

# 食品安全委員会評価技術企画ワーキンググループ

## (第25回) 議事録

1. 日時 令和4年7月20日(水) 14:00~16:56

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) 疫学研究で得られた用量反応データへのベンチマークドーズ法の適用について
- (2) 令和2年度~令和3年度食品健康影響評価技術研究成果の報告等
- (3) その他

### 4. 出席者

(専門委員)

川村座長、赤堀専門委員、猪飼専門委員、小坂専門委員、澤田専門委員、  
寒水専門委員、祖父江専門委員、竹内専門委員、西浦専門委員、松本専門委員、  
山田専門委員

(専門参考人)

広瀬専門参考人

(食品安全委員会)

山本委員長、浅野委員、川西委員、脇委員

(事務局)

紀平評価第一課長、前間評価第二課長、寺谷評価調整官、井上評価情報分析官、  
爲廣課長補佐、今木評価専門官、石神技術参与

### 5. 配布資料

資料1-1 前回のWGでのご議論を踏まえた疫学研究データにおけるBMR設定に関する  
記載案

資料1-2 第24回評価技術企画ワーキンググループ資料(BMRの設定に当たって考慮  
すべき点)

資料2 令和2年度~令和3年度 食品健康影響評価技術研究「ベイズ推定を活用  
したベンチマークドーズ法の評価手法検討と国際動向の研究」研究成果発  
表資料

- 参考資料 1 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 [動物試験で得られた用量反応データへの適用]  
(2019年10月食品安全委員会決定、2022年6月改正)
- 参考資料 2 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～  
(2018年7月食品安全委員会評価技術企画ワーキンググループ)
- 参考資料 3 評価技術企画WGにおける今期の進め方について
- 参考資料 4 第24回評価技術企画ワーキンググループ資料 (疫学研究データにBMD法を適用する際のBMRの設定についての既存情報)

## 6. 議事内容

○今木評価専門官 それでは、定刻となりましたので、第25回「評価技術企画ワーキンググループ」を開催いたします。

先生方におかれましては、お忙しい中、御出席いただき、ありがとうございます。

本ワーキンググループは、原則として公開となっておりますが、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、新型コロナウイルス感染症の蔓延防止のため、ウェブ会議形式で行います。なお、このような事情から、本日は傍聴者を入れずに開催することとし、本会合の様子を食品安全委員会のYouTubeチャンネルにおいて動画配信を行っております。通信環境等から議事進行に支障が生じる場合もあろうかと存じますが、何とぞ御理解のほどよろしくお願い申し上げます。

また、内閣府において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。

それでは、以降の進行につきましては川村座長にお願いしたいと思います。よろしくお願いいたします。

○川村座長 皆様、こんにちは。本日の評価技術企画ワーキンググループを始めたいと思います。

本日の議題は、「疫学研究で得られた用量反応データへのベンチマークドーズ法の適用について」と「令和2年度～令和3年度食品健康影響評価技術研究成果の報告等」です。

事務局から資料確認をお願いいたします。

○今木評価専門官 事務局でございます。資料確認の前に少し連絡がございます。事務局に人事異動がございましたので御報告させていただきます。

近藤に代わりまして、紀平が評価第一課長として着任しております。

○紀平評価第一課長 評価第一課長に着任いたしました紀平と申します。よろしくお願いいたします。

○今木評価専門官 石岡に代わりまして、前間が評価第二課長として着任しております。

○前間評価第二課長 こんにちは。評価第二課に7月7日に着任いたしました前間です。よろしくお願いいたします。

○今木評価専門官 高山に代わりまして、寺谷が評価調整官として着任しております。

○寺谷評価調整官 寺谷です。よろしくお願いします。会った方も何人かいらっしやると  
思いますけれども、どうぞよろしくお願いいたします。

○今木評価専門官 続きまして、いつものお願いではございますが、本会合はウェブ会議  
形式でございますので、御発言等に当たってお願いしたい事項が3点ございます。

1つ目、発言者の音質向上のため、発言しないときはマイクをオフにさせていただくよう  
お願いいたします。

2つ目、御発言いただく際、赤い挙手カードを御提示いただくか、ウェブ会議画面上の  
挙手ボタンを押していただきまして、発言の際には、冒頭にお名前を御発言いただいた上  
でお願いいたします。発言の最後には「以上です」と御発言いただきまして、マイクをオ  
フにしてください。また、座長より全員の先生に対して同意を求める場面があるかと思  
いますが、同意をする場合には手で大きな丸をつくっていただくか、青い同意カードをカメ  
ラに向けていただければと存じます。先生方のリアクションを見ることができるよう  
に、カメラも常にオンにさせていただければと思います。

3つ目、音声接続不良など通信環境に問題がある場合は、カメラをオフにすることや再  
入室により改善する場合がございます。マイクが使えない場合は、ウェブ会議システムの  
メッセージ機能により事務局まで御連絡いただきますか、全く入室できなくなってしま  
った場合には、事務局までお電話いただきますようよろしくお願いいたします。

以上、御協力のほどよろしくお願いいたします。

続きまして、資料の確認をさせていただきます。

本日の資料は、議事次第、専門委員名簿のほか、資料1-1、資料1-2、資料2の資  
料3点と、参考資料1から4までの4点をお配りしております。

不足等がございましたら、事務局までお知らせください。

以上でございます。

○川村座長 すみません。ちょっと進行表に印刷漏れがありまして、失礼いたしました。  
抜けたことをお伝えしておきます。

本日は11名の専門委員の先生方が御出席でございます。

竹内専門委員は所用のため14時45分まで、赤堀専門委員、小坂専門委員、祖父江専門委  
員は15時まで、寒水専門委員は16時40分までで退席されるとお聞きしております。

また、専門参考人として広瀬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名の委員が出席しておられます。

ウェブ形式なのでやりにくいところもあるかもしれませんが、率直な意見交換を  
してまいりたいと思っております。

資料は今御説明があったとおりですが、まず初めに、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」に基づいて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。

○今木評価専門官 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御報告いたします。本日の議事につきまして、専門委員の先生方から御提出いただきました確認書を確認しましたところ、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でした。

以上でございます。

○川村座長 ありがとうございます。先生方、御提出いただいた確認書について相違はございませんでしょうか。よろしいでしょうか。

ありがとうございます。

それでは、本題の議事のほうに移ります。1つ目の議事につきましては「疫学研究で得られた用量反応データへのベンチマークドーズ法の適用について」です。前回のワーキンググループの議論を踏まえまして、事務局で文案を作成していただきました。事務局から説明をお願いいたします。

○爲廣課長補佐 本日のワーキンググループでは、前回から引き続き、BMRの設定に当たって考慮すべき点について御審議いただきます。

では、資料1-1、前回のワーキンググループの御議論を踏まえた疫学研究データにおけるBMR設定に関する記載案と、資料1-2、BMRの設定に当たって考慮すべき点を御用意ください。資料1-2は前回5月の第24回評価技術ワーキンググループでお配りした資料になります。前回のワーキンググループで専門委員の先生方に御議論いただきました内容を事務局がまとめ、疫学研究データにBMD法を適用する際のBMRの設定について記載案とさせていただいたものが資料1-1となります。検討の際にまとめた国内外のガイドラインにおけるBMRに関する記載状況に関しましては、参考資料4としてお配りしてございます。

それでは、資料1-1を御覧ください。まず、5行目から12行目までは疫学研究データに対するBMRの設定について、二値データにも連続値データにも共通する考慮事項についてまとめたものでございます。1段落目に記載してございますように、疫学研究で得られたデータにBMD法を適用する場合のBMRについては、毒性学上または臨床上もしくは公衆衛生上の観点から専門家が設定すると記載させていただきました。

7行目以下には、具体的な考慮事項として、その際、対象とした健康影響等の種類及び重大性、観察された反応率、バックグラウンドの反応率及びそのばらつき等を踏まえた統計学的検出力及び用量反応データの特性等について考慮する。また、動物試験等で得られた結果がある場合は併せて検討する。なお、検査値等の代理指標を用いる場合は、ヒトの健康に対する臨床的な意義や公衆衛生上の意味を十分に考察するとしております。

ここまでの疫学研究データを用いる場合のBMRの設定に際し、二値データにも連続値データにも共通する考慮事項として事前に専門委員の先生方からコメントをいただいております。猪飼専門委員からは、11行目の「ヒトの健康に対する臨床的な意義や公衆衛生上の意味を十分に考察する」とあるとおり、疫学データを用いた評価の進め方についての記述

が抽象的であり、読み手にとっては大変難しいメッセージにならないかが心配ですと御意見いただきました。

寒水専門委員からは、9行目に記載がございます「データの特性」より「データの特徴」という表現のほうがしっくりきますと御意見がありました。こちらにつきましては、事務局としましては、特徴と特性は、どちらもほかと異なる性質という意味がございますが、特徴には目立つ性質という意味合いがございますので、ここでは特性とさせていただければと考えております。

祖父江専門委員からは、5行目の記載で、疫学研究で得られたデータにBMD法を適用する場合のBMRについては、毒性学上または臨床上もしくは公衆衛生上の観点から、専門家が設定するとなっておりますが、臨床・公衆衛生の専門家がBMRを正しく理解して値を設定することはかなり難しいと思いますと御意見いただきました。

その上で、現在、動物実験・疫学研究から観察されているNOAELの値から、アウトカムの頻度・サンプルサイズの組み合わせで逆算されるBMRを、幾つかのばく露・アウトカムの組み合わせで試算してみると、臨床・公衆衛生上の専門家からの理解が得やすいかと思わずとコメントいただいております。

一旦ここまでの部分について御審議いただければと思います。よろしく願いいたします。

○川村座長 ありがとうございます。ただいま事務局から御説明がありましたように、ここが総論的なところといえますか、基本的な考え方なのですけれども、BMRというPODを設定する際に非常に重要な指標を決めるに当たって、基本的には物質の毒性学的な観点、臨床もしくは公衆衛生の観点を踏まえまして、専門家が議論して決定するという趣旨でございます。

毒性学的という文言は、多分その物質が持つ、動物実験とか今までの世の中で起きた事件とかそういったものを踏まえた物質としての特性というところにフォーカスがあるように思いますし、臨床というクリニカルの中にはインディビジュアル、一人一人を見るという観点が強くにじむ言葉ですので、1人においてどのような確率的な影響があるかと。それから、公衆衛生というのも少し広い、パブリックですので日本全体あるいはコミュニティー全体でどのような健康影響が想定されるかといった視点の置き方なのですが、そういった様々な、しかし、重要な観点から専門家が議論するということになります。

ここには総論的な話ですので用語しか出ていませんけれども、具体的には、今申し上げたように物質的な面、それから個々の人、生活者でどれぐらいの影響があるか、あるいはは集団としてどのぐらいの影響があるか、どういう影響があるかというようなことを考慮することになるかと思えます。これは視点の問題です。

次は健康影響の種類と重大性ということで、これは生きる、死ぬといったクリティカルな問題から、検査値の変化といった間接的な、あるいは中間指標の変化まで様々なものがあって、それぞれBMRを決めるときは、BMRの持つ意味を考えれば異なるであろうと。設定

値も異なると思いますので、そういった健康影響の種類や重大性によってBMRを具体的に決めるに当たってその点を考慮するということ。

それから、さらに、バックグラウンドの反応率。つまり、無ばく露、厳密に無ばく露かどうかは難しい点がありますけれども、基本的に最小ばく露の領域においてどのくらいのアウトカムの発生があるか。しかも、そのアウトカムの発生が対象の集団によってどのくらいばらつくかといったようなこと。つまり、何もなくても起きる現象の多様性といったものを考慮するということ。さらには、統計学的な検出力や計画的な用量反応関係の特性について考慮するということになります。

動物実験で得られた結果がある場合は、併せて検討すると書かれています。むしろ私は動物実験の結果が初めにありきではないかと思っております、疫学研究で検討するといっても、今まで何回も議論してきましたように、様々な交絡があるとか、あるいはばく露についても、ばく露総量を測ることは現実には難しいので、部分的なばく露を切り取っているということから、これだけで因果の関係を断定することは非常に難しい。特殊な例を除いて困難ですので、動物実験のデータを踏まえて、それが無い場合は、例えばAIの力を借りる。(Q)SAR、リードアクロスのようなものの力を借りて、相当の因果関係が推認される場合において疫学研究データの利用に入っていくというふうに考えておりますので、むしろ動物実験はベースにあるのかなと考えております。

最後に、検査値等の代理指標の注意書きが1つ付記されておまして、十分それは意味があるものであると。単なる反応を見ているのではなくて、それ自体に意味があるものということを付記しております。

かなりエッセンスは詰まっていると思いますし、皆さんの御意見も十分に踏まえて事務局、それから私のほうでもんだものがございますが、しかし、言葉ですのでいろいろ行き届かない点、あるいは誤解を生むことがあってはいけないので、皆様の御意見を伺いたいのですが、いかがでしょうか。広瀬先生、お願いします。

○広瀬専門参考人 どうもありがとうございます。確かにBMRというのは以前、どこの会議だったか忘れたのですが、WHOとかの専門家の会議でもここをどう考えるというのはすごく長く議論をしていたというのを紹介した記憶があります。大変難しい問題だというのは理解しつつ、ただ、具体的に書くのはなかなか難しく、猪飼専門委員の言うように抽象的にならざるを得ないところもありますし、逆に祖父江専門委員からの指摘のように、時にはNOAELイコールBMDLという極端な仮定を置いてやるということは別にして、多かれ少なかれ、BMDLがNOAELに結果として今まで特に動物実験の場合は近かったということがあるので、イメージとしてはそういうこともあるのかなと思います。

しかし、順当に考えると、統計学的な検出力と毒性学的、あるいは生物学的な反応性の意味と量の大きさとか、そういったものとの兼ね合わせというか、両方兼ね備えているところで取らざるを得ないと思うのですが、それもそうきっちりと書くことができないといったところから、具体的方法を記述するとそれに固まってしまうところもあるし、

ある程度専門的判断というか、専門家のさじ加減をなるべく少なくするためということもあって、ベンチマークドーズによる数理モデルが出てきた経緯はあるのですが、それにしてもやはりクリティカルなところは、様々な観点を総合して検討するという表現にするしかなく、私としてはこの文章でエッセンスが詰まっている表現以外の表現で書くのはなかなか難しいのかなということと、これでいいのではないかと思います。

以上です。

○川村座長 ありがとうございます。一番ベンチマークドーズに今までの御経験がある広瀬先生からの御意見なのですが、ここにも少し苦渋がにじみ出ていると言うと変ですけども、やはり書き切れないのですね。ですから、私としては、言葉としてはこういうふうにさらっと書くしかないのですが、こういう観点から専門家が設定するという書き方にはなるのですが、これで終わりにするのではなくて、ここの議論で出たことを、具体的な留意事項をきちんと覚書みたいに記録しておいて、それについて具体的に特定の物質について決める際にチェックリストみたいにして、この点は議論したかどうかとか、この点は重視すべきかどうかとかいうようなことをきちんと確認する作業が要ると思うのです。

ですので、私も先ほど断片的には触れましたけれども、健康影響がこの物質、このアウトカムにおいてはどのくらい重要なのか。もっと重要なものがあるのか。そうではなくてどこまでこれだったら許容できるのかとか、そういった様々な一つ一つの点を覚書のような形でリストアップしておいて、それをきちんと潰していくというような作業が必要だろうと。そういう実施に当たっての実施要領みたいなものを多分、この文章の下に来るべき実施のマニュアルみたいなものについては、具体的にチェックすべきポイントを列記していく。それはどんどんやっていくうちに増えていくだろうし、あるいは不要ということで抹消されるものもあるかもしれませんが、そういったことをきちんと留めおく、確認をするということを進めていくことになるのかなと。

経験の蓄積も必要ですし、諸外国もそんなにすんなりきれいにやっているわけではないだろうし、そちらの経験も絶えず横目でにらみながら、日本で導入すべき点は導入しますしということで、ここは本当に骨子だけで、具体的なチェックポイントは列記していくべきではないかと考えているところです。

お願いします。

○澤田専門委員 がんセンターの澤田です。確認したいところとして、BMRについて、とても重要な健康影響だから1%の変化にしようとか、微小な変化だけれども人体には影響があるから10にしようとかいうのを決めるという意味で合っていますか。

○川村座長 それも含むというところです。

○澤田専門委員 やはり何を指しているのかが分かりづらいというのがありまして、BMRについて毒性学上または臨床上もしくは公衆衛生上の観点から、何を専門家が設定するか

という、その「何を」がよく分からないというのがこの最初の2行目にあります。そこを明確にできないということでしょうか。

○川村座長 いや、目的はベンチマークレスポンスを設定する。

○澤田専門委員 なので、例えばそれもということであれば、ベンチマークレスポンスの大きさ、とかをというふうには書けないという感じでしょうか。などをとかでもいいのですけれども。

○川村座長 目的はベンチマークレスポンスを設定したいわけですが、これが一義的に決まらないので、あれも考え、これも考えて、専門家がああだこうだといって、どこに落ち着くかということですね。だから、設定するものはBMRなのですけれども、BMRが一義的にこれだったらこれ、これだったらこれと決まらないので、例えばモータリティーを問題にしているとか、疾患のインシデンスを問題にしているとか、検査値で数値が何%動いたとか、いろいろなレベルのアウトカム指標がある。

二値のものでもそういうクリティカルなものから軽度の、病気には違いないかもしれないけれども、直ちに生命を左右するものではないといった、アウトカム自体にすごく重い軽いがあるし。

○澤田専門委員そこは分かります。それは多分、公衆衛生上とか毒性学上というところで表されていると思うのですけれども、それによって何を、BMRの01とか05、10を設定するために、そういうふうに言いたいのか、そもそもの値なら値、または幅とか大きさ、そういうのは書かなくても、誰もが分かっているのならそれでもいいのですけれども。

○川村座長 目的のBMRなので、これが例えば動物実験であれば10%だったのですよね。

○澤田専門委員 なので、人間だったらそれが10%ではないかもしれないというところも分かります。毒性学上とかということ。だから、10%とか5%というのを決めるという意味ですか。

○川村座長 そうそう。

○澤田専門委員 それを書いてもらったほうがよくないかなと思うのが私の意見です。

○川村座長 どういうふうには書けば分かりやすくなりますか。

○澤田専門委員 なので、BMRの幅なのですかね。反応の幅。

○川村座長 幅というか、ピンポイントで幾つにするとかなのですけれども。

○澤田専門委員 01とかの強さのレベルを決めるというのですか。

○川村座長 だから、何%かということを決める。

○澤田専門委員 それを書いたほうが良いような気がしますけれども、どうなのでしょう。

○川村座長 BMRというのは何%という意味なので。

○澤田専門委員 これがBMRに含まれているということで、不要だということですね。

○川村座長 そう思いますけれども。

では、ちょっと応援をお願いいたします。広瀬先生。



○広瀬専門参考人 澤田先生の御意見はもつともで、この文章だけを見るとBMRは何か分からないのですが、これは多分、最終的にできる文章の一部で、BMRをどうするかというところだけを先につくっているんで、動物の指針の最初の目的を見ていけば分かるのですけれども、目的はベンチマークドーズ法を使ってその一つのBMDLという指標を元にPODを決めることなのです。要するに動物の場合はNOAELから不確実係数を使って行政指標を決めるのですけれども、そのときのNOAELはPOD値として、要するに行政指標を決める前段階の基準点ですね。それをBMDLで設定しよう。だから、そのBMDLに決めるためにどんな反応率を選ばいいか。最終目的はもちろん基準値を決めるのですけれども、それに使うためのBMDL、PODを設定するためのBMRはどうあるべきか。その文章は多分別のところからつくられると思うので、ここではエッセンスだけ、BMRだけの議論が入っているので、その説明はここに入ってこないのですけれども、多分それはこの文章全体がもっと出来上がった段階では当然入れるべき内容だと思います。

以上です。

○澤田専門委員 ありがとうございます。この部分では必要がないということですね。その後の資料にも、BMRの設定についてリスクがその幅で上昇することによって国民全体にどう影響を与えるのかということをもっと説明できるように議論を深める必要があることを確認したというのがあるので、リスクの幅とかがあったほうがいいのかと思った次第ですが、ここで必要ないのであれば、分かりました。

○川村座長 ありがとうございます。後でまとめますね。

赤堀先生、御意見を先にお願いたします。

○赤堀専門委員 今、澤田先生のお話を聞いて、確かにそうだなというふうに思いました。BMR、レスポンスの大きさの話をしているのかなと思う一方で、それをどの指標にするのかということもここで決めるわけではない。要するに、一つの疫学データの中に、例えば死んでしまふとか幾つかのエンドポイント、表現型で、そのどれを取って、その上で10%にするのか、5%にするのかというのが本来あるべき議論なのですけれども、ここで今やろうとしているのは、エンドポイントそのものは含まなくて、あくまでRateだけという理解なのですかね。

前提条件なのですけれども、今ここで文書化しているのは、例えば1個しか疫学データがなければそれはそれでいいのですけれども、複数疫学データがあって、多分それぞれにBMRの設定が必要で、それぞれに必要なBMRの話について整理をしているという理解でいいですか。今さらのところすみません。

以上です。

○川村座長 ありがとうございます。当然のことながら、アウトカム、エンドポイントごとによってその意味が違うので、一つのデータセット、一つのアウトカム指標についてBMRが決まる。同じ一つの研究の中でも、これは澤田先生もよく御存じのようにいろいろなものをアウトカムにして研究成果が公表されます。だから、同じ疾患であってもインシデン

スで表現する場合、モータリティーで表現する場合、どちらもあるので、それぞれBMRは異なると思います。

だから、何をアウトカムにするかということは、データセットの選択と並んで、一つのデータセットの中でもいろいろなものが並列、併記される。基本的にはできるだけメインにするものは1本という原則はあるのですが、こういう観察研究では介入試験と違って事前に1本に決め切れなくて、いろいろものについて様々な検討をする。循環器系疾患の中でも心疾患であったり脳卒中であったりといろいろ分けますし、アウトカム自体がメインアウトカムメジャーとセカンダリーアウトカムメジャーズというふうに論文に書くことも多いので、そうなるとうBMRも実は一つの研究であっても幾つも出てきてしまうという可能性はあります。

なので、そういう点が不明確というか、疫学研究に伴う問題としてあると思いますが、分かりにくい点は表現をもう少し検討して、容易に分かるように工夫したいと思っておりますけれども、行き届かない点がありまして申し訳ありませんでした。

澤田先生、何とか。

○澤田専門委員 はい。今の議論は分かりまして、この前後に盛り込まれた全体としてはきっと分かるようになるのだらうなと思っておりましたので、ここの文章に無理やり入れてくださいという要望ではないです。必要なければよいです。分かりました。

○川村座長 ありがとうございます。広瀬先生がおっしゃったように、全体の流れがあって、ベンチマークドーズを決めるに当たって、ベンチマークレスポンスは避けて通れない要素なので、これを決めないとベンチマークドーズが決まらない。つまり、最終ゴールは社会的な規制、政府による規制値を決めるというのがゴールですけれども、その前に、その規制値を決めるための出発点になるような値を決めるのが食品安全委員会の報告書がすべきことで、それがポイントオブデパーチャー、PODと言っているものです。そのPODを決めるに当たって、ベンチマークドーズを使う方法とか、NOAELという統計学的有意性を使ったものがあるので、ここではベンチマークドーズを使って、どこからを問題ありとする領域にするかというのがベンチマークレスポンスになりますので、そのベンチマークレスポンスの数字を1%であるとか、0.1%であるとか、そういうことを決める作業は避けて通れない。そのベンチマークレスポンスの値を決めるということで、ちょっと言葉を補う必要があるかもしれませんね。御指摘ありがとうございます。

○澤田専門委員 すみません。ついでに今のお話でもう少し分かりやすくなるとするならば、PODを決めるためにというのも最初にあったほうが良いような気がして、PODを決めるために疫学研究で得られたデータにBMD法を適用する場合、毒性学何とかの観点からBMRの値を専門家が設定するというのが分かりいいかと思っておりました。

○川村座長 では、事務局のほうで文章をつける。ちょっと前後の文章との絡みがあるので、全体としてそれがちゃんと分かるようにしますね。

○澤田専門委員 お任せします。

○川村座長 ありがとうございます。

ほかに御意見いかがでしょうか。

○今木評価専門官 事務局でございますが、補足させていただいてもよろしいでしょうか。

○川村座長 お願いします。

○今木評価専門官 先ほど広瀬先生からも御説明いただいたとおり、今御議論いただいているのが、指針の骨子案となる一部分でございますので、少し全体像が捉えにくいところがあるかと存じます。先ほどチャットで猪飼先生からもお送りいただきましたが、参考資料1としまして、動物試験にBMD法を適用する場合の指針がございます。こちらがベースになって議論が進むものと思いますので、基本的には動物の指針と同じように、冒頭に目的、定義というような形でBMR等については記載がされると考えているところでございます。

以上でございます。

○川村座長 ありがとうございます。

では、全体の構成の流れでここがすんなり理解できるように事務局と一緒に考えますので、また出来上がったものをもう一度確認をお願いします。

では、猪飼先生、お願いします。

○猪飼専門委員 猪飼です。私はコメントで抽象的と書いて、改善の提案も書いていないというちょっと失礼なコメントをつけてしまいましたが、このガイドライン、恐らく参考資料2をベースに、今度は疫学データを踏まえた場合のところを增強される、もしくは別のものを並行してつくられるという位置づけだと理解していますが、これは規制当局である食安委がどのような判断でこの審査を進めるかという指針なのか、それとも、化学物質を製造する側も再現性を持ってこういう審査がされるから通るはずだということが十分予測できるような審査のプロセスを書き下すものなのか。どちらの目的かということによってちょっとその書きぶりが変わってくるかなという気がしました。

その結果として、川村先生が前半におっしゃったように、こういうチェックポイント、論点があるのではないかということを書き添えるというのが現実的な解決策だと思えますし、書き切ることが当然できないと思うのですが、我々がこんなことで気をつけてやっていますということを書き添えるだけであれば、このままでよいと思えます。もしこれが誰から見ても、こういうルールで話をすれば、我々の予想した結論を規制当局も下したという納得感まで出そうと思うと、澤田先生がおっしゃったように大分丁寧に、我々が何をやっているのかということを書き下していかなくてはいけないのではないかなと思った次第なので、その御整理をいただければと思います。

以上です。

○川村座長 ありがとうございます。大変重要な点だと思います。やはり国の機関として、具体的な規制は管理機関である農水省とか厚労省が行うことが多いわけですがけれども、そのための基準となるサイエンスベースの議論の結果をちゃんと伝えるということなので、結果的にはちゃんと社会規制値につながることでありますから、これは製造業者さんであ

れ、食べる側の一般国民であれ、影響を受ける指標ですので、審議のプロセスが明確で公表されているということは絶対必要条件になります。おっしゃるとおり何をやっているのか分からないということではいけないと思いますので、やったことはきちんと公開されるべきだし、何を重視したのかということも当然分かなければならないので、それは公表されるべき。実際に公表されています報告書には、今までの報告書が本当に議論の細かいところまで書かれていますし、会議録自体が公表されている。今のこの議論も公表されるわけですが、一応、何かあれば検証可能ということになっています。

ですので、当然、具体的に決めるに当たっては、何を重視したか、どういう観点から特に問題としたかというのは当然公表されますが、今、指針をつくるに当たって具体的にこの値はこうであるとか、これを最重視するとかいうことは書き切れないのでこういう表現になっている。でも、これで終わってしまうと何を言っているのだろうというふうに思うと思いますから、猪飼先生の御指摘はごもっともなので、それはこういった観定の例示としてきちんと書いておいたほうが良いと僕は思って、さっき事前の打合せでも事務局に長々とつぶやいておりましたので、拾ってくれると思います。

今日の議論もちゃんと問題点を拾ってくださいねということ強くお願いしておきましたから、やはりこういう形を、形式は国の機関としてのルールがいろいろあるので、私は任せようと思っていますけれども、考え方としてどういう観定で何を具体的な指標としてということが例示として分かるようなことはしておいたほうが良いのかなと思っています。

こんな感じでよろしいでしょうか。

○猪飼専門委員 ありがとうございます。

○川村座長 ありがとうございます。ほかに御意見はよろしいでしょうか。

それでは、一応、また後で議論の必要があれば戻りますので、次のほうに進めたいと思います。お願いします。

○爲廣課長補佐 それでは、説明させていただきます。続きまして、二値データを用いる際のBMRの設定に関する記載案について御説明申し上げます。BMD法を適用する用量反応データは主に（１）の二値データ、あるいは（２）の連続値データが用いられます。

14行目、二値データに関する記載につきましては、疫学研究は、研究対象や観察手法等が多様であり、動物試験データにおいて基本的に算出すべきとされる過剰リスク10%は必ずしも適当ではないので、データセットごとに専門家が設定するといったしました。

こちらは、本ワーキンググループで作成いただいた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、動物試験で得られた用量反応データへの適用、以下動物指針とさせていただきますが、それにおきまして、二値データではBMRとして過剰リスク10%を用いたBMD関連指標は基本的に算出するとされているため、記載してございます。

見え消し部分に関しましては、先生方に事前に御意見を頂戴する際にお送りしたのから変更となった部分になってございます。

二値データを用いる場合に考慮する点につきましては以上となります。ここまでの部分について御議論をお願いいたします。

○川村座長 ただいまの事務局の説明に関しても、遠慮なく御質問とか御意見をお願いいたします。

○祖父江専門委員 祖父江ですけれども、発言していいですか。移動中なので雑音が入るかもしれませんが。

○川村座長 お願いいたします。

○祖父江専門委員 二値データのところで参考資料1の動物の場合の記述のところにNOAELとの連続性というようなことが記述されていると思うのです。参考資料1の4ページ一番下の辺り、二値データでNOAEL法を用いた毒性評価との連続性も考慮しというようなことが書いてありますね。動物試験の場合は10%というのが、もちろんそれが独自で意味のある値なのかもしれませんが、従来のNOAELの値と連続性があるということで皆さん選んでいるということであれば、疫学もそこから独立して存在するわけではないので、何らか従来のNOAELとの連続性というようなことをここに記述すると、そこから目安になるような量が計算できて、専門家のほうの理解も深まるのではないかと思いますけれども、どうでしょうか。前段階で私の言った意見の、場所がちょっと違うのですけれども、その補足であります。

以上です。

○川村座長 ありがとうございます。NOAELとの連続性というのはちょっと私自身も引っかかる場所があって、前から言っていることなのですが、実はNOAEL法の問題点を克服するためにベンチマークドーズを採用するわけですね。統計学の問題点。これは西浦先生もおっしゃっているし、広瀬先生、そのほか今まで扱った方は皆御存じのことですけれども、統計学的有意かどうかで決める方法というのは、差の大きさではなくて、差があることの確かさで決めているだけなので、これ自体が例えば不確実性というか、データの信頼性が低いとNOAELがどんどん上がってしまう。よりアウトカムが発生数が多くないといけないから上がってしまうというようなことがあって、不具合があるということで、NOAEL法の問題点を改善するためにベンチマークドーズを取るということですから、基本的にはつながっているということは、できれば破棄したいわけです。ただし、国がやる行政で用いる数字につながるものを決めるので、今までの経緯を無視するべきではないということがあるのではないかと。

私は直接行政を担当することはなかったし、これからもないであろうから、そんなくやし過ぎかもしれないけれども、しかし、国の責任として値を決めて、そしてその値が社会の規制として強制力を持っていくというようなことで、方法が変わったのでいきなり例えば桁違いに厳しい値になったりすると、今まで通っていたものが通らなくなるとかいろいろなことがあると思うのです。それから、国民の理解を得るに当たっても、今までの値と余りにも違うと何でなのとかいう疑問が来ると思うので、そういう意味で、社会的な、あ

るいは行政的なものを考えてNOAELとの関連というか、差異や近似性といったものは考慮するけれども、それにあまり引っ張られると、わざわざベンチマークドーズを採用した意味がなくなってしまうので、違うなら違うで、なぜ違うのかということはちゃんと説明しないとイケないし、近似していれば近似しているで、大きな差は生じませんでしたということ言えばいいわけですね。

ですので、ここは単なる研究の話ではなくて、行政施策を実行するために研究をどう利用するかというのがここでの審議事項なので、その説明をきちんとするという意味でつながりというか、連続性というところとちょっと適切かどうか分からないのだけれども、差異とか近似性も含めて従来のNOAEL法で得られたPODとの関係性について記載するということだと理解しております。

このような感じですが、事務局側としてはどうですか。

○爲廣課長補佐 事務局としましては、なかなか何%かというようなところを実際に記載するのは難しいところかなということがありまして、このような記載をさせていただいているところなのですが、川村先生から今御意見がございましたように、その詳細についてどのように考えて決めたかということについては、その考え方を何かしらの形でまとめて公表できるような形にできればとは考えております。

川村先生に先ほどからおっしゃっていただいておりますように、いろいろな対象とするエンドポイントによりまして、BMRの設定値は大きくなったりも小さくなったりもしますので、それが代替指標を使った場合であるとかになると、前回お配りした一昨年度の調査報告書資料等では、ホウ素の取り込みを基準値としてBMRを設定したものは50%というようなことも記載されておりましたので、一概にデフォルト値として何%とかいうようなところを決めることは、エンドポイント、何を指標にするかによって変わってまいりますので、その辺のところをうまく説明できるようにしていければなと考えております。

以上です。

○川村座長 ありがとうございます。

祖父江先生、今、車を降りられたと思いますが、聞こえていましたでしょうか。

○祖父江専門委員 聞こえていますよ。

○川村座長 何かまだ疑問は残りますでしょうか。

○祖父江専門委員 BMRを純粹にこういう条件で設定していくということだけだと、恐らく公衆衛生とか臨床の先生方はなかなかイメージが湧かないのではないかと思います。だから、これぐらいの量がNOAELの値に対応しますみたいなことがあると、もちろんそれを採用するわけではないのですが、それに応じて臨床的な意義とか公衆衛生的な意義で決めていく出発点になるのではないかと。議論のきっかけになるような目安がないと、5%、10%、1%となかなか値自体を決められないと思うのです。

○川村座長 おっしゃるとおりですね。BMRが持つ意味は何かという決める前の段階がきちんと伝わらないと理解できないと思うので、そのところはきちんと、報告書にはもち

ろん書かれますでしょうし、パブリックリレーションというのですかね。説明のときにきちんとそこも含めて説明されないといけない。要するに、国民の理解、あるいは業者さんの理解も得られないといけないことなので、広い意味での公衆衛生ですね。社会の様々な活動を含めた公衆衛生の目的にかなうように説明しないとけない。

特にベネフィットとリスクの兼ね合いを議論しなければならないことは必ず出てくると思うのです。メチル水銀の議論のときもありましたけれども、メチル水銀へのばく露量は少ないほうがいいけれども、栄養的な観点も考えないとけないよという議論は専門調査会の席の中で出てきておりましたし、単に毒性の物質を抑制するという観点だけではなかったように思いますので、そういう点でベネフィットというか便益と比較衡量とか、多分祖父江先生がおっしゃったのは社会全体でどのくらいの人が不利益を被るかとか、一人の人においてはどのくらいの確率で不運が起こるかとか、そういうことにつなげられるようにする必要があるのかもしれないと思います。

○祖父江専門委員 ベネフィットとリスクのバランスで考えるといったら、ベネフィットのほうの値があるので、それが目安になると思うのですけれども、今、それはないのですよね。リスクだけの話なのです。だから、余計に値を決めにくいというか、ゼロに近づけばそれはいいのでしょうか。

ただ、川村先生がおっしゃるように、あまりNOAELに引っ張られ過ぎるのはよくない。あるべき決め方で決める方法を今考えているのだということは理解できますので、ただ、連続性というか、今まで決めていた値とそんなかけ離れたものを決めていいのかというようなことがあるので、目安としてそれが頭にあると出発点になるのかなという意味です。

以上です。

○川村座長 ありがとうございます。やはり両者の連続性とか理解のしやすさということは重視したいと思います。

ほかはよろしいでしょうか。お願いいたします。

○西浦専門委員 西浦です。祖父江先生がおっしゃっていたことにちゃんと説明をしておかないといけないなと思って発言をすることにしたのですけれども、もともとの参考資料1にあった記述は何かというと、参考資料1の4ページ目の10%のところではNOAELとの連続性も考慮しという話があったのは、動物実験データで二値のアウトカムに関して検討する場合は、日本のこういうガイダンスとか指針だけでなく、ヨーロッパのEFSAとかでも10%というものについて具体的な言及があったものと記憶しています。EFSAではデフォルトとはっきり言われていないのだけれども、積極的お勧めみたいな感じで10%と書いているのですけれども、その理由として、もともと『Environmental Health Perspectives』に載ったようなNOAEL相当とBMDL10が近いですよという研究があるので、それを言及しつつ、動物実験の二値ではそれが書かれているということだと思いますというのが1つ。

それで資料1-1の二値データに戻るのですけれども、そういうことを考えていると、(1)の今読んでいるところは2点論点があると思うのですが、「過剰リスク10%は必ず

しも適当でない」の後に、例えば動物実験の評価結果とか、それはNOAELでもいいですし、動物実験データにBMD法を適用したのも参照するとよ的な付記がどこかにあってもいいのかもしれないと思いますということと、日本語の付記されているものは1ページ目もそうなのですが、**「データセットごとに専門家が設定する」**というのは実質、今の最後の議論にもちょっとありましたが、無理ですよ。なので、これは付記しなくていいのではないかなと思いました。これはp-valueを決定する際に、極端な話として**「私が神であり、私が今日からp-valueの閾値を7%とする」**みたいな感じのことを言うようなことに相当する話ですので、そんなことは多分できないと思います。価値判断まで要されることになるので、今、コロナで僕とか尾身先生が苦しんでいる話はそういう話です。こういうところで**「データセットごとに専門家が設定する」という話が残っていると、後で実装できなくなってしまう**ので。さっきの参考資料はいい日本語できていて、専門家の関与の下で検討するとかそんな感じのことになっているのですけれども、**「データセットごとに専門家がp-valueとかでゴールポストを一々動かしてリスク評価ごとに決めるといのは、やはり指針としてでもよろしくないものだな」という印象は持ちました。**

一意見です。以上です。

○川村座長 ありがとうございます。非常に大事な御指摘だったと思います。今までの経緯ですね。動物実験における経緯、あるいは具体的な専門調査会における審議の経緯などを踏まえての御発言だったと思いますが、**「どうして今10%というのがいきなりぽんと出てきているか」というような前提もあるので、確かにそういうつもりで読むと、いきなりぽんと出てくる。**一応この前段階に動物実験の指針があるという前提なのだけれども、全ての人**「動物実験の指針を読んでから疫学研究的指針を読むとは限らないので、これを単独で読んでも分かるようにはしておかないといけない」と**思います。

それから、**「データセットごとに専門家が設定する」ということなのだけれども、最終的に規制値をどうするかは専門家が決める話ではないし、リスク評価機関が決める話ではなくて、リスク管理機関が決めるものだろうと**思います。そのときには**「ベネフィットとの兼ね合いであるとか、コストとか、生産効率化の問題とか代替物質の存在とかいろいろ考慮されるはずですので、一義的には多分、我々がレコメンデーションとする数字には実際にはなりにくいし、放射線のときに私は体験したのだけれども、基準値は専門家で決めたのですが、実際に例えば農水省や厚労省では、それを食品に案分比例して物質ごとに規制値を決めるということをやっていたわけ」**です。そんなことはこちらでは全然言っていないのだけれどもなというふうに、**「リスク管理機関は実践に当てはめるためにいろいろ知恵を使ってやっておられて、食品安全委員会としては生涯の通算許容量と言うと変ですけども、一応PODを一つ決めて、生涯100という値を決めて、それをリスク管理機関につないだわけ」**です。そこでそれぞれの観点から、**「農水は農水で、厚労は厚労で自分の所轄の分野に当てはめて規制を決めている」と**いうことなので、最終的には**「食品安全委員会が実際の行政の値を決定するわけでは**



ないのだけれども、それにつながるちゃんと根拠のある数字をきちんと出していくということだと認識しています。

しかしながら、そういったプロセスが分かるということと、それから、専門家が決定するというよりは、確かに今までさんざん使ってきた専門家の関与の下にというのはうまい表現だと思うので、その表現については工夫の余地があるなと思いました。

西浦先生、いい解決ではないかもしれませんが、よろしいでしょうか。

○西浦専門委員 承知しました。

○川村座長 ありがとうございます。

それでは、二値と連続量という切り分けの問題というか、セクションの問題なので、続けて連続量のほうに進みます。

その前に、ごめんなさい。データセットごとという文章をどうするかということと、それから、データセットごとに決まるだけではなくてアウトカム、エンドポイントごとに異なってしまう。1つの疫学研究は動物実験よりもはるかにたくさんアウトカムについてやるので、複数のアウトカムが出てくる可能性がある。それもモータリティーであったり、インシデンスであったりという次元の違うものがあったりして、それぞれひょっとしたら違うかもしれませんね。ですので、データセットだけではなくて、転帰あるいはエンドポイント、どちらの表現を使うか。エンドポイントを割と共通に使ってきましたけれども、データセット及びそこで扱われているエンドポイントごとに専門家の関与の下で合理的な決定をするといえますか、ちょっとそれに似たような表現にしたほうがいいのかもしいないなと思いました。

この点も含めて、今のところ、こんな方向でまとめてよろしいですか。あるいは何かもうちょっと、いやいやこれはということがあったらお願いします。

それでは、ここで仮に置いておいて、次の連続値データのほうに進めさせていただきます。事務局、お願いします。

○爲廣課長補佐 事務局です。それでは、18行目、続きまして、連続値データを用いる際のBMRの設定に関する記載案について御説明申し上げます。

連続値データに関する記載につきましては、①から④の4つの項目に分けさせていただきました。まず①の生物学的に意味のある反応量の変化を、科学的根拠に基づいて設定できる場合は、動物指針に相当する1行目の記載に加え、個人における数値の変化と集団における数値の変化の違いに関する考慮について記載してございます。

また、②の生物学的に意味のある反応量の変化が不明な場合は、健康影響等の種類及び重大性、バックグラウンドにおけるばらつき、当該影響が生じるまでのMOA等を考慮しつつ、用量反応データごとに専門家が設定することと記載いたしました。

③、④は動物指針とほぼ同じ記載となっており、③では、科学的に合理的と思われるカットオフ値を設定できる場合は、ハイブリッド法等を取り入れるということ。④では、情

報量の減少を伴う点に留意する必要があると思いますが、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もあるということに記載いたしております。

前回のワーキンググループで御議論いただいております、サンプルサイズに関する記載については、コメント枠中に記載してございます。用いた疫学研究の質の一要素として、BMD法を適用するデータの選択に当たって考慮すべき点として記載することを検討予定としてございます。

本件につきまして、澤田専門委員から、例えば統計学的検出力は十分あったので、既報のサンプルサイズで問題なく評価できたとか、既報のサンプルサイズでは統計学的検出力が不十分であったので、評価の結果に注意が必要などのように統計的検出力といった表現を使うことは好ましいのではないかとコメントいただいております。こちらは対象部分を御議論いただく際に御検討いただくことになろうかと考えております。

連続値データを用いる際の考慮すべき点に関する記載についてもコメントいただいておりますので、御紹介させていただきます。

小坂専門委員からは、22行目の「多様な分布」というのが少し分かりにくい。また、いつも想定されるというわけではないので、「示すこともある」といった表現ではいかがでしょうかということと、25行目の「専門家が決定する」はリスクミとして、専門家が決めるというのを強調する必要はないのではないのでしょうかという御意見がありました。

寒水専門委員からは、20行目に「公衆衛生上の意味合いが異なる点を十分に考慮する」とありますが、これは何の変化かという定義の違いであるため「公衆衛生上の違い」というわけではないと思います。その際、「十分に考慮する」という部分は混乱や誤解を与える可能性がありますので、なくてもよいと思いますと御意見いただいております。

関連して、竹内専門委員からも、「公衆衛生上の意味合い」は誤解を生じないように記載する場所と文章を少し変えて、例えば19行目「その際」から21行目に「考慮する」までの代わりに、6行目の最後に「なお、公衆衛生学的な観点とは、個人的にはほとんど意味のないレベルの数値の変化でも、集団においては健康影響等のアウトカムの発生頻度が大きく異なってくることを十分に考慮することを示します」と記載するのはいかがでしょうかと御意見いただいております。こちらの部分につきまして、記載をどのようにするか御議論いただけますと幸いです。

澤田専門委員からは、15行目、16行目の二値データの説明として、二値データの例として、死亡または生存、腫瘍発生の有無など補足があると分かりやすいですということと、18行目初め、また38行目終わりなどに、連続データの例として生化学的検査、免疫学的検査等の補足があると分かりやすいと思われましたと御意見いただいております。

こちらにつきましては、指針の一部を議論しているため、全体像をつかみづらいところがあるかと思えます。二値データや連続値データの説明につきましては、動物指針と同様に用語の定義として記載することを想定しております。

山田専門委員からは、①が優先で、①から④の順に段階的に進めていくのでしょうかと御意見をいただいております。事務局としましては、①が②よりも優先されますが、①と③はどちらが優先されるというわけではなく用いる研究データにより適切なほうが選択されるものと考えております。なお、④につきましては、情報量が減少するというデメリットを考えますと、使用はかなり特殊な事例に限定されると考えております。

先生方からの補足等がございますようでしたら、お願いいたします。

そのほか、BMRの設定に関する論点から少し外れますが、山田専門委員より、選択された各BMRの根拠、選択の正当性を示し、文章化する旨のことは記載されるのがよいと考えますとコメントいただいております。こちらは今後、ワーキンググループで議論いただくことになると思いますが、事務局としましては、動物指針と同様、結果の文書化という項を設けることで対応できるのではないかと考えております。

また、広瀬専門参考人より、連続値データの④のところで個別データが利用可能な場合は、ばく露量も連続値データとして得られるので、二値化したデータとしての解析に用いるためには、ばく露量も区分化する必要があると思います。BMRの設定に影響する論点ではないのですが、区分の設定の仕方によってはBMD関連指標の制度に影響を与えるように思われました。この論点について現時点ではどのような表現で、ガイダンス全体のどの部分に記載するか、アイデアはございませんが、どこかに記載したほうがよいかと思いましたがと御意見いただいております。

御説明は以上となります。よろしくお願いいたします。

○川村座長 ありがとうございます。

連続量のほうは記載が少し多くなっておりますが、ちょっと補足で、広瀬先生はもう出られたかな。もう席を外されたかもしれないですね。ハイブリッド法などを村田先生と一緒に検討されて、レポートといいますか、総説みたいなものを出されていますので、留意点を聞こうかなと思いましたが、いらっしゃらないようなので、今御出席の皆さんで議論したいと思います。

今の事務局の説明につきまして、御意見ございますでしょうか。①から④の記載について、あるいは委員からのコメントについて。1つの論点としては、個人レベルの意義と集団レベルの意義が異なるということだと思います。例えば血圧で言えば、個人のレベルでは2ミリの違いというのはほとんど誤差になってしまいますけれども、日本全体で2ミリといえ、脳卒中の発生が一万何千人だったか、無視できない数のアウトカムの発生に差が出るということも分かっておりますので、それが臨床的意味、個人レベルでの意味と公衆衛生、集団レベルでの意味の違いだと思います。

そういう意味では、①の記載は公衆衛生上のというのは要らないかもしれないですね。個人の数値の変化と集団における数値の変化は意味合いが異なるとも言えるかもしれないと思いましたが、ほかに何か御意見ございませんでしょうか。

では、一つ一つ検討していきましょうか。まず、①、②、③、④は必ずしもこの順番でということではなくて、①は最もエッセンシャルで、生物学的に意味のある、あるいは臨床医学的に、公衆衛生学的にでもいいのですけれども、反応量の変化がはっきりと生命なり生存、生活なりに、ライフという点においてしっかりと意義が説明できるようなものがあれば、それをBMRとするということなのですから、その際に個人における数値と集団における数値は意味が異なるので、そのことは考慮しますよというメッセージが最初です。

②は連続値の用量反応データは多様な分布を示す。実を言うと多様な分布を示すのは連続量に限らない話で、通常のイベントの分布においても多様な分布を示すものなので、必ずしもこういった連続量のアウトカムに固有のことではないのですが、ただ、本当に統計的に処理し切れないものが結構あるかもしれないですね。ここに書いていないのだけれども、例えば正規化できるのであれば正規化、ノーマライゼーションというのを行って、できるだけ関数に乗るような形にするということは必要だと思いますし、例えば二峰性になってしまったとか、そのようなときにバックグラウンドがそもそも単一の集団かというようなことは大事かと思うのですけれども、時々種類の異なる集団を併合するという事も行われることがあるので、そういったデータの由来にも遡らないといけないのだけれども、そういうことをやってもなお一つの関数にできないような分布を示した場合にどうするかということはあるかもしれないです。

ここも専門家の判断で、要するにケース・バイ・ケースだから、あらかじめこういう処理をするということは決められないので、こういう場合があり得るから、そういう場合はその都度審議するという、要するにプロセスしか書いていないのですね。どうするかという考え方しか書いていないので、こういう歯切れの悪い書き方になってしまうのですけれども、このときに何か気をつけるようなことがあるかどうかですね。この辺りはどうでしょうか。

寒水先生は、例えば分布の問題なんかにはどう向き合うか、御意見はお持ちでしょうか。  
○寒水専門委員 私がコメントしたところを先に説明してもよろしいでしょうか。①の個人における数値の変化と集団における数値の変化というところは川村先生に説明していただいて、多分皆さんすっかりしていると思うのですけれども、この文章を読んだときに、趣旨をすぐに理解できなかつたので、誤解とか混乱のもとになるのではないかと思いコメントさせていただきました。私の専門が統計学ですので、確率分布を集団について考えたときに、個人における数値の変化は一個一個のサンプリングした値で、それは当然ばらついた値になるわけで、ここでいう集団における数値の変化というのは確率分布の母平均そのものが変化するということかと思えます。BMRは集団における値と理解しているので、個人の数値と集団における数値の区別というのがBMRの設定における注意書きとして必要なかというところが気になっていました。そもそも集団における数値の変化としての設定される値だと思えますので、そこに個人の数値のばらつきの変化というものがどういう余地で考慮されるか理解できなかつたわけです。

○川村座長 ありがとうございます。了解しましたというか、意味が分かりました。要するに、BMRを含めてベンチマークドーズというか、食品安全委員会の仕事は公衆衛生の観点でやるものだから、個人の検査値の変動を直接問題にしているわけではなくて、日本全体としてどのくらいのハームなりベネフィットなりがあるかということを議論しているのだから。

○寒水専門委員 ばく露を受けたグループとばく露を受けなかったグループの平均値の違いを議論しているはずなので。

○川村座長 そうですね。もともと集団レベルの考え方だから、さっき例に出した血压で言うと、血压を毎日測っていれば10ミリや20ミリは簡単に変動するのだけれども、日本全体で国民健康・栄養調査で血压を調べたときに3年に1回やる検査で平均値が10ミリ上がっていますといったらこれは大問題になりますので、そういう意味で、ここでは当然のことながら集団のレベルでの数値の変化を議論しているから、個人で日によって違うよとかそういう話は関係なのだと思いますね。

ですので、ここは割り切ってしまうて、あくまでも集団レベルでの健康影響を見ているのだと書いてしまったほうがすっきりするということになります。

○寒水専門委員 この文章が何のための補足なのか、分からなかったということです。

○川村座長 ありがとうございます。多分、この文章を書く前のいろいろな議論の中で、個人のレベルでも集団レベルでは違いますからとかいう話は出てきていたので、それを事務局が丁寧に書いてくださったのだと思います。そういう意味では、これは座長のまとめ方の下手さかげんであって、おっしゃる意味は今明確になりましたので、ここで意味のある差異という、臨床ではMinimal Clinical Important Differenceという概念があるのだけれども、御存じのとおりですが、ここでは個人がどうのこうのうという話ではなくて、あくまでも日本全体として、あるいは日本全体とは限らずコミュニティー全体としてハザードがどのぐらい生じるかということを考慮する。そういう集団の視点ですよというふうにはっきり書いたほうがいいかもしれないですね。

この点について、ほかの先生方はいかがでしょうか。広瀬先生、お願いします。いらっしゃった。

○広瀬専門参考人 すみません。先ほどちょっと中座してしまって申し訳ありません。

個人の変化というのは、いわゆる調査したときの個人ではなくて、多分これは毒性学的に意味のある、個体にとって意味のある数字の変化と、ここでは集団を扱っているので集団で全体で血压が何ミリか上がったときの変化というのは、意味が違うのは当たり前なので、わざわざ書く必要もないと言われればもちろんそうなのですが、ハイブリッド法だと、例えば血压が10ミリ上がってしまう人が何%増えるというところを今度はBMRにしてしまうという考え方を取るときとの違いを説明しようとしているので、これだけ読むと確かに誤解というか、意味のないことを書いているように見えるのですけれども、次元

の違う話を1つの中で比べてしまっているのようになっていて、そういう意味では表現を変えたほうがいいのかとは思いました。

○川村座長 ありがとうございます。個人における数値の変化ではなくて、集団における数値の変化というのが正しいかどうか分からないのだけれども、影響ですね。個人における影響のことを指しているのではなくて、集団における健康影響のことをここでは指していると、そこでは検討するのだというニュアンスでよろしいですか、広瀬先生。

○広瀬専門参考人 個人における毒性が、血圧は有害影響とは言いませんが、何らかの病理的变化が起きる割合と、集団全体で。

それはまた違うな。連続量なので、それは二値データで、申し訳ありません。個人にとっては10ミリも20ミリも上がるのは確かに影響なので、そういつてしまう人が何人現れるかというハイブリッド法のような取り方と、公衆衛生で何%上がるという取り方、2種類の取り方があるので、それはそれぞれ意味が違うのですよということが言いたいのかなと。

○川村座長 個人における変化でBMRを決めるのではなくて、集団における数値の変化でBMRを決める。個人だと、例えば血圧だと、10ミリ違えば上がったと言えるかもしれないけれども、2ミリとか4ミリ上がったところで、そんなの誤差範囲ですよ。そんなのあまり気にしない。だけれども、そうではなくて集団全体として2ミリ動けば、国全体としてはアウトカムが変わってくるので、集団全体の変化の意味ですよ。個人において意味がある変化のことではなくて、集団において意味のある変化を取りますよという意味だと僕は理解したのです。生物学的に意味のある変化を、個人レベルで意味のある変化ではなくて、集団レベルで意味のある変化です。だから、個人の変化に比べると小さいですよ。

猪飼先生が何かコメントしてくださったので、口頭でお願いできますか。

○猪飼専門委員 猪飼です。書き方の文案として先ほどの①のところですが、後半、その際、個人における値の変動、例えば日内変動の幅と比べて集団間や前後比較における代表値（平均値や中央値）の差はわずかな差でも大きな公衆衛生上の意味を持ち得ることを考慮する。

○川村座長 具体的な説明だと思います。

○猪飼専門委員 御提案です。以上です。

○川村座長 では、今のがまるごと入るかどうかわからないので、少し詰めさせてもらってよろしいですか。しかし、おっしゃった意味はとてもよく分かって、個人内変動、個人の中でのチャンスによる変化、数値のばらつきであったり変動であったりしますが、個人内変動と違って、集団で同じ変動が、平均値や中央値が動くのは意味が異なるということをごここに上手にコンパクトに書き込んでみたいと思います。今、メモを書いていたと思うので、事務局で残せると思いますので、このところをコンパクトにうまく書き込めないかなと思います。今のままではどういう意味かが分かりにくいから、もう少し解説、補足の説明でもいいかもしれないし、本文自体にしてもいいかもしれないですが、ちょっ

と書き方を変えて分かりやすく、意味が違うのだと。個人内変動、個人として意味がある変動と集団として意味がある変動は違いますよということを書きたいと思います。

では、寒水先生、意見をお願いいたします。

○寒水専門委員 多分、猪飼先生は分かってこう書いていらっしゃると思うのですが、この前後比較というところは混乱を与えると思うのですが、集団内ということですね。集団内では、個人における値の変動と被るから、あえてこういうふうに書いたのではないかと思ったのです。

○猪飼専門委員 御指摘は受け入れます。集団内の前後比較と書けば、なお正確ですね。

○川村座長 僕もそのことについてよく考えたことがあって、どこからが意味があるかなと。要するに単回の測定は本当にすごくばらつくのだけれども、毎日定期的に測っている人の1年通じた変化というのもあるので、それは中間ぐらいですね。日本全体の一億の平均値というのも、連続的に分布するのだろうなと。だから、実は個人内と集団レベルというふうに言葉では2つに分かれるのだけれども、その中間にたくさんある自分みたいなのがあって。

○寒水専門委員 その平均値も年齢とともにずれていくので。

○川村座長 だから、いわゆる単回測定の持つ意味と、それから集合的な意味、平均値なり中央値なり代表値の変化というのは意味が異なる。ここでは通常、臨床で使われている1回ごとの測定の変化を意味することではなくて、単回測定の変化の意味ではなくて、あくまでも可及的にたくさんの人で構成される集団の代表値の変動であるということ。うまい表現でまとめられるかどうか分からないですけども、そこはとにかく個人もたくさん頻回に測れば集団の意味に近づいていくので、本当のもっとマスの集団と、そういうふうにここではかなりマスを意識したものであると。個人の1回の変動のこと、いわゆるMCIDのことを指しているのではないですよということが分かればいいのかないかなと思いました。

猪飼先生、納得していないかもしれないけれども、コメントがありましたら。

○猪飼専門委員 猪飼です。先ほどの寒水先生の集団内のというのを入れたらという話ですが、今の川村先生のお話を踏まえると、個人内での前後比較であっても、つまり代表値なのか、ばらつきを見ているのかというところが大事なのではないかということですね。

○寒水専門委員 そうです。統計の言葉を使えないからこうなってしまうと思いますが、統計の言葉を使うと、データと単なる母平均なのです。個人におけるというのはあくまでもデータのばらつきで、母平均が動くということ、統計の言葉で説明すればはっきりします。この種の文章にそういう言葉は使いにくいので、折衷案的にこういった表現をしてくれたのではないかと考えています。

○猪飼専門委員 集団内にこだわらないという意味で、正しく意味が伝わればいいですが、寒水先生と思いは同じです。

以上です。

○川村座長 確かに母平均という言葉を使えばそれだけで済んでしまうような話ですね。ちょっと考えます。臨床医は、母平均という概念を離れても、薬を飲んで血圧が下がったというためには、2ミリ、3ミリ下がったから血圧によく効いていますねとは言わないわけで、個人だと20ミリぐらい下がらないと下がったと言えないけれども、集団では2ミリ下がれば起きてくるイベントの数はかなり違うという意味で意味がある。そういうニュアンスかなと思ったのだけれども、どうでしょうね。

だから、確かに偶然誤差というのですかね。単回ごとの測定の誤差を超える差がないと駄目で、個人内だと相対的に誤差が大きいけれども、母平均になるとかなりステディーなものになるから、少しの変化でも意味があることになるのですが、統計学的用語をあまり使わないで表現しないとイケないから。

すみません。議論を止めてしまうのはいけないので、ちょっと頭をクールにしてもう一度考えます。

では、①の後半の文章は表現を変えますが、そのほか②、③、④についてはいかがでしょうか。

○寺谷評価調整官 寺谷です。すみません、ちょっと頭の整理をしたいのですが、今やっている作業は、動物試験で得られた参考資料1を踏まえつつ、そこそここういうところが今回のことに関して言ったら違うよねということを見ながら作業しているのだと僕は理解していたところです。そのような理解で大丈夫ですよ。

○川村座長 そうですね。動物実験と違うということを強く意識しているかということ、必ずしもそうではないけれども、もちろん動物実験との違いは踏まえないといけないので。だけれども、疫学研究で使う様々な検査値ですね。連続値データのほとんどは臨床検査値なので、臨床検査値の持つ特性について、その中で意味のある変化、BMRをどう捉えるかという議論をしているので、動物実験そのものとの対比には必ずしもなっていないかもしれないです。

○寺谷評価調整官 ありがとうございます。あとはこの疫学の中とBMRのことを今まさに議論した上で、これを踏まえてできるものは何かといたら、結局、まず指針をつくっていくということとか、それに付随する技術的文書とか解説書みたいなものが必要なのか、必要じゃないのかということも含めて作っていくのだろうなというイメージを持っているところなのですが、それもそういう理解で大丈夫ですか。

○川村座長 それはそれで正しいです。指針を作って、指針は骨格なので、ガイドンスとか、あるいは実施マニュアルといったものを具体的に書き込んでいく。作業用のマニュアルを作っていくことになりますので、それはお考えのとおりだと思います。

○寺谷評価調整官 そういう意味で言うと、今のところの母平均の数値の変化を捉えている話と、例えば個人において意味があるかないかみたいなことはとても重要なことで、特に僕は厚労省での経験でもあったのですが、ごちゃごちゃにされてしまって、行政側もい



ろいろ大変な目に遭ったりしていることは多々あるので、どこかに書いておくのはすごく大事だなと思っているところです。

書く場所とかは、逆に言えば、こういう議論をしたということを含めて、何かエッセンスが盛り込まれていけばいいというような理解でいいかなと思ったのですが、いかがですか。

○川村座長 そのとおりです。厚労省サイドというとな変ですけども、医療を抱えている側からいうと、例えば血圧2ミリの違いなんか誤差に入ってしまうから意味がないと言われてしまう。MCIDに入らないということになってしまうので、臨床家ではMCIDは多分比較的普遍的に理解してもらえらると思うのですけれども、そのことをここで言っているのではないですよ。パブリックヘルスの観点から、日本全体でどのぐらいの健康ロス、健康不利益があるかということ議論してBMRを決めるのですという趣旨なので、それをどういいう言葉を使って表すかということで、着任早々ですが、お知恵をお借りできればと思います。

○寺谷評価調整官 恐らく今いろいろ議論いただいたものを含めて、また事務局のほうでいろいろやっていくのだと思うのですが、最終的に今いただいたものをどこに書いていくかということも含めてですよね。なので、もしかしたら、今どうしてもここで作っている文章は、もともと動物実験で得られた用量反応データの適用というベンチマークドーズ法の活用に関する指針を直していく、ないしは意識してつくっているものなので、そこに無理やり全てのことを書き込もうと思うと、多分、何が言いたいのかよく分からない文章になってしまう気がしています。

ただ、いずれにしてもこういう議論をちゃんとやったということは、どこかに何か残るようにしていくということでもいいですね。そういうふうに理解しました。

○川村座長 ありがとうございます。全体、こういった規定類は一定のつくるルールとか作り込み方があるのは承知しているので、そこところは法令の専門家である事務局の人に委ねることです。そこに盛り込むエッセンスの科学的な妥当性をここで議論して、かなり雑然とした議論になってしまうところを交通整理して、所定の形に収めてほしい。何段階か必要だと思いますので、必ずしもこの指針で全部尽くす必要はなくて、こんな細かいことをここに書いたら全体がかえって見にくくなるということは下位の規則に落とし込んでいただく。そこはむしろ事務局の皆さんがプロでいらっしゃるから、そこは委ねます。もちろんみんなで共有しますけれどもね。それは当たり前ですが。

だから、今日の議論を上手に拾っていただくというのが事務局にお願いしたいことなのです。ただ、議論に漏れがあつてはいけないので、こういう意味だ、ああいう意味だということ様々な角度からしゃべっていて、それがとてつもなく雑然として聞こえると思いますけれども、自分でもしゃべっていてそう思うので、これはどこかで頭をクールにしてもう一回見直さないといけないと思っているので、収束しないのだけれども、一旦引き取るということで、それが座長の仕事かなと思っています。

ただ、今の議論の中で分かったように、個人と集団の話とか、個人と集団の間に多数の機会がある。frequently measuredの数値が出てくる。そうなると、サイズが増えるにしたがって安定的なものになる。それは要するに母平均的なものに近づくということで、数の論理があるということで、そういうことはお分かりになっていると思うけれども、そういったことを今、それぞれの立場から発言していただいているということなので、上手にまとめるという、さじを丸投げするわけではないですけれども、一緒に考えますので、うまく拾っていただければと思います。よろしいでしょうか。

○寺谷評価調整官 ありがとうございます。自分自身の頭の整理が、あと自分自身もどういう役割を果たすべきか、ちょっと整理できていなかったところがあるので、ありがとうございます。また皆さんのほうに戻して、ぜひ川村座長のほうで進めていただければと思います。よろしくをお願いします。

○川村座長 ありがとうございます。ここはいろいろな立場の人の集まりなので、分野を超えて、ここは議論する場ですから。ただ、行く行くはそれを国の機関としてきちんと一つの方針にまとめ、さらにはリスク評価機関たる食品安全委員会からリスク管理機関にお渡しして、具体的な政策に還元していってもらいたいということなので、そういう流れの中の一つで、集団の議論というのは必ずしも収束するとは限らないのだけれども、それはコロナでも一緒ですけれども、どこかで政策を実現するところでは一つに決めないといけなないので、それは分かった人が集約すべきことであって、万機公論に決すべきところと、それから実際に政策に転化するところとは部署が違ふと僕は思います。

ほかによろしいでしょうか。そういう意味で、いろいろな意見はぜひ言っていただきたいのですが、ほかに②、③、④も含めて御意見いただけませんかでしょうか。

さっき途中で広瀬先生が中座されているときに広瀬先生に振りそうになったのは、ハイブリッド法について、前に村田先生と御一緒にお書きになったものもあるし、実際に使ったときの感触ですね。要するに、数量の変化を人の人数に還元する仕組み、人の人数の変動に移す仕組みというふうに理解しているのですけれども、そのときの使われた感触とか留意点はございますでしょうか。

○広瀬専門参考人 先ほど少し説明したかもしれないですけれども、平均値で集団で2ミリ上がるという話と、あ那时候村田先生と議論というか、主に村田先生のアイデアですけれども、そうではなくて、例えば10ミリとか20ミリを超えた人数が何%増えるか。捉え方の側面が平均値だけではなくて、異常値になった人が全体の中で何%増えたかというのをベンチマークドーズにするというやり方をすると、二値データのような扱いができるようになる。そのときのBMRは、今度は集団が2ミリ上がる云々ではなくて、異常値をどこに設定するかが重要になってくるので、そういう意味です。

そこで、僕の最後のコメントになるけれども、気づいたのは、疫学データの部分はばく露量も実は連続値でして、レスポンス連続値とばく露量も連続値。動物実験は、レスポンスは連続値だけれども、ばく露量は規定値なのですね。4段階、5段階しかないのです。

でも、疫学データはそこも連続値になるので、そうすると多くの医学論文は区分分けして、どこからどの範囲はこの集団というふうなやり方をして用量反応関係を説明したりしているのです。その間の取り方を変えると、また微妙に結果も変わってくるので、それは今度、BMRを設定するのとは違う次元のハイブリッド法での考慮点になるので、それはここに書くのか、別のところに書くのかして、必要な観点ではないかというのをコメントさせていただきました。

○川村座長 おっしゃる意味は分かります。量的なものをカテゴライズして二値に近い、異常値と言っていかどうか分からないけれども、特定の領域に入ってくる人の人数で検討して、その人数の割合がどれだけ増えるかということベンチマークレスポンスにするということですね。

それから、そもそも研究自体も表現上の問題としてばく露量をカテゴライズして、それでその中の平均値とか中央値で表示する場合が多い。本当は、ばく露量もアウトカムも両方とも連続量なので、統計学的に言えば散布図になるはずで、その中で回帰をしたりすることができるのだけれども、何で回帰するかが決まっていないうるか、モデル化するしかないのだけれども、それをただ分かりやすさのために、研究としては割とカテゴライズして、その中にmean ±sdのような形で表現することが多いですね。それは見やすさのためだと思うのです。

だから、ベンチマークドーズでドーズレスポンスカーブを描く上には、本当は生データがあれば全部、カテゴライズも必要なくてそのまま値を使えばいいと思いますけれども、ただ、一回論文にするとときにカテゴライズしてしまっているものも多いので、それをどう扱うかということはあるかもしれませんね。

○広瀬専門参考人 追加で、多くはこういった期間で評価する場合、生データまで得られないときがあるので、論文のデータからやるという制限がある場合とそうではない個別データが取れた場合にどうするかという話と2つのやり方があるので、その辺、そういうやり方があるよということは、それは動物実験とは明らかに違う現象なので、書いていただければという意味です。

○川村座長 分かりました。実際に今はデータの供託ということをやって、RCT、臨床試験なんかは最初から供出するという前提でやることも多いのですけれども、ただ、疫学研究のデータはあまり公共財となっていないことが多いだろうから、そうすると論文になったものの範囲でやる。そうするとカテゴリー分けの影響がモデル化に当たって少し影響する可能性がないとは言えない。本質的にはうまくいっていただけないのだけれども、どうしても現実、きれいに見えるように書いてしまう場合もあるので、等間隔であるとか、等比的な間隔ではなくて、よく使うのは分位でやることも多いし、分位になると間隔はばらばらになってしまうし、そうなったときに出てくる結果に多少違いがあるかもしれませんね。だから、そういうときにカテゴライズの仕方によってちょっと影響を受けるかもしれませんよということは付記しておいてもいいかもしれないです。

あとはよろしいでしょうか。

それでは、これで文案が仕上がったわけではないのですけれども、議論を踏まえて文章をもう一回練りますので、一定の時間をいただければと思います。ただ、ここで言うておかないといけない、ここの表現は問題ですよと、あるいはこれを追加しないといけないとか、そういうことがありましたら。③と④は動物指針とあまり変わらないので、既にリファレンスとしてひな型があるわけですけれども、②は新たな文章になるのかもしれないので、特に御意見がありましたらいただきます。

これでまとめますというふうに示し切れていないところが座長の力不足ではあるのだけれども、私自身もやはりちょっと一旦頭をクールにしないと、それで皆さんの言ったことを十分整理して、事務局とやり取りする中で決めていく必要があるので、成文化というか、文章の骨格はこうなのですけれども、少し表現を変えるべき点についてはちょっと時間をいただければと思います。

それでは、その次の議題があるので、事務局に振ってもよろしいですか。

○爲廣課長補佐 今のところなのですけれども、この文案について、最初のところから確認だけ手短かにさせていただければと思うのですけれども、5行目、BMRの設定に関しまして、もうちょっと表現を工夫するということと、14行目、二値データに関しましては、記載できるようところをもう少し考慮しながら記載を考える。もしくは付随する覚書のようなものをつくれるかどうか検討するというところ。最後に、連続値データのところに関しましては①の個人における変化と集団における変化についての記載ぶりを少し考えるというところを修正するというような形でよろしいでしょうか。

○川村座長 一応、今出ている意見の中ではそういうことになります。

○爲廣課長補佐 ありがとうございます。

○川村座長 どういう表現がいいかは、私ももう一回、私としての整理を事務局に伝えますので。すごく頭の中が充実したのは充実しましたが、頭の中が整理されていない状態なので、これを澤田専門委員からも御指摘があったように分かりやすくする。あるいは赤堀専門委員からも御指摘があったようにぱっと分かるようにするということが大事なので、そのところは本質を伝えながら分かりやすくするというのをこれから作業していきたいと思います。

○爲廣課長補佐 ありがとうございます。事務局のほうでも、今回御議論いただいたお話の中に出てきたエッセンスをできるだけ拾いながらまとめていきたいと思いますので、よろしく願いいたします。

○川村座長 お願いします。

休憩を少し取りますか。

○爲廣課長補佐 それでは、5分間休憩を取っていただいて、16時から開始ということをお願いできればと思います。よろしく願いいたします。

○川村座長 では、一旦休憩に入ります。ありがとうございました。

(休 憩)

○川村座長 それでは、皆さんおそろいですので、続けてまいりたいと思います。

議事(2)に入ります。昨年度まで、西浦専門委員、広瀬専門参考人を中心に実施していただいていたベイズ推定を活用したベンチマークドーズ法の評価手法の検討と国際動向に関する検討について、西浦専門委員からその成果を御発表いただけるそうです。

それでは、西浦先生、よろしいでしょうか。お願いいたします。

○西浦専門委員 承知しました。私から画面を共有しながらお話しさせていただければと思います。前年度とその前の年度で、ここに示しているようなタイトルで「ベイズ推定を活用したベンチマークドーズ法の評価手法検討と国際動向の研究」ということで、研究班を展開させていただいていました。私だけではなくて、国立衛研の広瀬先生にも共同研究者で入っていただいて、あと、長崎県立大学の齋藤正也先生、特にベイズ推定に強いバックグラウンドをお持ちなのでということで一緒に研究をさせていただいてきましたので、その成果を報告させていただきます。

ベイズの話を一きなりしても分かりにくくなりますし、今回のワーキングをやっていても人が少し入れ替わった中で、新しい方法というのをいつも振り返りながらやったほうが良いと思うので、ちょっとrepetitiveですけれども、BMDの話のベースをちょっとします。用量反応関係を利用して、どこまで食品に特定の化学物質、食品添加物などを許容できるかという課題が主立った適応範囲なのですけれども、そういう用量反応関係を評価するときallowable doseを決めるために統計モデルをたくさん適合しながらリスク評価をするということが増えてきました。これがBMD法と呼ばれるわけなのですけれども、これまではNOAELといって、全く反応しない程度のドーズからどンドンドーズを上げつつ実験を行って行って、全く反応しない上限値の値(ドーズ)を評価するという動物実験をすることとか、あるいはLOAELといって反応する個体数が最も小さい値を取ること、動物実験の評価が別途あったのですけれども、OECDとかの動物実験ガイドラインを使って用量反応関係を見てもいつもサンプル数は少ないのが実体です。NOAELみたいに反応ゼロのドーズがあったとしても、統計学者の見方からすると、それはサンプリングエラーではないかというふうな話にもなることがあるので、統計モデルを利用してこういうデータ分析をするのがいいのだよということが広まっています。

今、指針で議論されているのは、そのステップとして使っていく運用方法がまだ透明化されていないからということだと私自身は理解しています。例えばステップ1でドーズレスポンスのモデルを大体、往々にして10個ぐらいフィットします。2個か3個のパラメータからできたスタンダードディストリビューションと呼ばれるものを適合して、そこにBMRが定義されていて、特定のBMR、例えばここで示しているのはBMDですけれども、BMDを計算すると、その片側の95%信頼区間、下限値というのがBMDL<sub>10</sub>と呼ばれるわけなのですけ

れども、これを取ると大体NOAELと一緒にすることが多いということで、よく利用されているものです。

NOAEL、LOAELとBMDの利点みたいなものをそれぞれ比較しながらもう一回振り返っておきたいのですが、NOAEL、LOAELは研究デザインでどこについて、特定のドーズのどこを特に分析しているかに強く依存してしまう。毒性試験結果の部分情報しか利用しないわけですね。ドーズレスポンスカーブというのはとてもリッチなのだけれども、特定のドーズの結果ばかりを集中的に見ていることになりまして、不確実性に不適切な対応と書いていますけれども、例えば3匹とか5匹しかサンプルを特定のドーズで検討していないのに、そこで無反応だったらそれを無毒性量と言ってしまえるという定義自体の問題がありました。

フィロソフィカルにそれとは全く違う方法ということでBMDが出てきたのですけれども、検討ドーズに制約を受けずに、でも、多くのドーズに満遍なくデータを分析されていたら、サンプリングエラーも考慮しながら信頼区間の下限値を計算できる。それはドーズ選択、ドーズ間隔とかサンプルサイズの制約を、データによっては数が少なかったら受けますけれども、豊富に観察されていると自信を持って言える限りは大きな制約を受けない。特に高い部分のドーズの影響も加味して検討できる。特定のBMRの反応レベルで比較可能になる。この研究だったらこれぐらいのBMRで、ドーズは幾らだったけれども、同じBMRで別の方法を取るとこうなったとか、そういうのがtransparentにできるということです。

全ての用量反応曲線データ、どの曲線が正しいということは残念ながら限られたデータで機能も分かっていないドーズレスポンスの曲線だと分からないのですけれども、少なくとも全てのモデルの不確実性を加味した上でどうするかという検討ができるし、実際にNOAELという観察自体が仮にデータ上になくてもBMDLを計算できるというのがBMD法の利点なので、使えるのではないかとということで急速に発展している方法です。

その中でもBayesian Benchmark Dose Methodというものが最近になって欧米で盛んに使われるようになったので、こうやってわざわざ研究班でいろいろな調査とか自分たちの研究もさせていただくことになりました。BMD法にベイズ推定を適用しているわけなのですが、ベイズ推定をすると、もちろん確率論的な推定が可能になります。いつも推定値という何らかのパラメータが分布として得られるということと、既存の情報を加味した推定が可能になる。つまり、例えば何らかの事前情報があることとか、過去にほかの場所でリスク評価をした結果を参照しながら分析することができるというもので、最尤法を大きく置き換えつつ広がろうとしています。だから、人によっては第1次用量反応評価革命というのはNOAELからBMDに移り変わったときだという話をする一方で、第2次はBMDの点推定から区間推定をするようになったということなのだと言われており、さらに、海外ではアメリカなんかで推定システム、推定できるサーバーをインターネットで公開して、一人一人登録したユーザーは誰でも使えるという状態にすることも始まっています。

なので、ベイズ推定を使った方法がどういったもので、どこまで使えて何がピットフォールなのかということをしっかり理解しておかないといけないので、今日はその話をカバ

一するのですけれども、そもそも頻度論者と言われている最尤推定をする人、frequentist と呼ばれる者と、ベイズ論者、bayesianと呼ばれている人たちは何が違うのかというのを理解しておかないと、その話にするっと入っていくことは難しいので、ちょっとだけカバーします。

頻度論者の人は、パラメータ ( $\theta$ ) というのは未知の固定された値であって、それは例えば先ほどちょっと話に出てきた母平均とか、平均値を推定したいのだったら母サンプルに依存するわけですね。観察データを使ってその固定値を推定したいというのがフィロソフィーです。異なるサンプルを採ると異なる推定値が得られて、それは標本分布になりますということです。ベイズ論者の場合はそこに対しての向かい方がちょっとだけ違うのです。パラメータというのはrandom variable、つまり確率変数と呼ばれているものです。分布を利用して確率変数を記述したいのですけれども、観察データを利用しながら、この辺で大体この値はありそうという話をアップデートしていくのだよというのがベイズ論者であるということの違いがあります。

それを踏まえてどういうふうにベイズ推定が特にこのBMD法に利用されているのかというのを知るために研究班を組織させていただきましたけれども、用量反応モデリングの手順とか判断基準の検討が中心課題です。既に導入されたBMD法に関しての既存のソフトウェアの使用手順を整理することだったり、諸外国でどのような動きになっているのかというのをしっかり理解することが第一目的で、もう一個は国外の主要なリスク評価機関でベイズ推定の活用状況とカリコメンデーションがどうなるのかというのをしっかりと理解することによって、日本でも最尤法でこういうのを実装してきたのですけれども、ベイズ推定に遅れないということが第一義的なものでした。ここでは読みませんが、それに対する個別の細目課題を設定してやってきました。

班研究を2年やらせていただいたのですけれども、コロナで走りながらの中だったので、なかなか大変だったのですが、大きな項目としては左側の列を見ていただくとおり3つありました。1が一番大事だと私が認識しているのですが、ベイズ推定をBMD法に活用にする手順と判断基準の詳細をしっかりと理解するということ。2番目は国外のリスク評価機関でそれがどういうふうに使われていて、今後どうなるのかということについてある程度把握するということと、3ですけれども、既にBMD法にベイズ推定を使うということを実装したソフトウェアがあるので、それがどんなものであって、どういう手順で今後の動向がどうなるのかという話をカバーします。

あまり時間がないのでさっといきますけれども、1のところが重要なので、テクニカルな部分は一箇一箇丁寧に見ていきたいと思います。

先ほどもお話したわけですがけれども、まず、ちゃんと頻度論とベイズ論というものの差異とか計算手法の把握を理解するところからスタートです。このスライドは報告書から取っているのでこういうフォーマルな形で書いているのですけれども、さっきのスライドで説明したものを表にして書き換えているだけです。でも、ベイズ推定をBMD法に適用する

上では点推定と違って区間推定をする利点があります。1つとても大きなことなのかもしれませんが、すごく小さいサンプル数の実験データとかでも対応が可能であるということです。妥当に区間推定をするので、不確実性が最も妥当な方法で定量化可能だからと考えられます。

異なるデータを足していけば推定精度が向上します。過去のデータを参照可能だからということですね。正規性とかを区間推定の際に仮定しなくてよいので、最尤推定の際にはブートストラップとかで漸近正規性が必要になるときがありますけれども、より区間推定を正確に実施することができる。使い慣れるととても便利な方法であって、不適解を簡単に回避できる。あとはベイズ更新をするためのMCMCというチェーンのコンピューテーション、計算プロセスさえ勉強すれば比較的簡単に複雑なモデルもフィットすることができる。

一方で、役所として困るだろうということは、事前情報からスタートするわけなのですが、事前分布をどのように扱えばいいかというのは、この後議論しますが、必ずしも一義的に解決することは難しいと思います。推定時間が長いので、コンピューター機能を要します。この研究を介して事務局からも、自分たちのコンピューターで大丈夫ですかという相談をいただきました。多くの事務局に置かれているコンピューターでは計算できなさそうであるということも途中で分かりましたので、サーバーを手に入れてくださいというアドバイスをしましたけれども、ひどい場合は、事後分布が収束しない。計算は、デスクトップが凍るといようなことにありますし、解釈には慣れと修練を要します。どのようにして収束を判断するのかという指標もいっぱいあるのです。ゲルマン・ルービンズ・クライテリアというのに始まって、ほかにもいっぱい、こうしたら収束だよというようなものとかは使いながら覚えなさいといけないし、日進月歩で使用メトリックが発展するのも結構スピードが速いということがありますし、ソフトウェアが限定されていて、プログラミング言語でもR言語ぐらいはできないと厳しいということがあります。

ただ、ベイズ推定をBMD法に使うと、推定というものがとてもリッチな不確実性のビジュアルの表現とともに得られていくということです。

左手に出しているように、確率分布としてBMDが得られているというのが左上のほうです。左下のほうはパラメータの事後分布です。それと、チェーンが安定していたのかというのをそれぞれのパラメータの横で見えています。

右側に書かれているのですが、ベイズ推定を使うと信頼性が高い。これは何かというと、同じデータが与えられて、反復して同じ手法で計算したときに大体同じ結果が得られる度合いが高いですよということが、二値データはもちろん、連続データに関してもそういう結果が得られていて、frequentistよりもいいということが言われています。

なので、ベイズ推定をBMD法に適用する特徴とか基本手順をしっかり踏まえた上で次に進んでいかないといけないわけなのですが、ここは大事なのでちょっとだけ読みます。ベイズ推定では事前分布に関する想定が一定の役割を果たすということですが、この後ちょっとお話ししますが、事前分布をしっかり設定しないとイケません。推定パラメー



タの事後分布に代表されるように、区間推定の結果が分布として得られることになる。その見方とイメージについて実用的な事例を通じて理解することが望ましい。

ベイズ推定の中でもベイズモデル平均化手法というものが特にBMD法の中でモデル平均化をする中心的手法となっています。今までモデル平均化というとAICを重みにつけるかとかいろいろな距離を取る方法があったのですが、ベイズ推定の場合は、それぞれのベイズのモデルがどれくらい正しいかというメトリックを直接定量化できるので、そのチェーンから得られたモデルが正しそうなものを重みに使った平均化ができますので、それが今後、中心的役割を果たすと考えられます。

モデルパラメータの分布というのもベイズ推定を使うと、こんな感じで事後分布を分布として得られるので、例えばパラメータ  $b$  のようなものだったら、何か左右対称だな、なのですけれども、パラメータ  $a$  のようなちょっとだけ片方にスキュードな分布が得られた場合は、不確実性として左側よりも右側に裾が長いので、右側のほうに向かってまだまだあり得るのだなというようなこととかをビジュアルに理解することができます。

それだけではなくて、図の下に書いていますけれども、事後分布が平坦であればあるほど不確実性が高い推定結果であることを反映しますので、平坦という形状そのものもベイズ推定の結果として重要な役割を果たすことになります。もちろんそういう分布を、ちゃんとソフトウェアを使えば1行書くだけでこうやって分位点に直すことができます。それぞれのパラメータの不確実性を基に分位というのを見て、95%信頼区間をそのパラメータについて、2.5パーセンタイルから97.5パーセンタイルのところを取って求めていくわけなのですけれども、それだけではなくて、どんな分布をしているのかなど。

この右端にRhatとありますね。こういう辺りが、収束がこのパラメータに関してはちゃんと分かったかどうかという判定もしながら見ていくようなことにも使えるということですね。

そうやって得られたパラメータから、リサンプリングエクスペリメントというのですけれども、先ほど提示した分布にランダムにパラメータが従うとして、それぞれのパラメータを抽出するという実験を例えば1万回とかやると、シェードというのですか、水色の影のようなものが得られます。水色の影の中央値を取ると、この赤線が得られるわけなのですけれども、観察値を95%の信用区間に含んでいるかどうかというビジュアルアセスメントも、今までよりも格段にきれいな形で見ることができずし、特定のドーズでのレスポンスレートに関しても不確実性をとても便利に出すことができるようになってきています。

BMDの推定値のグラフがこちらなのですが、added riskの場合もextra riskの場合も分布として見ることはできますし、確率密度分布を基にBMDLが得られるということになります。なので、ベイズ法を利用したBMD法を実装するというのでは、皆さん見ていただいとおり、全て確率分布として、random variableなので確率変数の推定結果として得られる確率分布を扱います。

ベイズ推定は区間推定なので、BMDLというのとはもともと信頼区間下限値ですから、区間推定を便利にできる方法があると、やはり便利に使えます。特に小サンプルとかですごく乏しい実験のときに、頻度論と比べたらいい結果になると考えています。

頻度論における区間推定では、例えばブートストラップ法なんかは今、私たちが日本語でつくったソフトウェアで実装しているものですが、考え方自体が異なるのですが、推定値自体は実は似通っていることが知られています。ただ、一致はしません。

頻度論に似た p 値とかモデル平均化方法があるというのが、ここの後のPODを決める上で重要になります。

大事なことなのですが、BMDSの頻度論ベースよりも信頼性が高いというのは、幾つかのデータセットで何回もそういう信頼性のためのランダムシミュレーションみたいな実験をしている人たちが何グループかあるのですけれども、全てにおいて一致しているのはベイズのほうが信頼区間が高いということが分かっています。

事前に所有する情報ではどうやって活用していくとか、初期設定をどうやってやるかというのが恐らく役所で一番悩むことであり、かつベイズ推定でもベイズの初心者の人たちはどうしてもプライヤのところの恣意性というもので一回止まります。これはしょうがないことで、慣れていくしかないみたいなのです。ベイズ推定は初期設定で不明なパラメータについて、初期値を与えるのではなくて分布を与えます。それをデータで逐次的に更新していくということです。だから、とてもいいかげんな推定なのです。

そのときに参考になるのですけれども、ある化学物質、食品添加物がどれだけ悪いかということを知るために同様の対象とか、あるいは近い投与量で同一アウトカムに関して検討した既存の研究があるかどうかを必ず探すことになるのですが、今回からは探すというだけではなくて、探してBMD法を適合するということが必要になります。そこで得られた情報があるのだったら、それを使ってさらなるサンプルについて検討すればいいんじゃないということになるのですね。

ただ、サンプル数が過度に小さ過ぎて適合できないとか、BMD法が使えないとか、BMD法による検討が執り行われたことが全くないと。自ら再検討したくてもドーズのデータも得られていない。そういった場合は事前情報として用いることはできません。自ら得られる情報がある場合は自分で一回やればいいのですね。

新しい分析では2つのことをやります。既存のデータに基づく分析結果を事前情報に利用して推定するというときと、全く既存データを利用しなかった場合の両方を検討します。後者では事前分布にはフラットプライヤといって一様分布というものを使いますが、これはベイズの教科書でいつも載っています。もともとの初期設定の分布が真っ平らですと、両端の値だけ決めますということなのですけれども、それで推定をしてみる。2者の間で事後分布の分布を比較します。全く分散が異なるようなところまで収束している場合は事前情報にほぼ独立な結果が得られているということになるので、どちらの結果を

取ってもいいので、ここでは事前情報を利用しなかった結果と言っていますけれども、どちらでも同じ結果になります。

でも、一方で、事前情報を使った場合とそうでない場合で事後分布の分散が異なる場合は往々にしてあるのですけれども、推定結果が事前分布の想定に対して感度が高いと判断します。なので、事前情報が相当影響を与えてしまうことになるので、そのような場合はサンプルサイズが限られている場合に頻繁に起こるので、追加の観察情報の入手を検討することが望ましいので判定を保留することも視野に入れて考える。観察情報の限界点を明示しながらリスク評価を習慣化するとよいでしょうということだけここには書いています。

例えばなのですけれども、プライヤがどれだけ推定値を変えるかということのいい例がここにあります。左上ですけれども、同じ分布について左上の2つのパラメータは一様分布だと想定しました。右上の2つのパラメータでは正規分布とログノーマルみたいなものを想定しました。それぞれの事後分布で得られたBMDの曲線のリアライゼーションがどうなっているのかというのを見ているのが左下と右下です。真ん中は最尤推定法に基づくものですね。

ユニフォームで計算するとBMDは941ぐらいになるけれども、スペシフィックプライヤでやると570ぐらいになる。相当違いが出てくる場合があるので、そのときはエビデンス不十分ということになるのですけれども、感度が高いということを少なくとも報告した上で、その詳細を検討することが必要になります。

だから、そこまで神経質になる必要はないかもしれないけれども、ただ、全部文章化しておいて責任を取らないという状況をつくる役所のやり方とは相当そごがあるということです。

結果の見方ですけれども、最尤法とベイズ法の間で妥当性は大きく異なることが知られています。信頼性の点でベイズモデル平均化を使ったほうがいいということが研究ベースでたくさん知られています。ベイズモデル平均化が頻度論で用いられる平均化よりも適切に観察データとモデルの間の距離を恐らく定量化しているものだと考えられています。マニアックに書いてすみませんが、そういうことなので、これがよさそうですということだけです。

この見解は二値データでも連続値データでもほぼ変わらない。ユニバーサルかどうかということの証明自体は解析的にはできないわけなのですけれども、いっぱい数値的なエクスペリメントをしているとどうもこの結果で大体一致していますということが分かっています。

近似解を用いたベイズ推定は、この後また言いますけれども、ラプラス法というものが実装されているソフトウェアがアメリカで出ていますが、そのときの再現性みたいなものは担保されていない。ただ、妥当性は必ずしも悪くないということが知られています。

ほかにもベイズ単調セミパラメトリック回帰の利用によりモデル不確実性を最小化可能であるオプションつきですよとか、あと、モデルパフォーマンスの観点からやはり最尤法

よりはよさそうです。ただ、最尤法もすぐ悪いというわけではないですということが書かれています。

そういう中で、ベイズ推定を用いることと、これまでのGUIとあって、マウスでカチカチできるような最尤法ベースのものどどちらがいいかということであったりとか、あるいはこれは結構難しいのですが、ラプラス法という近似解ですぐ結果が出るようなソフトウェアを実装しているのはアメリカのBMDSとあって、アメリカの政府関連の研究機関が発行しているソフトウェアがそれをやっていて、一方でBBMDとあって国際的にサーバーをフリーで開放しながら計算を可能にしている、MCMCを可能にしているのはインディアナ大学とか大学の研究者がそういうのをやっているのですけれども、研究者の観点からいうと、区間推定はチェーンの更新でやれるからいいのですよねということのプリンシプルとかに戻ると、やはりマルコフ・チェーン・モンテカルロをちゃんと実装した方法でやるといいですねということが知られていて、例えば二値データに対しても、失敗が何回あるとかいうのをBMDSとBBMDとかで比較をしているような検討結果があるのですけれども、BBMDが自慢しているのは、BMDSが失敗しているようなケースが多いデータを使っても、BBMDは必ず収束に至っている。区間推定をちゃんとチェーンを回しながらやっていたら安定しているよということができるのでは全然違います。

でも、ベイズ推定を使用する上では留意点が必要ですよということなのですから、ベイズ推定を活用可能なのは、最近こうやって発展しているのはやはりコンピューターがよくなっているからです。マルコフ・チェーン・モンテカルロを利用するとか、安定性を分析可能、解がどれだけ安定しているかということのをちゃんとビジュアルに検討しながら分析できる研究体制を、例えばこれを使うのが事務局になるのだったら、事務局で備えてもらうということが長期的に必要です。単にコンピューターを買えばおしまいというわけではなくて、そのチェーンを自分たちで見ながら、いつも出力できないと困ります。事前分布として活用可能なエビデンスを判定可能であるということも重要になります。使うかどうかみたいなこともしっかりと文献をひもといてみないといけないということですね。

計算結果は二次元分布として描述できて、BMDLとかの参照用量の根拠となる計算結果をまとめることができるので、とてもいいですよという意味ではいいのだけれども、ただ、(iv)で書いていますけれども、平均化をするのですが、各モデルが仮定する式の妥当性評価だとか、BMDL以外の細部の結果に関して説明可能でないとこの手法を用いたときに、事後にそのデータのの一つ一つを説明することができないとちょっと困るというのも厳しい点では事実かと思えます。

事後分布について事前分布に対する感度分析を実施することが可能であるというのはミニマムだと思います。

利点と欠点はそういうことを踏まえてまとめるとこんな感じです。利点としては小サンプルに対応可能、異なるデータを足すことで推定精度が向上、正規性を仮定する必要がない、不適解を数値的に回避できる、MCMCを理解することで複雑なモデルが使える。

ただし、欠点としては、事前分布をどのように扱えばいいのかというのは永遠の課題のままになると思います。推定時間が長い。コンピューター機能を要する。事後分布が収束しない場合がある。適合度やp値、頻度論と異なるので、解釈に慣れと修練を要するといえますけれども、これは1か月ラボでやれば大体大丈夫です。ソフトウェアが限定されているということですね。

最後のサブの課題では、疫学データの分析をテーマにしました。先ほどハイブリッド法の話がありましたけれども、特に疫学データといっても連続データがあったときにどのように推定をすればいいかというところで、今のコンテキストから一旦外れるのですが、ベイズ分位点機能障害閾値というのが使われるようになっていきます。これはここでまず出しているのは、正規性を仮定するハイブリッド法というものです。Dakeishi et al. がリスクアナリシスに出しているもので、横軸が1日当たりのアルコールの摂取量、縦軸が血圧でも特に収縮血圧、上の血圧なのですけれども、ハイブリッド法というのは全てのアルコール摂取量の下でデータが正規分布に従うと仮定して、その正規分布の平均値が一定のドーズレスポンス性を持っている。そういうふうにしなごらデータをフィットする。もともとの140というところから超えている $P_0$ というものからすると、 $P_0$ プラスBMR、ベンチマークレスポンスレート分だけ分位が上がったというところをBMDにしますよということです。

だから、疫学データ分析でハイブリッド法というのがいいなというのは、これは1995年にできている方法論なわけですがけれども、そのように正規分布を利用すると分位点に直すのがとてもシンプルで、その分位を利用すると平均反応率の式を基にBMDもすごく単純に計算することができますよと。式で書けちゃうよというようなことも分かっていたということです。

ただし、データというのは必ずしもこうやってノーマルに分布するはずがないので、ハイブリッド法というのは10年以上前の方法で、必ず問題が生じるということが分かります。これは上限の計算が簡単に可能ですよということを言っているだけなのですけれども、単純に平均値に関する式があったら、平均値プラス $1.65\sigma$ を足したところが信頼区間の下限のカーブになりますよということなのですけれども、これをしていると気づくのですが、分位点に直すと同じようなアーギュメントができるのだったらということなのですけれども、実を言うとベイズには分位点回帰というのがあります。ノーマルディストリビューションを仮定せずとも分布をそれぞれのドーズで分位に直す。連続値で全然ノーマルディストリビューションに従わないようなデータでも、分位点回帰をすれば同じようなアーギュメントでBMD法が適用できますよということです。

例えば、分位点回帰とはどんなものなのかというのをここに出していますけれども、横軸がドーズ、縦軸が反応なのですけれども、左側の図のように、どういうふうに分布しているかという、ドーズが上がるとどうも上振れするところの変な値が増えていそうだな

と。ただ、どんな分布になっているかよく分からないなというので、やはり分位点回帰をしっかりとすることは必要になります。

右側のデータがとても面白いものなのですけれども、正規性が保たれないどころか、スキューネスもドーズによってちょっと複雑な形をしている。ただ、これを立体で見られる方々にとっては、こんなのモデル化したらいいのではないかとすぐ思うわけです。

そういうのでベイズの分位点回帰を使うと、これはまた横軸がドーズですけれども、全く形状がドーズによって違うことにも対応した上で分位点回帰をして、そのときのmeanというのはこの点線が上がっていつている線を見るのと、また、分位を見るという部分はハイブリッド法と一緒になので、 $P_0$ にBMRを足したところを閾値にするとBMDがここですねというのを簡単に出せるということが最近の研究で、といってももう5年前ぐらいですけれども、知られてくるようになりました。

ほかに疫学データの検討では、観察研究に基づくので交絡が無限に存在とか、高用量ばく露領域が観察されにくいとかいろいろなことがありますけれども、今回はこの評価技術では今、テーマに出している話ではないので短く割愛しますが、ほかにも問題がいっぱいあるのですよというようなことも踏まえた上で、疫学データに関しては締めくくりたいと思います。

ハイブリッド法と比較して分位点回帰ができると、データがどんな感じで分布していてもよいし、計算コードの記述であったり実装できる限りではBMRに対応するドーズを分位点回帰の解から簡便に計算できるというのは変わらないし、モデルが変な想定をすることはない中で柔軟に解が求まりますということです。

以上までのところが手順と判断基準で、残っている部分、国外でどうなっているかというのとソフトウェアがどうなっているかというのはちょっと走りながら説明します。

BMD法というものにベイズ推定を活用することについては、国際的にももう動きが現れています。ガイダンスとしてベイズ推定を使いましょうというのに関しては、国連機関と欧州で言及があります。Environmental Health CriteriaはWHOが出しているもので、広瀬先生もこれの策定に関与されていますけれども、その240というのはもう公表されましたが、その中でBMD法にはベイズ推定を利用することが推奨されるというふうに言及されることになりました。

EFSA、ヨーロッパの欧州食品安全機関でもガイダンスの改訂案が、今、案でコメントを求めている段階がちょうど終わったぐらいですけれども、まだ公表にはなっていないと理解していますが、そこでもベイズ推定を積極的に推奨することになっています。なので、それに対応したソフトウェアが出てきてはいるのですけれども、BMDSは近似解、ラプラス法を用いたベイズ推定をやっている。BBMDは完全なMCMCということです。PROASTは今、頻度論ベースということになっています。

こういうこともあったので、BBMDという最もエクソプリシットなMCMCを使っているソフトウェアはウェブセミナーをたくさんやっているのです、私たちも夜中、何回か徹夜をして

セミナーに出ました。ワークショップに参加して、ユーティリティーとか二値データ、連続値データとかを使っていて、ああ便利だなと思われながらやっていました。またその話は後でカバーします。

それと、ToxicRというソフトウェアがRのパッケージでできそうです。これは広瀬先生がSociety for Risk Analysis 2022に参加して入手してくれてくれた情報なのですが、アメリカのEnvironment Health Scienceの国立研究所にもともとBMD法にとっても見識の深い研究者のグループがありますけれども、その人たちがBMDSの問題を解決した、BMDSとほとんど同じ機能を実装したものをR言語で書いたよというものをパッケージ化しました。パッケージの査読も確か終わったと思うのですが、その中にベイズ推定をMCMCでやるということもできるようになりました。ただ、これはアメリカのBMDSとは独立のものではあるのですが、それを使うとBMDS風のソフトウェアを利用しながらMCMCもできるようになりました。

ウェビナーイベント、これは研究班があったので出させていただいて、オンライントレーニングを受けさせていただいたのですが、ユーティリティーとかサーバーの利用に当たっての注意点も含めてやっていて、国際的にはたくさんの方々が出席されていて、ユーティリティーを知るためには出ていてとても有益なものだなと思われました。

EHCとEFSAでベイズ推定を使ったほうがいいですよということに関してはお話ししたとおりで、その詳細の文章はスライドが皆さんのお手元にあるので、御興味のある方だけ読んでいただければと思っています。

ソフトウェアを用いる場合の手段、運用なのですが、試験運用とインプット、アウトプットの調査とかを自分たちでしましたということと、既存ソフトウェアがどういうものかというのを実際に足を使って調べてきました。ラプラス法というのが使われているのは、BMDSを起動するとこんな感じのものが開くのですが、GUIといってマウスでカチカチとアプリケーションを開いて、こういうウインドウが開くとやはり使いやすいですね。そのために実装されたソフトウェアで、メニューバーがある場合もありますけれども、プルダウンで何をやりたいかを入れて、あとは「Run」というのをクリックするだけで済む。

連続データに対してのBMDSとBBMDの比較がたくさん行われているのですが、BMDSというのは連続データを使った場合でも微妙なデータだったらちょっと失敗が多いということが問題になっています。それを恐らく変えるだろうと思われているのがToxicRです。ToxicRのトレーニングコースは広瀬先生が出られているので、また追加があれば発表後に追加していただければと思いますけれども、BMDSとかを大きく超えたGUI制限を受けないようなパッケージということをつくっています。Rでパッケージをこれは自分のコンピューターで走らせているのですが、ファミコンを彷彿とさせるような起動画面が出てくるわけなのですが、これを打つ暇とかもあるぐらい、これが好きなんだということが

伝わってくる初期画面ではあるのですが、使わせていただいた限りでは、ベイズ推定とかをやっている、確かにBBMDみたいに誰かのサーバーにデータを投げることをせずに、自分のコンピューターの中だけでやることになる、このToxicRを使っていくと全部が終わるのだなということとかも把握をすることぐらいまでは何とかできましたということでした。

長くなり過ぎてはいけないので、まとめさせていただきますけれども、3点あります。ベイズ推定をBMD法に活用する手順と判断基準とかを検討したのですが、ベイズ推定は、少ないサンプル数のデータというのは動物実験ではすごく多いのですが、十分に不確実性を定量化しながらBMDLを推定できるという点ではやはりアドバンテージがあります。役所的に困ることは事前分布の想定で別でリスク評価結果に影響があるということなのではあるのですが、この点に関してはこれから解決されていくと思います。手順書というものが海外でも出ざるを得ないということになると思うので、ベイズのモデル平均化に関しても、モデル平均化の後のベイズのp値の見方とかに関しても、恐らくガイダンスがこれから出てくることになると思います。それに対してのトレーニングを一定の度合いで日本でもやらざるを得ないと思います。

もう一個、国外のリスク評価機関では、ベイズモデル平均化を用いましょうというのが『Environmental Health Perspectives』とかで広く議論されていますので、強く推奨されるという流れ自体はもう止まらないと思いますので、今後、リスク評価機関では実装が少なくとも必要になってくると思うので、コンピューターとできれば長期的には人材を準備していく必要があるのかもしれない。

ソフトウェアに関してですけれども、MCMCを実装していてとても高い信頼性を誇っている。分布としてとても便利に推定結果をビジュアルに見せてくれるという点で、ToxicRであったりBBMDがとてもよく実装されているので、もうそれはどんどん使えるようになっていくというのが重要なのではないかと考えています。

以上で終わります。ありがとうございました。

○川村座長 ありがとうございました。いろいろ御苦勞も含めて御紹介いただきまして、ありがとうございます。ただいまの成果の御発表につきまして、何か御意見とか御質問はございませんでしょうか。

念のための確認を私からさせていただきますけれども、分布の形による影響ですが、その辺りは影響の大きさ、最初の分布を与えるということから始まるとお聞きしましたけれども、その与え方。実際に先生もスライドでフラットなものと特定の型の分布がありましたけれども、あれは極端な例だと思っておりますが、実際にあり得る分布の形の中で、得られる結果の振れ幅というのは感触としてつかめますでしょうか。

○西浦専門委員 もともと一つ一つのパラメータに分布を仮定するのは、そのパラメータについて推定をする前に何か知っているからというときに相当します。ですので、もともと知っている限りではこれだけでもというのでスタートをせざるを得ないわけなのですが、何も知らなかったらフラットにやる。何か知っていたらそれに相応する分布を



仮定しているというのが根底にある考え方なので、まずはそこからスタートということなのだと思います。結果に対する影響というのは、事後分布の分散を見たりとか、解の安定性を見たりとかして検討していくことになると思います。

○川村座長 ありがとうございます。確かに臨床でベイズを使うときもいわゆる経験値というのが事前の分布なり事前確率になったりするのですけれども、それなりの知識、例えばEFSAが先行していればEFSAのやったものが事前分布、結極一緒になってしまうかもしれないのですけれども、何らかの専門家が持つ知見とか、あるいは海外で評価されて使われたものが事前の分布として実際に使われ得るものと考えてよろしいでしょうか。

○西浦専門委員 実際の毒性評価でまだフォーマルにBayesian BMDというのが使われたことはないわけなのですけれども、でも、実際にやる場合は、恐らく一個一個のデータベースを見ていくことになると思います。毒性試験の結果をデータベースにしているものはアメリカでもありますし、あるいは日本だと国立衛研でもあるわけなのですが、対応する化学物質に関しての動物実験のエビデンスがあれば、まずそれにフラットで適合した結果があって、その上で新しいデータの分析をするというコンテキストで使うことになるのではないかと思います。

○川村座長 国際的にまだ経験が浅いけれども、論理的に優れているので、あるいは試行結果が優れているので強く推奨されるという経緯でしょうか。

○西浦専門委員 結果が安定的だということが一番の理由だといつも言われています。どうしてもランダムなチェーンを走らせてやるいいかげんな推定なので、得られる結果というのは、物すごくたくさんチェーンを回さない限りは必ずしもいつも一緒というわけではないのですね。ですけれども、本当に真実とされる平均値だとかをいつも信頼区間の中に入っているというような、そういうパフォーマンスとして再現性がとてもいいので、小サンプルでも信頼に足る結果が得られますというふうな表現がいつもされているのだと思います。

○川村座長 ありがとうございます。人間の日常の行動はベイズのほうに近いかなと感ったりします。

ほかの委員の先生方から御質問、御意見はございませんか。広瀬先生、何か補足といたしますか、コメントはございますでしょうか。

○広瀬専門参考人 実は私自身はベイズというところについてはあまり専門でないので、難しく、まだ完全に理解はできていませんけれども、ソフトを触った感じでは、BBMDとか、操作性はいまいちですがToxicRというのかなり課題が整理されてできているかなという感じは受けています。それに似たシステムは、恐らく、EFSAでも今、開発中なのだろうなと思って、まだできていないのですけれども、そういうプラットフォームが今整備されつつあるという、国際的にはそういう状況ですね。

ただ、事前確率については私もちよっと勉強不足というか、どうしたらいいものかという話なのですけれども、同じ研究班で西浦先生に質問するのはあれなのですが、一度フラ

ットな事前分布で計算した結果を用いて、その次からはそれを用いたのを使っていくのがいいのですかね。今、現実的に解決しようと思うと、多分同じようなデータセットを探してきて、まずそういうのを幾つか整備したほうがいいのかなと思ったりしたのですけれども、どうですかね。

○西浦専門委員 どれだけその結果が違うかによって、その指針の詳細がまた出てくると思うのですけれども、一番簡単なのは、特定の食品添加物とか化学物質についての毒性を検討するときに全く同じ化学物質の動物実験データで、全く同じアウトカムの過去のデータがあれば、それにフラットな事前分布を仮定して推定した結果のポステリアをそのまま使うとかいうのが健康的なのです。

一方で、ほかの類似の生理的な反応を起こすということが既知であって、ほとんど同じような作用をする化学物質のときにもその範囲を広げるのかとか、あるいは全然生理的な反応とか病理も違うけれども、過去にワイブル分布で特定の化学物質だったらこういう形状になったというそれぞれのパラメータの形状が分かっていたら、それを全く違う化学物質にも適用可能とある程度仮定してよいのかというような辺りが今後詰められてくることになると思います。

○広瀬専門参考人 ありがとうございます。

○川村座長 ありがとうございます。ほかに御質問はよろしいでしょうか。

それでは、そろそろお時間ですので、本当に西浦先生、いろいろ御苦勞をありがとうございました。これから取り入れることを前提にというか、具体的にどういうふうにするか、事前の手順とか事前分布をどういうふうに捉えるかといったことを具体的に考えていかななくてはならない。それから、最後で詳細には触れられなかったのですが、たくさんの問題点がありますと言われたので、その問題点についてきちんと、右往左往しないように問題点をしっかり意識して、現時点で考え得る対応を検討していかななくてはいけないかもしれないなと思いました。

したがって、このワーキンググループとして引き続き御指導いただくことになると思いますけれども、くれぐれもよろしく願いいたします。

今日は限られた時間の中、いろいろな御意見を頂戴し、また、最新の情報を御提供いただきましてありがとうございました。それでは、本日の審議は以上とさせていただきます。

事務局から事務連絡をお願いいたします。

○今木評価専門官 次回のワーキングの日程につきましては、8月31日水曜日午前中を予定しておりますが、詳細につきましては、座長とも御相談の上、決まり次第、事務局から先生方にお知らせいたします。

以上でございます。

○川村座長 ありがとうございます。

それでは、長時間にわたりましたけれども、これで第25回会合を終了いたします。

ウェブシステムでつなぎにくいところもあったかもしれませんが、御協力ありがとうございました。それでは、皆様、どうもお疲れさまでした。失礼いたします。