかによって進めるのも、私はちょっと書いていますけれども、BMDSというのはGUIといって、コンピューターでマウスでダブルクリックするとウインドウが開いて、そこにデータを入れるだけで結果が返ってくるので、ウインドウズが使えるれば、そのマニュアルさえ覚えれば使えるというものです。一方で、信頼区間の計算とかで一部雑な部分だとか、更新が滞っている部分だとかというのがのも事実です。

一方で、PROASTというのは、「R」という統計パッケージを一定程度使い使った経験がないと使用できないのですが、それを走らせることに抵抗がないというレベルまでいかないといけない。それは一定レベルでなれてもらわないといけないので、そこまで例えばマニュアル化するとなるとまあ大変になりますので、現状で何らかの化学物質の評価が得られているときというのは、よりBMDSのほうが行政でコモンに使われているものだと、そういう理由でなっているものだと認識しています。

〇〇〇 今のところわかっている限りは、クリティカルな違いみたいなものはないという理解でいいですか。

〇〇〇 専門家の中でBMDLみたいなポイント・オブ・ディパーチャーの値がソフトウェアによって結構変わるといふ議論はあります。BMDSというのはよくできたソフトウェアなのですが、BMDLの計算過程で必ずプロファイル尤度というものを使うのです。それと漸近的正規性という、統計学的には一定の想定が必要で、小サンプルに向いていない計算の仕方をしています。PROASTの場合はブートストラップ法というのを使っていて、より精密に計算できる一方で、過剰にBMDLを過小評価する可能性があったりする。最近はそのようなもの折衷案をとったような、ベイズ法を使いましょうだとか、新しいソフトウェアとかも登場してきたりしていて、そこも突き詰めるとちょっと議論をしていくことは必要だと思っております。

〇〇〇 実際にソフトをどうするかというのは最後には避けられない問題になるかと思うので。シミュレーションの研究の中で両方使われるということはあるのですか。

〇〇〇 はい。多くの物質で、例えば国で、一度食品安全委員会として評価したことがあ

るものとか、BMDSで一度やられている記録があるので、それでやり直した上で、ソフトウェアごとに私たちは閾値を出してダブルチェックしたりしています。

〇〇〇 研究班からのそういう御報告や情報をもとに反映することがあるかもしれないのですけれども、その取り扱いは今すぐは難しいかもしれないので、また検討課題に。

〇〇〇 今決める問題でもないかもしれないし、やはりこれもプラクティカルに、BMDSで今までやってきていてほとんど差し支えなければ、それでとりあえずはいいと思うのです。ただ、そういう問題点があるかどうかというのは認識しておいたほうがいいなという感じはしました。

〇〇〇 どうぞ。

〇事務局 我々も動物試験に関しては特定のエンドポイントに対して、なかなかふんだんに試験、ここではデータセットという言い方をしておりますけれども、そういったものが存在しないというのは我々も承知しております、〇〇がおっしゃったようにデータセットが一つしかエンドポイントに対してない。逆なのかもしれないのですけれども、試験に対して複数のエンドポイントがあるようなものしかないという場合は、右から入るのか左から入るのか、見方がちょっと違うのかもしれませんが、案2に実際のプラクティカルな作業としては落ちていくという理解をしたのです。

実際に今回、例示という形で示しましたが、実際のある化学物質で系統までは確認できていませんが、同じ動物種でこういったデータが得られている、複数のデータセットが得られた場合の扱いというのも指針上、プラクティカルに落とし込めればいいのですが、どういった考え方のもと、指針の文章を整理すればよいのかというところが、もう少し御意見をいただけないかなと思ったのですけれども、いかがでしょうか。

〇〇〇 ちょっと振り出しに戻すかもしれないのですけれども、〇〇への質問かもしれませんが、そもそもモデルはエイヤツで決めてはいけないのかどうか。

〇〇〇 エイヤツの意図するところを教えてくださいませんか。

〇〇〇 今、いろいろなシミュレーションをやって、普遍的にこのモデルがいいとかいうのが決まればいいのですけれども、そういうことを期待して、何らかの根拠がないと決めるににくいので、もちろんそういうシミュレーションスタディーでこのモデルは普遍性が高いとか、安定度がいいとか、そのような話になればそれを優先的に使うことはあるのかなと。一々実験セットごとにやっても事務局から用意していただいたように、外れてはいないのだけれども、これはすごくフィットがいいよねと、絶対これがモデルだよねというほどの心象は余り得られない。AICも近い数字になってしまうので、やってみないと決められないというほど迷わなければいけないものかしらという気がするのです。

〇〇〇 ものすごく迷うものなのだと理解しています。例えば、〇〇の案1のものだと、わからなかったらAICをとりあえず使ったらいいではないかということだったのですけれども、AICだけの問題とかでばっと出すと、このリスク解析の話だと、低用量領域のデータが少なかったらAICがよくても低用量領域評価がよいとは限らないとか、AICが2未満とか

だったらどのモデルがベストかというのは最終的に決断を下せないとか、すぐAICの問題点が列挙できるのですね。なので、どの基準を持ってこのモデルで行こうというプロセスをえいと決められるかというのが難しいので、今たくさんシミュレーションをしています。モデル選択のプロセスには様々な要素が関わるので、こういった基準に基づいて選択するのがベストかという点に関して十分な科学的見解を出さないと、決断は難しいのだと思います。

〇〇〇 わかりました。班研究によって特にプライオリティーが決まらない場合は、施策としての観点からやってみて一番目的にかなうものを選ぶというストーリーになる。これを事務局でうまく文章にまとめる。

だから、対立的に書いていただいたのですけれども、複数のデータセット、同質性が期待できるものが複数あった場合はこういうことも考慮するけれども、そうでない場合がほとんどなので、基本的には各実験のセットごとにこういうことを出して、こういう数字の絞り込み方をするという流れをつくる。注釈として、同じようなものが複数あった場合に留意すべき点というような書き方があろうかとは思いますが。同質的なものですね。もちろん同じエンドポイントでも、同じ物質でも、実験環境が大いに違ってもともと同一レベルで議論できないとなったら、最初から別物として扱うということです。その後は、一つの物質につきエンドポイントや実験手法が異なるものの中で数値を決めるということなので、そのときに数値が低いものを選ぶという考え方と、まずは実験の質をもっと評価して、実験の質が高いものを優先的に使うとかという、その辺の選択が新たに出てくるかもしれないです。また別の悩みが出てきたかもしれないのですけれども、ちょっとその構成を、きょうの議論だと結果的には一つの流れに収れんしそうなので、そのときに注釈の形で少し複数のものを取り扱うとかとなりますし、やってみていろいろな結果が出てきたときに、質を優先した選び方、質を加味して選ぶのか、それとも数字の最小だけで選んでいくのかというところが議論になっていくと思うのです。

どうしてもサイエンスマインドが強いので、信頼性の評価をして、やはり信頼性の高い、極端な言い方をすると質が一番高い実験のデータのBMDLを採用するという考え方もあるし、どっちみち全部不確実性をいっぱい含むので、これは最小のもの、値が一番低いものを選ぶべきだという考えもあると思います。その辺を今後、道筋をつくらないといけないかなと。質と数値をどうバランスさせるかという問題があるかもしれないです。

〇事務局 また事務局で文章を整理したいと思っています。今のお話を一旦整理させていただくと、基本的にはデータセットについて、案2にあるような、書き方はありますけれども、最終的にはデータセットをもとにした真のモデルといたしますか、よく説明できているモデルをもとにBMDなりBMDLを算出するという流れ自体はありつつも、先ほど〇〇からお話があったように同質性が期待できるようなデータセットが複数あった場合について、質を加味してBMDなどを選ぶとか、BMDLの値、その他のBMD関連指標をもとに選んでいくのかというところは、また研究のデータも待ちつつというところでよろしいでしょうか。

〇〇〇 そうですね。研究班の成果によって、例えばあるモデルは比較的、安定的にいい精度で出してくるというようなことがもし得られたら、それを優先的に使う。データセットに当てはめて、逸脱になってしまったりしたらはねることになるので、無批判に使うという意味ではありませんけれども、このモデルがいろいろな意味でロバストであって、逸脱が少なくてというような一定の報告が得られたら、それを加味するということです。

でも、そうでない限りは結果を見て選ぶということなのだけれども、質の評価との折り合いをどうつけるかというのがまだはっきりと定まってはいません。

〇事務局 何度も申しわけございません。同質性が期待できるデータセットが複数あるというのは、どちらかというところオプショナル的な位置づけとしつつ、さらに案1で示されているような複数のデータセットに対して共通である意味普遍的なモデルを選択する基準というのは、数字的に何か具体的なものをもって設定するのは、今お伺いしている中で〇〇からの御指摘があったようになかなか難しいのかなと。やはりそこはエキスパートジャッジになってしまうのかなという印象を受けたのですけれども、そういったことで。

〇〇〇 そうですね。その前に、シミュレーション、班研究の結果で何か重要な示唆があれば、それを酌んで、それにプライオリティーを与えることがあるかもしれないけれども、そうでなければ結果を見て選んでくるという道しかないでしょうから、何か注釈が、こういう条件のときはこれが入りますよというのが、まだ何を入れるかは決まっていなくても、もくもくというマークだけがつくといいますか、入るかもしれないと。そんな感じです。

〇〇〇 1個いいですか。シミュレーションしながらの予告的な話なのですけれども、単一の動物実験のときも複数のときもそうなのですが、こうやってモデルがたくさんあって、どれがいいというのが決まらないので、今、世界の、特にEFSAでのトレンドはモデルアベレージングになっています。モデルアベレージングのソフトウェアさえ出てきているのです。それで、エキスパートジャッジで決められないぐらい近いモデルが多くなって、BMDLも多くなるので、よいモデルを平均化しようという結構強い動きが研究で進んでいて、シミュレーションをやっているでもそうやっている平均化がよい場合が多いなという結果が出がちなのが、今、自分の見ているところなので、そういったこともあると想定しながら考えていただければと思います。

〇〇〇 ありがとうございます。すごく大事なほのめかし、ほのめかしを超えた御示唆だと思いますけれども、結局、どれをやってもかなり近接してしまうという現実を用意していただいたもの、今までの経験でもそうなので、最終的には全てをのみ込む形のアベレージングとなってくる可能性もあると。正式に今決めるわけではありませんが、そうすると、研究班のお勧めとしてアベレージしたものが一番いいですよということがあればそれを優先的に使うことになる。もちろんその際も、個別のものはやってみて確認をするということは要るかもしれませんが、結果的にアベレージングを採用するということは十分に可能性があるもので、それは一つのレコメンデーション。研究班からのレコメンデーション

ンとして、特定のモデルではなくて、全体を平均したものが一番使い勝手の面とか安定性の面によろしいという答申のようなものが得られたら、それを含めて記載していただくということになるかと思います。平均化もモデルの一つとしてね。

○事務局 ありがとうございます。

○○○ ほかにいかがでしょうか。

収束したのやら、散ってしまったのやら、よくわかりませんが、ただ、○○の班研究で既に着手されているところから重要な御示唆も得られました。研究の質と数値との折り合いとかその辺はまだ、きょう新たに出てきたところもありますけれども、意見交換の中で一つの筋道がより明確になったような気もいたします。

この方向で事務局にもう一度まとめていただいて、私も深くコミットしていきたいと思えますし、私自身も整理し直しますので、一緒に作業をしていただきたいと思います。特にその際に注意すべきこととかはありますでしょうか。よろしいですか。

それでは、きょう出した方向性をもとに事務局にお骨折りいただくこととなりますが、よろしく願いいたします。

では、この後は事務連絡等で、事務局にお返しします。

○○○ 承知いたしました。

現時点で、12月14日を開催予定としておりまして、先生方の御予定をいただいておりますが、次回ワーキンググループにつきましては、事務局での作業状況も踏まえ、改めてお知らせいたします。

○○○ それでは、今日の議事はこれにて終了いたしました。ちょっと時間を超過してしまいましたけれども、長時間にわたり、どうもありがとうございました。