

食品安全委員会評価技術企画ワーキンググループ

(第32回) 議事録

1. 日時 令和5年9月6日(水) 10:00~11:50

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) 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

(専門参考人)

広瀬専門参考人

(食品安全委員会)

川西委員、脇委員

(事務局)

及川事務局次長、紀平評価第一課長、前間評価第二課長、井上評価分析官、
寺谷評価調整官、爲廣課長補佐、岡谷評価専門職、北澤技術参与

5. 配布資料

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の一部改正案

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)

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 Guidance on the use of the Benchmark Dose approach in risk assessment (EFSA)

参考資料4 Benchmark Dose Technical Guidance (U.S. EPA, 2012)

参考資料5 Environmental Health Criteria 240 (EHC 240) PRINCIPLES AND METHODS FOR THE RISK ASSESSMENT OF CHEMICALS IN FOOD chapter5 Dose Response Assessment and Derivation of Health Based Guidance Values

6. 議事内容

○岡谷評価専門職 定刻となりましたので、第32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。

開催通知等で御連絡いたしましたように、本日の会議は公開で行います。

また、本会議は当会議室への参集及びウェブ会議システムを併用して行います。傍聴につきましても食品安全委員会のYouTubeチャンネルにおける動画配信により行っております。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。

本日は10名の専門委員（山田専門委員が欠席）が御出席されていますが、2名少し遅れていらっしゃるようです。

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、2名の委員が出席しております。

御発言に当たりまして、2点お願いがございます。

1つ目、御発言の際には、会場にいらっしゃる先生は挙手にてお知らせください。ウェブシステムの先生は、挙手していただくか、「挙手」のボタンでお知らせください。

2つ目、御発言の際には、冒頭にお名前をおっしゃってください。傍聴者や速記者にどなたが発言されているかを明確にするため、御協力をよろしくお願いいたします。

それでは、以降の進行は川村座長、よろしくお願いいたします。

○川村座長 皆様、おはようございます。

それでは、お手元の議事次第を御覧ください。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」です。

事務局より資料確認をお願いいたします。

○岡谷評価専門職 本日の資料は、議事次第、専門委員等名簿、座席表、資料1、参考資料として1から5までお配りしております。座席表につきましては、今朝ほど資料を差し替えさせていただきましたので、よろしくお願いいたします。

会場の皆様は、参考資料につきましては、お手元のiPadのブックというアイコンから御覧いただけます。

不足等がございましたら、事務局までお知らせください。

以上です。

○川村座長 ありがとうございました。

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を行ってください。

○岡谷評価専門職 それでは、御報告いたします。

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でした。

以上です。

○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。

よろしいですね。

それでは、議事に移ります。

「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」、前回の第31回ワーキンググループに引き続き、審議を進めます。

前回ワーキングでは指針全体を通して審議を行い、幾つかの修正の御意見や宿題の出たところもありました。また、再度全体を見渡して気になった点なども出てまいりましたので、本日はそれらの修正点を中心に議論を進めます。

それでは、資料1について2つに区切って議論していきたいと思います。

まず前半部分、冒頭から第3のI. 動物試験への適用について、事務局から説明をお願いいたします。

○爲廣課長補佐 事務局でございます。

それでは、資料1について御説明させていただきます。お手元に資料を御用意ください。

こちらの資料は、現行の指針からの修正部分を見え消しで示しております。そのうち、青色でハイライトした部分が前回お示しした資料からの修正箇所でございます。

では、表紙をおめくりいただきまして、1ページ目、注釈1を御覧ください。これはPODの用語の説明部分ですが、寒水専門委員から、説明文中にある初出の略語についてフルスペルが必要ではないかと御指摘をいただいておりますので、NOAEL、HBGV、MOEのフルスペルを追記してございます。

次は第2の定義についてでございます。ページをおめくりいただきますと、10行目の6ポツから8ポツまでデータの種類について記載がございますが、現行の指針には同腹仔影響があるデータや有害事象の数をカウントした計数データについて言及しておりませんでしたので、8ポツの連続値データの後段に計数データに関する記載を追記させていただきました。

21行目から「なお、乳頭腫の数等の単位（時間や空間等）あたりの発生数について測定した計数データを連続値データとして扱う場合もある」と記載しております。

続きまして、4ページ目を御覧ください。

ここから第3としてBMD法の活用について記載してございます。

まず、11行目、12行目は言葉の体裁を整えた修正となります。

そして、14行目から動物試験で得られた用量反応データの適用に関する記載となります。

1ポツの用量反応データの収集、選択について、以下に(1)から(4)の項目に分けて記載してございます。

まず、16行目の(1)の見出しの部分を御覧ください。こちらは前回のワーキング時には「データの収集」としておりましたが、24行目下の枠内にございますように、手順の流れがより分かりやすくなるよう、「動物試験の結果の収集」と修正いたしました。また、これに伴い、18行目から21行目にかけて本文の用語を整えております。

続きまして26行目、(2)の見出しにつきましても、前回ワーキンググループでの議論を踏まえ、「データセットの選択」と修正しております。

では、次に5ページ目、下の脚注4を御覧ください。こちらは、同じページの4行目にある共変量の説明をした注釈でございます。寒水専門委員から、モデリングと解析という用語が並列で扱われていることに違和感があると御指摘いただきました。御提案のとおりモデリングを削除し、「解析のために付け加える変数」と修正しております。

続きまして、8行目から順序カテゴリーデータに関する記載がございます。前回ワーキンググループで御議論いただきましたように、括弧内の順序カテゴリーデータの二値化に関する記載は脚注といたしました。なお、この脚注はグレードについての注釈としております。コメント枠内にございますように、竹内専門委員より、順序尺度の意味を考えるとグレードは「反応程度」というより「反応程度の順序」が正しいのではないかと御意見をいただいております。御指摘を踏まえ、説明文を「順序づけされた観察された影響の程度のこと」と修正しております。

また、寒水専門委員から注釈最後の一文の二値データに関する記載が不適切と御指摘いただいております。御提案いただいている案のとおり、「病理所見の有無を二値データとして取り扱うことを検討する」と修正しております。

続きまして6ページ目9行目、(3)を御覧ください。以前はデータセットの統合と原データの入手が同じ項目になっておりましたが、下のコメント枠にございますように、竹内専門委員から、疫学研究の項の記載とほぼ同じであるため、同じにしてはいかかかと御提案いただきました。そこで、御提案のとおり、疫学研究の項との整合性を考え、「(3)原データの入手」、「(4)データセットの統合」に分けて記載することとし、文章も一部整えております。

続きまして、11ページまで進んでいただきまして、2行目を御覧ください。こちらは4ポツの結果の文書化に関する記載で、連続値データをBMD法に適用した場合に必要な試験結果の概要を記しております。マーカーの部分については、寒水専門委員から、ばらつきの指標は「四分位範囲」が適切と御指摘いただいておりますので、御指摘のとおり修正いたしました。

冒頭から動物試験データに関する記載までの指針の前半部分について、前回ワーキンググループからの変更点は以上となります。

御説明は一旦ここで終了させていただきます。

○川村座長 ありがとうございます。

前回のこの会議でいただいた意見を基に、事務局と一緒に修正案を作ったのですが、いかがでしょうか。内容的にあるいは言葉の選び方などにつきまして、何か御意見はございますでしょうか。

根幹を揺るがすような修正はしていないつもりですけれども、少し注意すべき点とすれば、2ページ目の8ポツ、19行目から始まる連続値データという記載がありますが、その中で離散量である、例えば児の数とか腫瘍の数とか、そういった飛び飛びの数、非常に数が多くなれば連続量と一緒にすけれども、10未満ぐらいの数個とか数匹といったものにつきましては飛び飛びの値になるので、この取扱いをどうすべきかということで検討してもらいましたが、連続値の中に入れて、その中で説明しておくという形を取っております。補正項が入ったりするので、取扱い上、一部異なる場面もあると思いますが、考え方に極端な違いがあるわけではないので、データの種類としてはここに含めて記載しております。

それから、4ページの中ほどからですが、データという言葉が錯綜してしまいましたので、データというのは一人一人あるいは一匹一匹が示した生の値を指しますので、データセットであるとか、それから、研究あるいは試験という言葉で包括的なものは用語を置き換えたということで、区別をはっきりさせたというところです。

あとは、6ページ目の9行目から、原データの入手とデータセットの統合は別々に記載するということですが、これは分けたということで、これも大きな内容の変更ではありません。

こういった辺りですけれども、あとは用語の使い方などは先生方の、特に御専門の先生方の御指導に従いまして修正をしております。

ここまでのところで何か御指摘はございますでしょうか。よろしいでしょうか。

ありがとうございます。

それでは、ここまでのところは特になければ、微調整という程度ですので、考え方を根本的に変えなくてはいけないというところはないと思いますけれども、よろしければ、ここまでのところはこういう記述でまとめさせていただきたいと思います。

続きまして、後半部分を進めましょう。

○爲廣課長補佐 事務局でございます。

それでは、11ページ目を御覧ください。

22行目になりますが、ここから疫学研究で得られた用量反応データの適用についての記載になってございます。

24行目、(1)の見出しは、前回ワーキンググループでのデータの用語に関する議論を踏まえて、手順の流れが分かりやすくなるよう、「疫学研究結果の収集」と修正いたしま

した。これに伴い、本文中の用語も整えております。

続きまして12ページ5行目、(2)につきましても、前回ワーキンググループの議論を踏まえて、見出しを「データセットの選択」と修正し、本文中の用語も整理いたしました。

9行目、適切なデータセットの選択については、用語を補いまして、「信頼性が高く、PODを求めることが妥当と判断し得る適切なデータセットを選択する」と修正してございます。

13行目から、データセットの選択の際に整理する項目として①から⑥まで列挙してございます。

13ページ目の4行目、⑥のバイアスに関する記載のところでございますが、もともと「バイアスの存否」と記載してございましたが、寒水専門委員から、バイアスは程度の問題であって、ないことを保証することは難しいと御意見をいただいております。御指摘を踏まえて、こちらの記載は「バイアス並びにそれに対する対策若しくは考察」と修正いたしました。

18行目のマーカー部分は、竹内専門委員より句点の脱字を御指摘いただいておりますので、修正いたしました。

20行目から2ポツのBMRの設定についての記載でございます。

21行目から23行目までは、用語について整えた修正となります。

24行目からの記載につきましては、次の14ページのボックスにございますが、前回ワーキンググループでの議論を踏まえ、健康影響の社会への許容度と規制値を導くための信頼性の2つの観点からBMRの設定に関わってくる事が分かるように、説明文を修正しております。こちらの説明文は、同じボックス内がございます寒水専門委員からの御意見を踏まえて、事前に先生方に確認していたバージョンから一部修正させていただいております。13ページ目の24行目から読み上げさせていただきますと、「その際、対象とした健康影響等の大きさ」、括弧は飛ばしますが、「とその精度を考慮する。さらに、ばく露領域等の用量反応関係を的確に反映した数理モデルを用い、用量反応曲線の立ち上がり点等が合理的に推定できるようにBMRを設定する。なお、ヒトの疫学研究であってもBMRによって導出されたBMD関連指標がそのまま健康影響に基づく指標値とはならないことに留意する」と修正いたしました。

14ページ目、3行目に記載があるBMRについて、脚注にBMRの設定方法を例示してございます。疫学研究で得られる用量反応データの傾向を踏まえて、設定方法に関する記載をこちらに記載してございますが、コメント枠内がございますように、祖父江専門委員から定義のところでも統一的に説明したほうがよいのではないかと意見をいただいているところで、定義のほうに記載すべきかどうかについては、ワーキングの中で御議論いただきたく存じます。

続きまして16ページ、1行目につきましては、前回動物試験の項で御議論いただいた内容と同じですので、同様に修正しております。

ページをおめくりいただきまして17ページ目、23行目から4ポツ、結果の文書化の記載になります。こちらは疫学研究が対象となるため、31行目の③の項目名は「研究結果概要」に修正してございます。

また、32行目から記載しております概要に関する記載については、前回ワーキンググループでの御議論を踏まえ、二値データや順序カテゴリーデータといったデータ変数ごとの区分ではなく、扱う情報ごとに整理をし直しまして、a 評価対象物質のばく露に関する情報、b 各ばく露量における研究対象者数、c 各ばく露量における健康影響等の反応量に関する情報、d cに関するばらつきのように修正しております。

また、③の内容を変更したことに伴い、④にデータセットに含まれる情報を変換した場合について追加し、これに伴い、以後の付番を修正しております。

説明は以上となります。

○川村座長 ありがとうございます。

疫学研究のところですが、ここはいろいろ議論の余地があったかと思います。用語の統一は先ほどの動物と同じような感じですが、特に重要と思われる点は、13ページ目の20行目からあるBMRの設定、それから、祖父江委員から御提案のありました、BMRの言葉の説明についてどこでどういうふうに表記するかという問題あたりから議論してまいりたいと思います。

疫学研究というのは生のヒトでやっているのだから、BMRで得られたものが社会的な規制につながりやすいという印象を持っていらっしゃる方も多いと思います。しかし、それをストレートにやりだすと、どこまで社会が許容するかという問題とか、あるいは実際に到達が可能かどうか、それから、そもそも微妙なところを数理モデルで描いた曲線で決定できるかというところがあるので、いろいろ考えまして、14ページの上の記述は熟慮した結果なのですが、基本的にはそういった社会的な要請、許容するという言い方が適切かどうかは分かりませんが、社会的な観点から影響の大きさをどう認めるか、どう把握するかということと、もう一つ、数理モデルで用量反応関係を得るわけですが、非常に微量のところ、実測値があまり存在しないレベルのところは真の値といえますか、推定が十分にできるかということもありまして、ただ、ここで得られたBMRの値で導出されたばく露量というのがそのまま規制値になるわけではなくて、用量反応関係が明確に示されるところを得ると考えております。

要するに、健康影響の社会への許容度といえますか、社会的な影響と、それから、もう一つは規制値を導くための信頼性の観点からBMRを決定する。それはどういうことかというところ、例えば10のマイナス5乗とかマイナス6乗とかという確率のところは、多分通常のモデルからゼロに近いところのレベルを決定するというのはなかなか難しいのではないかと、ここだったら確実な用量反応関係が得られるミニマムな線を得て、がんでよくやられているように、そことゼロ点とを直線でつないで、社会的に認められるポイントの健康影響の大きさのところのばく露量を採択するというような考え方もできようか

と思っております、したがって、BMRでそのままHBGVの点とかあるいはリファレンスドーズをストレートに決めてしまうわけではなくて、その間に吟味する余地があるということを考えてこのような表現にしたところです。

ここについては大いに議論があるかと思うので、皆さん、御意見を出していただきたいのですけれども、いかがでしょうか。

祖父江先生、どうぞ。

○祖父江専門委員 祖父江です。

これは、疫学のほうだけにBMRの意味合いの説明があって、動物のほうにはあまりないですね。これは質問ですけれども、BMRというのは用量反応曲線の立ち上がりの点が合理的に推定できるように設定するものなのですか。用量反応関係では独立にこういうものだと決めるものではないのですか。だから、説明できるように決めてしまったら本末転倒のような気が僕はするのですけれども、意味のある変化で、それは立ち上がりがどうかということとはまた関係なく決めるものではないのですか。

○川村座長 だから、それは2つの意味があって、ここまでは問題なしとするというような健康影響の社会的な影響の大きさから、ここまでは社会的に大きな問題はないというレベルという決め方と、もう一つは、そのレベルは非常に微量のところでは決め切れないので、代わりの方法として、変化としてはっきり認識できるポイントを決めて、一つの例示だけでも、そこを0点と直線で結んで、社会的な許容されるであろう点をモデルとは別に導くという考え方もできるかなというところです。

つまり、10のマイナス5乗とか6乗とかというのを数理モデルの中で議論することは困難ではないかと思うので、この辺り、後で西浦先生のコメントもいただきたいのですが、この用量に対する反応は確かなものであるであろうというところの見え得る最小の点を決めて、そこをPODにはするのだけれども、しかし、そこから実際のHBGVを出すときはもう一作業する。微量のところでは、例えば1%というところは決められるかもしれないのだけれども、発がんで、しかも頻度の高いがんで1%というところ、例えば肺がん死亡だと10万人ぐらいいるわけですね。1%というところ、1,000人とか、そこら辺を許容し得ないとする、でも、そこから下は、1%より低いところというのはなかなか曲線の形が、モデルだから仮定しているのだけれども、実測値がないところのデータを読み切れるかということがあると思うのです。

そこら辺が皆さんの意見を聞きたいところではあるのだけれども、そうすると、確実にここは言えるという点を決めて、そこから安全サイドに立って直線でつなぐ。本当は凹型にへこんだ形になるから、もっと反応量は少ないはずなのだけれども、それが言えるかどうか分からないので、直線で結ぶといった作業。直線がいいかどうか分かりませんが、数理モデルが平均化とか何かしたとしても、それがごく微量のところでは当てはまっているかどうかということが分からない場合に、例としてですけれども、直線で結んで、安全サイドに立った許容できる量を決めるというような、そういうやり方もあり得るのでは

ないかということでこういう文章にしたのですが、西浦先生、どうですか。微量のところのモデルの適用性というか、確実さとか、その辺、コメントをいただきたいのですが、お願いします。

○西浦専門委員 ありがとうございます。

先に祖父江先生のお話にお答えしたほうが多分スムーズに行くと思うのですが、そもそもBMRの定義を振り返っておきたいのですが、BMRというのは基本的には単なる超過リスクだと理解しています。だから、そもそもばく露がなかったときに、何らかのアウトカムに対するバックグラウンドレスポンスレートというものがあります。だから、ばく露が全くなくても、これぐらいの率で病気が起こりますよというような率があって、それに対して、エクセスで1%とか、エクセスで5%、エクセスで10%というような度合いで、エクセスリスクがどれぐらい超過があるかというのがBMRの定義そのものだと思いますので、そういうふうにならば何らかのドーズレスポンスの曲線があったときに、バックグラウンドが何らかのコントロールとしてあって、それと比較してどれぐらいかというものと理解をしていますというのがまず1個目です。

その上で、BMRというのはBMDを定義するために重要になるということですよ。5%とか、10%とか、どこで切るかによってリファレンスとする場所が変わるということなので、まず14ページ目の1～3行目で書いていただいている辺りがBMD法の考え方とちょっと齟齬があるかもしれないと考えています。どういうことかということ、そもそも論的な話をしますけれども、ドーズレスポンスモデルというのでこのBMD法をやっているときというのは、今日までにほかの方法で実験医学とかに頼ってやってきたようなNOAELとかLOAELと比較して何がいいかということ、ハイドーズとかミッドドーズとかも含めて、ロードーズだけではなくて、結構高いドーズのときのアウトカムの発現率も含めて、ドーズレスポンスカーブがこんなにきれいにある簡単なモデルでとらえられるとしたときに、ドーズが減ったときというものをフィットしたドーズレスポンスカーブから外挿できるというようなフィロソフィーで使われているものだと理解しています。だから、ロードーズ領域ばかりを見ているわけではないのです。

一方で、でも、ロードーズの領域を正しくモデルしておいてほしいなというのがこういうリスク評価でドーズレスポンスモデルを使うときの願いなのですが、今のところ、ロードーズ領域だけがそれでうまくいっているとか、あるいは目に見える形でリスクが立ち上がるような話でありますように、何らかの閾値を置いて立ち上がるようなモデルだっているわけなのですが、そういうのがうまくいっていることを保証できるような感じの枠組みのモデルがいっぱいあるかということ、そうでもないということなので、14ページ目の1～3行目で書いていただいている部分は2つ節があるのですが、低ばく露領域の用量反応関係を的確に反映するというのを、そこだけの評価するのはまあまあ難しいと考えていますということと、もう一つ、立ち上がり点が合理的に推定できているということも、立ち上がるというような閾値があるというようなモデルをするとか、ある

いは目に見えている観察データでのリスクを記述することに優れているとか、何らかの条件設定がそこには必要だと思えるのですけれども、その辺りをルーチンに評価する仕組みは恐らく今の時点ではないのかなと思っています。

なので、ドーズレスポンスをこういう評価に用いる上でのフィロソフィーと部分的にここは反駁している感もあって、こういう感じでBMRを定義するのは難しいなと感じながら今の議論をお伺いしていました。

取りあえず以上です。

○川村座長 ありがとうございます。

中用量とか高用量だとぶれの範囲が相対的に小さくなるけれども、微量のところだとほんのちょっと曲線がずれたただけですごく大きく動いてしまう。だから、微量のところのちょっとしたぶれというのは、高用量と同じ程度のぶれであっても、実際に出てくるBMDの絶対値の大きさというのはものすごく変わってきてしまうので、要するに、もともと回帰曲線自体というかモデル自体が一定の誤差を含むものなので、微量レベルになると誤差がすごく大きな問題になってしまわないかという気がしたのです。それで、立ち上がりと言ったのはミスリードだったので訂正しますけれども、立ち上がりというよりは用量反応関係が確かに得られる、この用量反応関係は確かであろうという点を決めて、そこから先は別の演算でベンチマークドーズあるいはリファレンスドーズを決めていくという考えで、微量のほうも高用量のほうの絶対値は変わらないのだけれども、元の値に比べて相対的な誤差が大きくなってしまって、ここの微妙なところを日本として政府として決めるといときに影響が強過ぎはしないかということで、実際に10のマイナス5乗とかマイナス6乗という社会的に容認されるとされる確率のところの議論がどのくらい確実にできるのかということがあって追加した記載なのです。

確かになじんでいない、こういう記載はないと思いますが、動物実験だとそもそも動物の頭数が限定的なので、そういう議論に入る前に10%レベルの議論になってしまって、その代わり、後で安全係数とかをしっかりと掛けて不確実性を打ち消すのですが、人間だと(HBGVに)生の値を使ってしまいそうな気がするので、合理的と思われる数理モデル、あるいは場合によっては平均化によってリスクをヘッジしたモデルであっても、やはり微量のところというのはちょっとした動きがすごく結果に大きく影響してしまうので、一応見るところまでの点を選んでおいて、そこから後は措置として、ベンチマークドーズと言いかどうかは別として、合理的なリファレンスドーズにしていくと考えてこういう文章にしたのです。

確かに立ち上がりというのは閾値になってしまうので、これは表現がまずかったなと思いますので、用量反応関係のデータが比較的確実であるという趣旨のものです。立ち上がり点というよりは、このレベルになればより反応関係が比較的確かという点を決めておいて、あとは用量反応関係のモデルを数理モデルを離れて、あとは人為的な措置として直線でつなぐなり、それ以外の合理的なつなぎ方をして、社会が許容するレベルのところのド

ーズを決めていくということなのですが、そういう考え方は今まであまりなかったかもしれないので、EFSAなどもそういうところまでは書いていないかもしれないのだけれども、実際にこういった確率論的なものの考え方をしていくと、どうしてもぶち当たってしまうかなと思った次第なのですが、西浦先生、これについてさらに追加の御意見はございますでしょうか。お願いします。

○西浦専門委員 先生のおっしゃるような考え方というのは、放射線の健康影響とかがんの健康影響でばく露をどこまで認めるのかという話で、今までも実践で考えられていることだと思うのですが、それはこのBMD法を使う上では、恐らくですけれども、BMRをどこで設定したかというのを、この書類の最後にもありますが、設定根拠とかを書かないといけないということが後からあるのですけれども、そういうポストホックのエバリュエーションのときに重要になる話なのではないかと思っています。

どういうことかという、BMR設定というのをやる段階では、ロードーズでどれだけうまくいっているかという評価指標が現状でない限りは、決まった手続で、えいっとやっていくしかないというのが実情なのだと思います。一方で、その不確実性を動物実験の結果だとか先行研究だとか幾つかのものと比較すると齟齬がないということを確認しながら、ここで設定します。一方で、結果が出た後にこのBMRでよかったということを最終的に説明しないといけないときに、許容に関するフィロソフィーという話が出てくると思われるのです。なので、ポストホックエバリュエーションのときに、例えばですけれども、NOAELと直線ゼロとの間をつないでみた結果と今回のBMRで設定されている分位点というのがあまり齟齬がないということが評価指標の一個に入っていると、それでスムーズにこれよかったという解釈ができるという性質のものではないのかなとは感じました。

感想みたいになっていますけれども、以上です。

○川村座長 ありがとうございます。

祖父江先生、コメントはありますか。

○祖父江専門委員 西浦先生のBMRの定義としては、超過リスクであると。バックグラウンドに比べての相対的な増分ということですかね。そういうことなのか、この脚注に書いてあるのは、それが場合もあるということで、原則的には最大、最小といますか、最低、最高の振り幅に対しての何%かというのが一般的であって、恐らく動物実験だとそちらのほうではないかなと僕は思うのですけれども、バックグラウンドレスポンスレートを動物実験の場合はどう考えるのですか。

○川村座長 それは多分話がちょっと別というか、BMRの決め方には2通りある。動物実験の場合は変動幅の中の相対的に、つまり、ゼロのときはゼロということなので、バックグラウンドに対してどれだけ増えるかというのは言いにくいことが多いから、変動幅の中のどの程度かということBMRにするけれども、疫学研究は最大までばく露させられないので、主には低ばく露、せいぜいアクシデントのようなものがあっても中程度のばく露量までの範囲で見なくてはいけないから、最大値が分からない、振れ幅が分からない。そうな

ると、いわゆるバックグラウンド、自然ばく露のレベルに対して相対的にどれだけ増えるかという増分ですよね。いわゆるバックグラウンドレベルの反応に対する増分で見るのが、通常疫学ではそうになってしまわざるを得ないと思うのです。

○祖父江専門委員 そのことを動物と疫学とで違う点ですと書いてもらったほうが物すごく分かるのですけれども。

○川村座長 それは後で議論しようかと思ったのですけれども、先生の御指摘はもっともだと思ったので、実は用語の定義のところでも最初に言うておいて、座長の腹案ですけれども、疫学研究の場合はどういう根拠で何に基づいてどちらのBMRを取ったかと。それを特出しの部分だけ書いて、2通りあるということは用語の説明とかに移したほうがいいかなというのが、まだ事務局とも打ち合わせ切れていないのですけれども、私の腹案なのです。

だから、2通り考え方があって、最小ばく露と最大ばく露の間の反応量の振れ幅を100とか全体として、それに対する相対的な変動の量と、最小ばく露、バックグラウンドレベルの反応量に対してどれだけ相対的に増えるかというのと2通り考え方がある。どちらを採用したかというのは、動物実験だとほぼ自動的に決まってしまうと思うので、疫学研究の場合はどういう考え方に基づいてBMRを設定したかというのをきちんと記述しないとイケないということを書いておこうかなと。そうすると、用語自体はちゃんと最初の用語の説明で両方の概念が分かるけれども、疫学のところは特に注記して、どういう根拠でどちらのBMRなのというのが分かるようにしておかないと。本当はBMR1とBMR2というのがあって、どちらをどういう考え方で選んだかというのは記述しないとイケないということを書いておくということです。でも、どちらでやるにしても、BMRをどう決めるかという決め方のところで、一番望ましくは、それで社会的に許容される、許容という言い方がいいかどうかは別として、許容される値のところを設定したいのだけれども、そうすると、非常に小さな値になることがある。そのときに測定し切れないようなレベルのところの用量反応関係のカーブが、誤差が相対的に非常に大きくなってしまいます。要するに、用量反応関係の信頼区間は低域で小さく、真ん中が太ってきて、高用量になるとまた狭まるのですけれども、しかし、絶対値が非常に小さいところで、例えば日本でわずか全体で10人しかイベントが起きないところへ5人とかという影響があると、50%増えてしまう。1万人起きるところで5人増えても相対的に問題にならない小ささですけれども、どうしても小さいところはもともとイベントの発生確率が低いので、ちょっとした変化ですごくぶれるのではないかなという懸念があって、それでより確かなドーズに対してこの発生は確かですよというところを選んで、そこから後は人為的な操作で、西浦先生がおっしゃるように、それは後処理としてやる。だから、それをBMDと表現するかどうかは難しいですけれども、社会が許容するレベルとして、例えばがんが10人増えます、100人増えますといったレベルにしようと思うと、物すごく小さいBMRになってしまうので、そんなところをグラフから読めますかという感じなのですよね。

お願いします。

○祖父江専門委員 人におけるイベントの発生確率が低いものを扱うのは、がんとか重篤な疾患で低くてもやはり大きな意味があるからですよ。そのことに対応するために、疫学研究は物すごくでかいポピュレーションを扱うわけだから、動物実験とちよつと違うというか、その点が低い発生確率のものが扱える研究なわけですよ。だから、絶対的に低いということではなくて、先ほどのエクセスリスクなども相対的にバックグラウンドから10%とかというのであれば、別に薄っぺらいところを見ているのではなくて、それを拡大して10%というところを見るようにすればいいのではないですか。

だから、動物と疫学でやはり違うのは、サンプルサイズを小さな確率のイベントを扱うために大きくしているわけだから、そこは発生確率が低いからそこで配慮するというわけではないような気がするのですけれども。

○川村座長 そのとおりなのですよけれども、10%だったら問題なく見えるわけですよ。ただ、10%だとクリティカルな病気は当然許容されないから、そうすると、社会的に今、バーチャルセーフドーズが10のマイナス5乗になっていますけれども、10のマイナス5乗といった発生確率を限りなくゼロに近いところのものをグラフから読めるか。視覚的に見るわけではないけれども、グラフが表現できているかというところが気になるので、例えば10%だったら確実だから、あるいは10%は大き過ぎますけれども、1%の点だったら読めるから、1%を読んでおいて、そこから後は後処理で対応しましょうというようなことで、そういう意味で、ここに書いたのは、社会的な許容される水準がストレートに導き出せればそれでいいし、それができない場合はばく露量に対する健康影響が確かに見られるという点の中で最初の点を選んでおいて、そこから後は数値の処理、動物実験における、疫学研究でもあり得るのですけれども、安全係数を掛けるとか、そういった直線閾値なり仮説的なものを準用して直線で結ぶとか、そういうようなことが考えられる。

しかし、これはこのベンチマークドーズの根本からはずれるというか、後の話なので、ここには書き切れないうえけれども、そういう趣旨もあって、2つの考え方でBMRが選べるのではないかという趣旨で書いて提案してみたところなのです。

ほかの委員の先生方、こういう考えで御理解いただけますか。疑問とか意見がありましたら遠慮なく出してください。

寒水先生、お願いします。

○寒水専門委員 BMRの定義は、もともと定義のところに書かれているように、また、西浦先生が説明してくれたようにバックグラウンドレベルからのプラスとしての超過リスクとして解釈してきました。その上で、2つの考え方があるとしてしまうと、一つ気になるのは、国際的なBMRの定義と整合性があるかどうかということです。もし、日本には2つの考え方があるということになると、その後のBMDの解釈も定義に合わせて変える必要があります。そのあたりを心配していますが、いかがでしょうか。

○川村座長 ヨーロッパはあまり書いていないですよ。

事務局、お願いします。

○爲廣課長補佐 事務局のほうから御説明させていただきます。

ヨーロッパのほうではバックグラウンドレベルからの変化量としているので、後者のほうも含まれた表記の仕方になっていると思いますし、EPAのほうですとSDをBMRとして設定するという方法がございまして、その場合、バックグラウンドの標準偏差をBMRとして設定しておりますので、そちらも後者のほうに含まれてくるかなと考えております。

○川村座長 バックグラウンドの、つまり、ばく露ゼロで存在するのを基準にするという考え方は私も持っていますが、ばく露ゼロで反応量がゼロということが動物実験などでよくあるので、ゼロに対して何%という増え方は出せないから、変動幅の中でどれだけ増えたかということを議論することになるし、疫学研究の場合は最大量の振れ幅なんて分からないので、最大負荷なんていうのはできないから分からないし、でも、バックグラウンドは通常ゼロではない。ただ、バックグラウンドレベルは、置かれた疫学研究のコホートならコホートがどこにあるかによって随分変わってしまうのですよね。例えば放射線であれば、日本のバックグラウンドレベルと、それから、中国とかヒマラヤのあたりだったかな。それから、日本の中でも温泉地と非温泉地とでは基礎量が違うのですよね。だから、ベースラインも多発国と比して非多発国で全然違うので、どちらにしてもBMR自体がかなりいろいろな要素が加味されてしまって、なかなか一つの物差しで言い切れないところがどうしてもあるわけですよね。だから、どういう前提での議論なのかということをはっきりしていけないといけないし、確かに複数の定義が存在するというのは混乱を来すので、どういうものをBMRとして扱ったかとか、そもそも研究の母体になる疫学研究はどこでどういう人たちを対象に行っているかとか、そういうことがきちんと記述されないといけないと思います。そのくらいヒトを使った研究もばらばらだし、使えるBMRもどちらが使えるかがケース・バイ・ケースになってくると思うのですが、その辺がちょっと難しいかもしれませんね。

事務局、何かコメントがあれば。

○爲廣課長補佐 事務局でございます。

今、座長がおっしゃっていただいたのは疫学に考慮してというようなところなのですが、例えば検査の指標値を影響としたとき、検査の測定値がどれだけ健康影響に影響を及ぼしているか反応の変動が分からないような場合も動物試験の中でもあつたりすると思いますので、そういうことを踏まえてのEPAのほうでのSDの仕様だったことを記憶しております。

そういった背景を踏まえますと、特に疫学だけでそういったものを使うというようなことを限定するものではないのかなとは考えております。

○川村座長 おっしゃるとおり、事務局から今説明がありましたように、検査値などで見る場合は、当然ベースラインがゼロとは全然限らないわけで、ゼロが正常というのもありますけれども、例えば尿中のアルブミンとかそういうのは基本ゼロが正常でしょうけれども、血圧などだったら100とか、どこがゼロベースかは分からないのだけれども、当然ゼロ

が基準にはならないわけで、したがって、ものによるということですよ。ゼロが基本になるものと、ゼロが基本にならないものとある。イベントの数は理論的にはゼロベースなのだけれども、人間社会ではゼロにならないことも多いということですが、西浦先生、お願いします。

○西浦専門委員 今、皆さんが議論してくださっていることは、大体よくやられるリスク評価でBMRを定義するときの話に相当するのですけれども、多分14ページ目の脚注のところに寒水先生の今のコメントのオリジンとかもあると思うのですが、最初のセンテンスが難しいなと僕も思っていたのですけれども、要するに、バックグラウンドレスポンスレートと言われているコントロールをどういうふうに定義しているかということによって、BMRの定義をするときどこをベースにしているかというのが違うというのが事務局からもあったヨーロッパの対応の意味合いだと思います。だから、サンプル数が物すごく大きくて、コントロールもとてもぶれが小さいというときは、コントロールミーンを使って、それからの相対的超過ということでBMRを定義する。

一方で、EPAの話でおっしゃっていたのは、コントロールにちょっとぶれがあるので、コントロールをプラス1SDというところを出発点にして、そこから超過のリスクを取るのが実際の毒性評価とかでもやっているのを見たことがあるのですけれども、そういうことなので、そういうふうな書き方に変えともうちょっと分かりやすいのかなと思いましたがというコメントです。

以上です。

○川村座長 ありがとうございます。ここは表現が難しい。

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。

BMRの話は、WHOの会議のときも、やはりそういう統計的に反応が立ち上がるということだけではなかったですね。反応が有意になるときのバックグラウンドをどう設定するかという問題でそこは変わるのですけれども、数理モデル、統計学的にBMRというのを単純に決めると、寒水先生が言われたところなんですとんと落ちるのですけれども、結局、ポストホック的に説明が必要になってしまうだろうということを実はあらかじめ織り込んでBMRを決めるときがあるのです。だから、BMRを設定するときの順番が、理想的にはまず数字を見てこうだろうと言った後に生物学的な反応性、さらにはその後の影響までを加味してマージンが取れるか取れないかとか、有意な影響が出るところから直線外挿するとかを考えてBMRを決めるのか、手順の話も一緒くたになっているので、それがきつとこのガイドラインに記載する場所にも影響してきて、最初に書くのか、それを後に書くのかということと関わっているのかなと思っていて、解決は難しいのですけれども。

あと、寒水先生がおっしゃられた国際的にどうなのかという話で、確かにEFSAとかEPAはまずは数字を見てというところがあるのですけれども、WHOのガイドラインは、二値データは割とそれでいいのですけれども、コンティニューアスデータ、これは動物でも、疫学ほど

ではないのですけれども、母数の問題もあるかもしれませんが、例えば10%上がったコレステロールはどうかというような生物学的な重要性も加味しようとして書いてありますので、必ずしも統計だけでBMRを決めているわけではないという意見が、WHOのガイドラインはそういう書き方になっています。もし書くのだったら、WHOのガイドラインのほうを少し参考にするといいのかなと今思ったというところではあります。

解決というか、意見としてまとめ切れていないのですけれども、生物学的な有意性というのは、動物はまだ分かりやすいのかもしれないのですが、疫学の場合は、私は専門ではないのですけれども、得てして正常値は生物学的な反応というよりは、平均値を正常値としてそこから外れるのを異常値としていると決めている。要は、よくよく考えると、何も生物学的なバイオロジカルなレスポンスだけで決めているわけではないというところもあると伺っているので、そういうところが動物より疫学のほうが難しい点が多いのかなと。そちらは私の感想です。

すみません。参考にならなかったかもしれません。

○川村座長 ありがとうございます。

ちなみに、検査値の決め方は3通りありまして、今、広瀬先生がおっしゃったように、平均値を95%信頼区間などで決める方法は基準範囲と言って検査業界が使っている方法です。臨床では新しく開発されたマーカーみたいなものはそういうのを基準にすることがありますが、通常は長期に追跡するコホート研究でイベントの発生を見ていって、例えば血圧で言えば120まではイベントが増えないので、そこまでを正常とか理想値と言って、そこが閾値になって、あとは指数関数的に上がっていくので、140になるとリスクは2倍、160になるとリスクは4倍ということで、120までをよしとしましょう、130までを許容して正常にしましょうとか決めていて、長期予後との関連で決まっています。コレステロールも同じで、LDLコレステロールは閾値があるわけではないのだけれども、高ければ高いほど、冠動脈疾患、太い動脈の動脈硬化性病変は多くなりますし、がんの場合は反対方向になります。けれども、そういうことで、臨床的というか、ヒトの場合は基準値の決め方が3通りあって、もう一つが病理組織などと対応してこの辺から病的に変化が見られるということで、肝機能などはそんなようなものを病理と対比していると思いますし、要するに、平均値でやる方法もあるけれども、主流ではなくて、ほかに決めようがない場合に平均値でやっているということになります。

それは追加のコメントだけなのですが、いずれにしても、動物実験の場合もいわゆる数量データで検査値のようなものを使う場合は、もちろんバックグラウンドというかばく露ゼロでも健康影響のマーカーのほうはゼロにはならないことが多いので、ゼロのものもありますけれども、そうすると、ものによってベンチマークレスポンスの決め方が変わってくるということになりますよね。2通り考え方があって、そのマーカーに適したものであると。疫学研究の場合、どこにBMRを設定するかということ、もともとBMRが何を表すものであるかということが2通りあるということと、もう一つは、社会的な影響で見るというの

が原則なのだけれども、それだと非常に小さいレベルを議論することになって、果たして数理モデルでそれをピンポイントで決め切ってしまうていいかどうか分からないので、もしそれに不具合がある場合は、はっきりと用量反応関係があるという点まで高いところを選んでおいて、あとは安全サイドに立つ処理をするというような決め方がある、と今の私は整理しているのですけれども、これに対して何か疑問とか別の意見とかがございましたら、お聞かせいただきたいと思うのですが。

そもそも、そうやってBMR自体も一義的には決められないものであるということもだんだんよりはっきりしましたし、疫学研究であるから社会的影響だけで決められるというものでもないとも思いますし、結局、その都度何で決めるかということも議論して、広瀬先生の御意見を参考にすると、あらかじめある程度は設定するのだけれども、後から調整はする余地がある、みたいには理解しましたけれども、ケース・バイ・ケースということになってくる面も大きいとも思いました。

御意見はありますでしょうか。

では、基本線を座長が走っていいと思っているわけではないのですけれども、ここでの先生方の御議論を包括的にまとめるとすると、用語の定義のところ、ベンチマークレスポンスは、ばく露量にしても、それから、健康影響についてもそうですけれども、その数値が持つ、数値とは限らないですが、指標の持つ特性によってベンチマークレスポンスの意味合いは2通りあるということ。それから、もちろんさっき言った決められない場合にSD分だけ上乘せしてそこを基準にするというような変法もあり得ると思いますけれども、そういうふうに複数の定義の仕方があるということ。それから、疫学研究では動物実験のようにまず10%を検討してということが多分なされにくい。疫学研究もサンプルサイズがいろいろで、数百人レベルの研究から何万、何十万というレベルの研究もありますし、最近のようにいろいろなことが電子化で、アイスランドみたいに登録制になると膨大な数のデータになるので、悉皆調査に近いような研究まであるので、そうなる、それによって研究の精度というのが随分変わってきますから、シンプルに国民レベルで議論できるようなデータセットもあれば、推定するので信頼区間が非常に広がるという研究まであるから、これはまた一義的には決められないということになるので、そうなる、社会的な影響という観点を原則にしながら、それでは十分な検討ができないとか不確実性が強過ぎるという場合には、統計学的に合理的なベンチマークレスポンスを決定するというような書き方になるのかなと思っています。

具体的な文章は後で考えますけれども、考え方として、定義のところ複数あるということ、それを列記するという。だから、実際に使うときは、動物実験のところにも書かないといけないかもしれませんけれども、反応量の種類によってどちらのBMRを採択するかが変わってくるので、そうなる、どういう根拠でどういう考え方のBMRを選択したかをどちらも書かないといけないということになるかと思えます。

いかがでしょうか。なかなかずっとこれでいいですとも言いにくいし、さりとて代案は

と言っても出てきにくいかもしれませんが、いかがですか。

小坂先生、いかがですか。

○小坂専門委員 いろいろ聞いていて、そうだなと思いながら聞いているだけで、それ以上何かコメントするほどの知見はありませんので、こういうのも読む人が具体例がないまま議論していると、やはり分かりにくい部分もあるのかなと思っています。先生方の議論をそれぞれ聞くといろいろなことが見えてきて、図表が頭に浮かぶと少し理解しやすいのだと思うのです。多分オーディエンスというか、これを読む人も同じような状況になっているときに具体的な事例とかがあると、もうちょっと理解しやすいのかなと思って聞いていました。

以上です。

○川村座長 ありがとうございます。

この指針本体になかなか具体例を全部書き込めないですけれども、折に触れて何々などと括弧の中で例示みたいなものは今までもしているので、できるだけそういうものが簡単に書けるものであれば、例示をしながら記述を進めていくということは考えてもいいかなと思います。

竹内先生、いかがですか。

○竹内専門委員 聞いていて、ベンチマークドーズ法を使うというところと結果を解釈して社会実装するところの2つにまたがった話を一緒にやる羽目になっているところだと思うので、そこが難しいのかなと。特に疫学の結果だと、それを用いたベンチマークドーズ法の結果をそのまま使ってしまうそうと最初のほうで座長がおっしゃられた懸念が大きいのかなというのもあるのですけれども、社会的実装に耐えられる結果が出なかった場合というのを多分一番気にされているのかなと思って聞いていました。なので、例えば値が小さ過ぎるとかですけれども、そこはひょっとするとモデルではこうでしたという話をしておいて、別の話を入れたほうがスムーズなのかなと。ここから先の値は小さ過ぎてぶれが大きいから別の何かを外挿するというよりは、モデルではこう出ましたという話をした上で、社会的実装においてというエクスキューズをつけて1つ議論を増やしたほうが、社会的実装に対応するとか、そのまま結果を使ってしまうリスクとかというのに対応するのに、シンプルかどうかは分からないですけれども、ひょっとしたらごっちゃにならなくていいかなと感じながら聞いていました。ただ、僕もどうしていいかなというのは決め切れませんでした。すみません。

○川村座長 ありがとうございます。趣旨はよく感覚としてつかめましたけれども、書くのが難しいですね。

○竹内専門委員 全くそのとおりだと思います。

○川村座長 広瀬先生、お願いします。

○広瀬専門参考人 EHC240のchapter5にSelection of BMRというところがあって、そのやり方は川村座長が言っているのに少し近くて、段階的アプローチを取っていて、まず決

まったBMRがあるならそれを使おう。それは、動物の場合はそうですね。10%とか割とみんなのコンセンサスが得られているものがある場合は、それを使う。その次に、エキスパートがadversity-based BMR、生物学的に何か影響のあるところを、統計的だけではなくて、ここから本当に有意な生物学的な反応がジャッジできるならそれを使おう。次に、もうちょっと社会実装を考えたところで、MOEとかBMD以外のそういったアプローチもやって設定しようというようなアプローチが書いてあるので、そんなようなのが参考になるのかなと思いました。

○川村座長 ありがとうございます。

デフォルトみたいなものをつくるような感覚になるのだと思うのですが、疫学研究の場合もサイズもいろいろだし、なかなかあらかじめ何%というのを決めていいかどうか、非常に心理的抵抗が大きいのですけれども。

○広瀬専門参考人 これは動物も疫学も関係なく、まずあったらそれを使いましょう。だから、動物の場合はまずそこが当てはまるのですが、疫学の場合はそこがないので、いきなりステップ2に行ってしまう。BMRの設定の手順みたいなものをつくって、動物も疫学も通して設定したという意味で書いてあります。

○川村座長 エキスパートジャッジと書いてしまえばそれで終わってしまうのですけれども、そうすると、どういう観点を積み上げていくのかというのがなくて、その他のところでエキスパートジャッジは許容できるのですが、最初からエキスパートジャッジけれども、ほかのところにもいっぱい出てきてはいるのですが、やはり考え方の幾つかの道は書いておいて、その中で最適なものを選んでもらうというようなやり方がいいのかなと。でないと、誰か強い意見を持った人が決めてしまって、右へ倣えで終わってしまったということでもまずいでしょうし、やはり観点は書いておいたほうがいいのかということなのですが、その際に、ヒトの場合だと社会的影響というのは一つの観点だと思うのですが、しかし、もう一つ今回付け加えたのが、統計学的観点と言っていいかどうか分からないのですけれども、精度の問題なので、この辺を書いていいかどうかということです。従来入れていなかったのが、今回のワーキングから入れているので、こういうことを書いて構わないのかどうかと。あるいは、こういうことは書かないほうがいいのか、そういうことで何か方向性はありますでしょうか。

座長があまりしゃべってはいけないのですけれども、ステップ1、2というのは確かに実施のしやすさで言えばとても実施しやすいのですけれども、何が優先されるかもまだ決め切れないので、候補となるものの考え方を列記して、これらの中からこういう点を判断基準にエキスパートが決めていくというまとめ方になるかと思うのですが、その際に用量反応曲線の持つ精度というものを入れているかどうかということなのだと思いますけれども、そこが一つのポイントかと思っています。

赤堀先生、手を挙げていただいていますね。お願いします。

○赤堀専門委員 門外漢なので的外れかもしれないのですけれども、EHC240の文書には、

どういう観点でエキスパートジャッジをするのかということが書かれているので、その要素を取り込んで文章をリバイスするという事は可能と思いました。既にきちんと有害影響について専門的に考えると、どこからが有害とみなせるとか、重篤さ等も加味してそこでは判断するという事とかが書いてあるので、その部分はそれでよいと思いました。

川村座長がご質問されていたモデルのほうからやっていいかというところに対しては、個人的には難しいところを含んでいるという印象を持ちました。適用するモデルそのものが十分に信頼に足るというところをどこかでまた評価しなければいけないというところで、個人的に懸念を感じました。

そういう意味では、アメリカが取っているような形のアプローチ、バックグラウンドのSDを例えば基準としてやるというほうが、実際に評価をするときにはやりやすいであったり、説明のしやすさという観点でも少し合理性があるのかなという印象を持ちました。

以上です。

○川村座長 ありがとうございます。

検査値のようなものであればSDを使うということは結構行われていますし、それは臨床の世界でもミニマルクリティカルインポートアントディファレンスが出せないものがたくさんあって、そのときには検査値であればSDを基準にして評価するという事はしばしば行われています。ですけれども、それが本当にいいかどうかというのはやはり疑問で、つまり、先送りですが、それを暫定的には一つのマーカーとするのだけれども、マーカーというか一つの基準にするのだけれども、それでいいかどうかはもう少し長期に検証するという事が必要になるかと思えます。

赤堀先生、もう一度お願いします。

○赤堀専門委員 一つ言い忘れていたのが、必ずしもベンチマークドーズでPODを決められないケースがあるという可能性はないのかなと思ったのです。川村先生が懸念されているように、かなり低用量のところでは振れ幅が大き過ぎて、それでも無理やりにでも決定しなければならないものなのかどうかというところなんです。こういう場合には、例えば決定できないかということはある程度得たりするのではないかと。BMDが適用できない場合があるということに関しては、この文書の中で一切議論しない点が気になりました。的外れなのかなと思うのですけれども。

以上です。

○川村座長 PODは決めても、実際に社会の中でどう適用するかは多分その後考えることになると思います。例えばMOE、Margin of Exposureみたいな考え方というのはストレートではなくて、当座の逃げみたいな、逃げと言ったら言い方は悪いですが、そういう表現の仕方もあるので、実際に国として規制を行う際の指標としては、ストレートに真正面から向えない場合は代替りの方法で表示するという事も十分あるかとは思えます。

WHOとかアメリカとかの記述をもう一度吟味しないとはいけないと思うのですが、できれば、優先順位がつけられるところまでは難しいかもしれないけれども、考慮すべき要素

についてと、そもそもベンチマークレスポンスというのはこういう考え方のものですかということ、言葉の定義のところと、それから、運用に当たって、BMRの決定ということにおいて、こういう観点で専門家により決定するというようなところまで書けたらと思っておりますので、そこに書き込む内容についてはもう一度精査させていただいて、皆さんにまたお示ししたいと思えます。

西浦先生、お願いします。

○西浦専門委員 このBMRの話は今の話が結構重要で、このガイドでどう書くのかは慎重にやっておいたほうが良いというのが私のお勧めです。どういうことかということ、今、リンクとかを送りますけれども、EPAとかが一つ一つの化学物質とかほかの物質のリスク評価をするときに、パブリックコメントを必ず取りますよね。そういうときにいつも論点になっていることのひとつが、このBMRはないのではないかとというような話が必ず届いています。

その激しいものの事例が、今お届けしたような甲状腺ホルモンに関連する毒性評価のコメントなのですが、元のものだったら、BMRというのを相対的なインクリースで20%とかで定義していたのだけれども、いろいろやっていたらそれではリスク評価はならないと。やさし過ぎるので、EPAのスタンダードというのが途中で0.5SDとかになったのです。それが、その後さらに1SDもいよいよかという話になって、混乱を極めている。化学物質によって経験を繰り返すごとに基準が異なるというようなことがあって、こういう話になっていくと思われるので、これは放置すると混迷を極めるリスクがあると思えます。

なので、ここでお勧めとしては、先ほど広瀬先生がWHOのEHCの改訂版の話をしてくださいましたけれども、その和訳が付表的にこういうガイダンスの後に参考資料として入っているだけで、全く違うことになると思えます。一種のフローチャートに近いようなぐらいい分かりやすくEHCでは書かれているわけなのですが、BMRを決めるときは、平たく言うと、これはいつもゴールポストが動くようにできているということにほかならないわけなのですが、そのときに最も合理的にゴールポストを決めている基本的考え方ということでアグリーメントが取れる部分がWHOのEHCにまとめられているので、それがあるだけで大いに助けになるのではないかなとは思いました。

以上です。

○川村座長 ありがとうございます。

先生は、国際的に流通しているとか認知されているものの記述を尊重して、どういう観点とか、そのイベントとか健康影響のものにより、がんであるとか、そのほかクリティカルな病気であるのと、それから、検査値のように間接的であってサロゲートマーカーでのものとは当然書き方も基準が変わってくると思えますし、それから、クリティカルかどうかという一人一人の重症度のほかに何人が発症するかというインシデンスのサイズの問題も影響するので、そういった要素でどこまで許容されるか。例えばBSE、前にあった狂牛病みたいに期待値が0.5ぐらいの小さいもののお話と、肺がんみたいに毎年10万人

とかというサイズのものとは、単なるパーセントでは議論できない違いがあるので、その辺のところもそういう要素について書いて、国際的な資料の事実と矛盾したりしないように、できればある程度歩調をそろえるということも必要かとは思っております。

ということで、今ここで成文を得るといのはちょっと難しそうなので、このところはもう一度考えさせていただけますでしょうか。

用語の説明のところに複数の考え方があることを説明するのはよろしいですか。

では、祖父江先生、そのようにさせていただいて、最初の用語の定義のところでは説明して、その運用についてはこの中に記述する。どういう観点でそれを決めるかということは、あるいは、手順はどうするかということは本文の中で書くということにしたいと思います。

では、このところはそういうことで宿題になっていますが、そのほか、この後のところで何か気になる点はございますでしょうか。あとはよろしいですか。

これまた議論を吹っかけるようで申し訳ないけれども、文章を修正するという意味ではなくて、念のための確認をするという意味ですが、17ページの15行目からPODの決定というところがあるかと思えます。ここで、記述はこれでよろしいのですけれども、ものの考え方として、ここにはBMDとBMDLを併記しています。書き方としてはこれでいいのですけれども、考え方としては、一部挿入してあるように、BMDとその信頼下限値であるBMDLというこのせりふが非常に大事であって、そもそもなのですが、いわゆる各データの代表値とばらつきの指標みたいに、平均が幾つで標準偏差が幾つというのはそれぞれ対等ですよ。平均値と標準偏差は対等で、どちらも独立して使えるものなのですが、信頼区間というのはp値の裏返しで、ある推定についてまず最尤値を出して、その最尤値のぶれを示す。どのくらいこれがばらけないかということを示すものなので、信頼区間というのは付随したものである。付随しているものなので、ここは今90%になっていますけれども、95%でも99%でも別に構わない。何%で表記しても全然問題はないもので、ただ、便宜的に95%がよく使われる。ここでは90%を使っているというだけの話なので、便利さだけで選んでいるだけですから、90%に深い意味はあまりないわけですよ。

だから、BMDとBMDLは推定値なので、対等ではなくて、BMDの確かさを示すものをBMDLとBMDHで表現しているという意味合いなので、そのところは、文章を直すということではありませんが、そのところの認識を、生データの平均値と標準偏差、あるいは中央値と四分位範囲というものとは性質が違うということで、記述が似ているので混乱を来すといけないので、念のためのコメントなのです。どうしてそういうことを言うかということ、これから実際に運用していく場合に、初めにBMDLありきではないと。BMDを決めて、それに付随してBMDLを書いて、実質的にはBMDLを採用するという手順なので、BMDLを採用するのはあくまでも措置としてであって、サイエンスで求めるのはBMDであって、もちろんBMDLもサイエンスで出てくるんだけど、採択するのは研究が持つ不確実性を補うためにLを採用するという措置であって、付随した指標であるということを確認しておきたく、

これは私が昔から言っていることなのですからけれども、あえてその辺りは念押しとして申し上げたいと思っています。

記述自体はこれで構わないので、実質的にはBMDLを採択するのですけれども、どうしてこれが問題になるかということ、動物実験でもそうですが、いろいろなエンドポイントがあって一番低いものを選ぶというときに、BMDLで一番低いものを選ぶと、当然信頼区間が広い、研究としては質の低いものを選んでしまうリスクがあるから、あくまでもBMDLは付随的指標なので、実際には採用するのだけれども、BMDが基幹となる指標ですと。運用上安全サイドに立つということでBMDLを取りますよという考え方で物事を運用するときには進めていただきたいと思いますので、あえて申し上げた次第です。

これはプラスアルファのコメントなのですが、これにつきまして、あるいはそれ以外のことににつきまして、何か御意見とかはございますでしょうか。よろしいですか。

今の後半の部分の記述について、特に御意見、御異議とかはございませんでしょうか。

松本先生、よろしいですか。（松本専門委員首肯）ありがとうございます。

澤田先生、よろしいですか。（澤田専門委員首肯）ありがとうございます。

では、全体を通して、一番疫学研究で特に要となるBMRについては、宿題というか、もう一度精査する必要があるかとは思いますが、そこは大事なところなので、もう一度皆さんに何らかの形でお示します。この部分は座長一任という形は取りませんので、皆さんにお諮りして総意としてまとめたいと思います。

必要に応じて、またワーキング、この会議を開く、あるいは御意見がすんなりといくようであれば会議は開かない可能性もありますけれども、一応ワーキングのメンバー全員にきちんとお示しして、納得がいく形で示したいと思っております。

そういうことですが、全体を通して何か問題点とか御提案とかはございますでしょうか。

よろしいですか。皆さん御発言いただいていますか。

猪飼先生、全体を通してコメントはありますか。

○猪飼専門委員 猪飼です。

BMRの話でこんなに大きくなると想像してなかったもので、初めての議論でいろいろ勉強になりました。ただ、BMRの議論で、例えばIQが幾らであるべきだという鉛の例とかを昔見せていただいたこともありましたが、そこがあまり大きな問題になるようであれば、そもそもベンチマークを使うべきなのか、NOAELなのか、そこがどちらでもいいぐらい話がぐらぐらしているのではないかという気もするのですよね。BMRを使うと決めたときに、諸外国が例えば10%をデフォルトとして使ってはくれるなど、例えばEFSAだったかが書いていたような気がしますが、ある程度の基準を決めて、バックグラウンド若しくはもっと強い臨床的なカットオフがあるならそこを使ってもいいけれども、その後の処理は淡々と進めるのだと。でも、そのためのモデル平均化の方法もいろいろ議論されているので、ある種の前提を置いて、その後をかつちりとした方法で進めるというのが今のここでのBMDの指針をつくっている目標ではないかなと思うのです。

BMRの議論はすごくぐらぐらするので、その値が変わればすごく結果には影響が出るのだけれども、その話は外でやって、WHOでも悩んでいるとおりに、原則は示しながらも、ここではあまり言及しないというぐらいの話のほうが、この指針を取りまとめるという目的が、そちらのゴールポストがどんどん動いてしまうことを心配しながら伺っていました。

ですので、結果的には頭の定義のところでは諸説ありと書いていただきながら、後ろのところでは、14ページのところでしたか、BMRを決めたから即HBGVとはならないとか、この辺の議論を省けば、このままでもいいのではないかなと思ったのが私の意見です。

あまり長くならないほうがいいなと思ったというぐらいの意見で、以上です。ありがとうございました。

○川村座長 ありがとうございます。

そういうBMRという一つの統計学的な指標なのですけれども、ここにいろいろな思いがこもるし、パブリックコメントを出したときにここに批判が集中すると思うので、説明責任といいますか、きちんと説明できる、多面的に考えていくと、それこそ猪飼先生が得意とする社会的な要素ですね。社会的というか、国民レベルの目線とかそういうのも必要になってくるので、このところはこれから実際に運用に入るときに多面的な検討をして議論を尽くして説明していくということが重要だなと思います。

ですので、ここでは書き切れないし、物質や健康影響の大きさによって変わってくると思うので、実際には、指針はつくるんだけれども、やってみて、やりながら相当詰めていかないといけないところが大きくなると思うのです。そういう意味で、基準というか、骨格の部分は今回の指針で示すのですけれども、もちろん実際の評価の段階ではここに書き切れていないいろいろなことを議論して、説明できる形できちんと収束させていくということになるので、またそちらのほうで皆さん活躍いただきたいなと思っています。

ありがとうございました。

ほかに御意見はよろしいでしょうか。

親委員会の委員の先生方、よろしいですか。コメントはございますでしょうか。

○川西委員 どうも御議論ありがとうございます。委員の川西でございます。

今日、委員が先週のシンポジウムで頑張り過ぎて体調を崩していて、大変失礼して申し訳ありません。代表して謝らせていただきます。

実はこのベンチマークドーズは、この前、事務局の寺谷さんから説明があったように、今、研究事業の二次募集でベンチマークドーズ法のリスク評価への活用における課題を明らかにする研究というのを公募して、締め切っている状態です。それで、今の議論を聞かせていただいて、BMRの問題はベンチマークドーズ法の活用を図る上でとても大切な問題ですし、継続的に、もう一回今まで実際に使われている例に立ち戻って眺めたり、今回の指針そのものは取りあえず仕上げるということが必要であっても、今後これを活用する上で、特に疫学研究などの場合にとっても重要な視点だなとしみじみ思いました。

そんなこともあって、委員のほうからもそういう視点も組み入れて、研究事業の場も使

いつつ、継続的に議論を続けて、より全体が頭の整理をしていくということではできのかなと思ひながら聞かせていただきました。

私の感想なのですけれども、そういう機会もあるのだということでコメントさせていただきました。ありがとうございます。

○川村座長 ありがとうございます。

確かに勇気づけられたと言うと変ですけれども、決め切れない、これは長くかかる課題であるということで、ここで明言し切ってしまうのではなくて、考え方の基本は少し書き込みますけれども、そういう評価技術の研究であるとか、あるいは実際の各専門調査会における検討であるとかを通してブラッシュアップしていくものであると思います。だから、EFSAもEPAもこうだと言では書き切れていないと思うので、そのところは3局合同の検討課題にもなるかと思ひますし、我々も積極的な提言ができればと思ひています。

一気にここで完成版をつくるという考えではなくて、将来に課題を残しながら、一応仮の締めということで仕上げさせていただくということにさせていただければと思ひております。ありがとうございました。

それでは、川西委員にまとめをしていただいたので、ちょっとほっとして進めたいと思ひます。

ほかに御意見はよろしいですね。「ね」と言って誘導してはいけないので、よろしいですか。（専門委員首肯）

では、よろしいようでしたら、本日の審議はここまでとさせていただきます。

そのほか、事務局から事務連絡等をお願いいたします。

○寺谷評価調整官 事務連絡というよりは、少しコメントをお伝えしたいと思ひます。

本日もこのような非常に活発な御審議をいただきまして、深く感謝申し上げます。

今、川西委員からもコメントがありましたし、座長にもまとめていただいたように、BMD指針の御議論については一区切りついてきたのかなと考えております。2019年にできました動物試験への適用という文書に加えて、この間にベイズ推定のことをオンしまして、それから、今回疫学研究の適用というものが加わりまして、一区切りついてきたかなと考えているところです。

このような動きは食品安全行政にとってとても大きな動きですし、さらに大きな視野としても、このようにリスクを評価したり管理するという科学技術に密着した行政にとってもとても大きな一歩になっているのだらうと認識しているところでございます。

御議論いただきました専門委員の皆様、それから、参考人、座長の川村座長にも深く深く御礼を申し上げます。誠にありがとうございます。

この後の作業としましては、まず座長一任とはなっていませんが、中身をさらに詰めまして、その後にパブリックコメントなどをかけていく作業になっていきます。引き続き御協力いただければと思ひます。どうぞよろしくをお願いいたします。

また、川西先生からもおっしゃっていただいたように、研究事業なども活用しながら、

実践しながらまたこの文章を磨いていくような作業を今後やっていったり、それから、BMDの普及啓発、人材育成、その辺についても、我々事務局としてもしっかり捉えていきたいと思ひます。このような場面においても、まさにここにいる先生方に、このワーキンググループという枠組みなのか、そうでないとしても、いろいろな形で御協力いただく場面が多々あると思ひますので、そのときもぜひ力添えをいただければと思ひます。

改めまして、このように私たち食品安全委員会はまさに科学に基づいた行政をやっているのですが、各専門調査会の中でも我々はまさに科学の中でもコアとなる科学技術というものを議論いただいているものだと認識しています。このように深く真摯な議論をいただいたことについて、深く感謝申し上げます。誠にありがとうございました。今後ともどうぞよろしくお願ひいたします。

以上、コメントでございます。

○岡谷評価専門職

今回のワーキングの日程につきましては、決まり次第先生方にお知らせいたしますので、よろしくお願ひいたします。

この後、事務連絡がございますので、ウェブの先生方は接続を切らずにもう少しお残りいただきますようお願ひいたします。

以上です。

○川村座長 ありがとうございます。

これで本日の議事は終了いたしました。

通算すると相当長い期間委員の先生方に御尽力いただきまして、かなり思考を深めることができました。成果物としては量的に分厚いものではありませんけれども、考え方というものを突き詰めるということで、評価技術企画ワーキングとしては使命を果たしつつあると認識しております。

そういう意味で、皆さんの御尽力に感謝しつつ、今回のワーキンググループは閉会とさせていただきます。どうもありがとうございました。