

〔仮訳〕

2006年5月

米国における牛海綿状脳症（BSE）有病率の推定に関するピアレビュー

最終報告書

提出先：

米国農務省（U.S. Department of Agriculture）

動植物検疫局（Animal and Plant Health Inspection Service）

メリーランド州リバーデール（Riverdale, DM）

契約公務員の技術代表者：

**チュアンファ・グォ（Chuanfa Guo）**

米国農務省（U.S. Department of Agriculture）

米国食品安全検査局（Food Safety and Inspection Service）

ワシントン D.C.（Washington, DC）

作成者：

**サミート・パティル（Sumeet Patil）**

RTI インターナショナル保健・社会・経済調査部

（RTI International

Health, Social, and Economics Research）

ノースカロライナ州リサーチトライアングルパーク（Research Triangle Park, NC 27709）

RTI プロジェクト番号 0208893.023

RTI プロジェクト番号

0208893.023

## 米国における牛海綿状脳症（BSE）有病率の推定に関するピアレビュー

### 最終報告書

2006年5月

提出先：

米国農務省

動植物検疫局

メリーランド州リバーデール

契約公務員の技術代表者：

**チュアンファ・グォ**

米国農務省

米国食品安全検査局

ワシントン D.C.

作成者：

**サミート・パティル**

RTI インターナショナル保健・社会・経済調査部

ノースカロライナ州リサーチトライアングルパーク

\* RTI インターナショナル (RTI International) は、リサーチトライアングル研究所 (Research Triangle Institute) の登録商標である。

## 目次

セクション	ページ
要約	ES-1
1 背景および目的	1-1
2 米国における BSE 有病率の推定に用いた方法	2-1
3 レビュー過程の解説	3-1
4 ピアレビューアに対する課題	4-1
5 参考文献	5-1
6 ピアレビュー報告書	6-1

## 要約

OMB (Office of Management and Budget : 行政管理予算局) では、連邦政府機関が草案を公表する前に、科学技術研究の品質を保証し草案に改善を加えるため、重要な科学的情報についてのピアレビューを行うよう要求している (OMB、2004)。APHIS (動植物検疫局) は、BSE (Bovine Spongiform Encephalopathy : 牛海綿状脳症) サーベイランスのデータに基づく BSE 有病率の推定手法に関するピアレビューの実施に関与している。これらのデータおよび分析結果はサーベイランス計画の特性を表しうるとともに、BSE リスク低減に関する政策・意思決定にも情報を提供するものである。APHIS は OMB のガイドライン (OMB、2002 ; 2004) に従ってピアレビューを実施するべく、FSIS (Food Safety and Inspection Service : 米国食品安全検査局) が RTI インターナショナル (RTI) と結んだ業務指示契約に基づいて RTI に支援を要請した。

APHIS は特に、2 種類の有病率推定法 / モデル (BSurvE モデルおよびベイズ解析) に関するレビュー、および分析のアプローチ、データ、仮定、得られた結論を解説した報告を必要としている。RTI は 3 名の専門家を指名し、作業内容書に従ってピアレビューを実施した。本報告書ではこの 3 件のレビューを提示する。

レビューアは全員、米国における BSE の推定有病率が牛 100 万頭あたり 1 頭であるということで見解が一致している。各レビューアはモデルを改善すべく意見を述べたものの、3 名とも、APHIS が実施した詳細な感度分析において提示された変更点によって推定有病率が変動することはないであろうことを認めている。レビューアのうち 2 名は、結果の信頼性を高めるため、他のパラメータおよび仮定に関する感度分析を実施するよう提言している。推定有病率が変動しにくいことを考慮して、レビューアはモデルの他の側面 (別の入力仮定の利用、サーベイランス標本における制限事項の確認または考察、他のパラメータの感度分析の提言、報告書中での適切な技術用語および定義の使用、報告書中の特定箇所の明確化など) に関心を集中させた。

モデルの適合性、透明性、頑健性という観点から見て、全レビューアとも、BSurvE モデルおよび Bayesian 出生コホート (BBC) モデルが統計的にも疫学的にも適切であることを見いだした。例えば、レビューアのうち 2 名は、BSurvE モデルが欧州食品安全機関のレビューにおいて是認されており、BSE の国内有病率の推定に適切なモデルと判定されていることを明確に記している。全レビューアとも、BSurvE モデルの適用によって十分に再現性の良い結果を得ることができた。1997 年の飼料規制に関する情報を用いる BBC モデルも、透明性および頑健性を備えた適切なモデルであることが見いだされた。しかしレビューアのうち 2 名は、BBC の結果が BSurvE モデルの結果に由来するとの事実を用いて、BBC の結果は BSurvE モデルと同様に重要であり、また、BBC の結果が BSurvE の結果と相関するの

は驚くことではない、という 2 つの関連した推論を下している。全レビューアとも、WinBUGS プログラムを用いて BBC モデルを適用することができたが、うち 1 名は報告書中のコードについてより詳しい説明の必要があるとした。

モデルの完全性という観点において全レビューアは、APHIS が重要な要因を適切に考慮しており、モデルのパラメータは適切と考えられるという点で意見が一致している。全レビューアとも、特に仮定に関してはおおむね意見が一致している。なぜなら、これらの仮定は感度分析によって後に評価されており、結果にほとんど（またはまったく）影響しないことが判明しているためである。しかしながら、一部の入力仮定に関してはより詳細な考察が必要であると述べた。

レビューアのうち 2 名は、BBC モデルにおいて、米国における発生数の相対的減少率を得るために英国のデータを利用することの精度について論じたが、現在の APHIS の仮定に反するような強い勧告は行わなかった。一方のレビューアは、米国における発生減少率は英国における流行の後期段階で見られる発生減少傾向に近いと主張した。もう 1 人のレビューアは、年変動補正因子を英国のデータから導かれた事前分布を有するモデルパラメータであるとする、BBC モデルの敷衍を示唆した。

分析に関しておそらくもっとも意味のあるコメントは、他のパラメータが結果にもたらず影響を評価するため感度分析を実施するようとのコメントであろう。例えばレビューアのうち 2 名は、サーベイランス期間の前期と後期では標本抽出方法が大きく異なるため、標本の年齢分布の影響を評価するよう強く提言した。レビューアはさらに、標本集合の代表性、感染動物の退出確率、潜伏期間の分布といった要因の影響を評価する必要性を指摘した。これらのパラメータが結果の頑健性に及ぼす影響は無視できる程度と考えられる。それでもなお、結果の信頼性をさらに増し完全を期するために、こうした不確定性の分析を行うことが勧められる。

## 1 背景および目的

RTI インターナショナル (RTI) は、USDA (U.S. Department of Agriculture : 米国農務省) の FSIS (Food Safety and Inspection Service : 米国食品安全検査局) および APHIS (Animal and Plant Health Inspection Service : 米国動植物保健検査局) の要請により、米国における BSE (Bovine Spongiform Encephalopathy : 牛海綿状脳症) 有病率の推定に関する外部ピアレビューのコーディネートを実施した (業務指示番号 0208893.023)。本報告書では、ピアレビューの背景的情報の提示、レビュー過程の解説、重要な疑問点すなわちレビューアに対する課題の列挙を行った上で、3 件のピアレビュー報告書を掲載する。

APHIS は 1990 年以來、米国の牛群における BSE サーベイランスを鋭意実施してきた。2004 年 6 月初旬、これらの取り組みはさらに強化され、12~18 ヶ月間にわたって対象牛群からできるだけ多くの標本を得るという目標がおかれた。これらのサーベイランスから得られたデータを分析し、米国における BSE 有病率の推定が行われた。この分析の結論は複数の目的に役立つであろう。USDA は強化サーベイランスで得られたデータを用いて、維持サーベイランス計画を立てるであろう。有病率を理解することは、BSE リスク低減策に関して科学に基づく政策決定・規則決定に寄与するであろう。また、有病率の推定値は、海外での市場を取り戻すための貿易に関する議論および取り組みの一助ともなるであろう。したがって、この分析および報告書は科学的に重要であり、OMB (Office of Management and Budget : 行政管理予算局) のガイドライン (2004) に従って外部ピアレビューを行うに値する。

ピアレビューは科学的情報の品質が技術界の基準に合致していることを保証する助けとなりうる重要な過程であり、分析を強化および明確化する助けとなりうる。APHIS は、BSE 有病率推定手法に関する正式かつ独立したピアレビューを、ピアレビューおよび情報の品質に関する OMB のガイドライン (OMB、2002、2004) に従って実施するための支援を RTI に要請した。

## 2 米国における BSE 有病率の推定に用いた方法

分析では、7 年間にわたる BSE サーベイランスの取り組みから得られたデータを用いて、米国における BSE 有病率の推定を行った。有病率は、時間軸のある一点における牛群に存在する疾患の量に関する推定値を与える。有病率は、成牛群（感染レベルが検出限界以下の成牛群を含む）における感染牛の総数または比率を表している。牛群の差異および牛群統計値を考慮した BSE 有病率の推定方法が開発されてきた。今回の分析では次節に述べる 2 種類の方法を用いて有病率の推定を行った。さらに、使用したモデルに有意な影響を及ぼす可能性があるいくつかのパラメータに関しては不確定性分析を実施した。

BSE 有病率の推定に用いた方法の 1 つは BSurvE モデルであった。BSurvE モデルは最近開発された手法であり、牛群統計値、牛群からの退出率、疾患に関する既知の特性を利用するよう設計された。このモデルは BSE 発見の尤度を直接的に推定するものであり、各サーベイランス標本に対し、無作為標本から得られる情報と比較した値を割り当てる。

一方、Bayesian 出生コホート (BBC) 法は、疾患対策（飼料規制など）の有無または有効性に関する情報を現存牛群における有病率の推定値に反映させるために用いられた。飼料規制が有効であるとの知識を考慮すれば、未来のある時点における有病率に関して、十分な情報を得た上での結論に達することだろう。ベイズ分析法は、事前知識をサーベイランスの情報とともに数学的に織り込むことで、利用できる情報すべてを反映した最終的な推定値を与えることができる。

推定値の結果に影響する可能性がある 5 つの要因について、不確定性分析を実施した。最終的な結論は、有病率推定に用いた両手法の結果および不確定性分析の結果に基づいている。

### 3 レビュー過程の解説

RTI は OMB のガイドライン (OMB、2004) に従ってレビューの過程を実施した。レビューの過程は、レビューアの選定、レビューの意図の説明、レビューの支援、各レビューを 1 冊の報告書にまとめる作業で構成された。

第 1 段階として、専門知識に基づいて 3 名のレビューアを選定した。まず FSIS および APHIS からピアレビューの背景および目的を把握してから、14 名のレビューア候補者を特定した。その後、候補者の能力、および BSE の科学における専門分野 (特に動物衛生、数学的モデルおよびベイズ理論、有病率の推定方法に関連する専門分野) の重なり具合に基づき、最終的に 3 名のレビューアを選定した。選定過程では利害の衝突についても考慮した。

第 2 段階では、APHIS が立案したレビューアに対する課題という観点からレビューの意図を説明した。RTI は課題とともに、報告書、BSurvE モデル (Excel 形式) BBC モデル実行のための WinBUGS プログラムを提供した。課題は 5 セットの質問で構成されており、詳細は本報告書の第 4 節で述べる。

第 3 段階で RTI はレビューアと意見交換し、レビューアがレビューの目的または分析自体について抱いたすべての疑問を明確化した。RTI はまた、APHIS および FSIS に対して定期的にレビューの進行状況を伝え、また、レビューアがピアレビューの目的を確実に果たすようにした。さらに、レビューアが各問題点だけを挙げるのではなく、その問題点についてとりうる対策についても記述するよう確認した。

最終的に、3 編のレビューを 1 冊の報告書にとりまとめた。本報告書では、ピアレビュー対象である分析に関する背景情報の概要、ピアレビューの過程、ピアレビュー報告書を提示する。オリジナルのピアレビューを第 6 節に添付する。

レビューの整合性を維持するため、課題質問ごとにコメントをまとめて掲載するのではなく、各レビューを別々の章として示している。課題質問に関して各レビューアが注目した側面は、レビューアの専門分野によって異なっていた。また、報告書の書式および記述スタイルも各自異なっている。そのため、各レビューを別々に読む方がコメントを理解しやすい。RTI では軽微なタイプミスを修正し、本報告書の記述に最小限の統一性をもたせるために報告書の書式を若干変更した。



#### 4 ピアレビューアに対する課題

APHIS は各レビューアに対し、以下に示す個別の質問事項に注目して記述を構成するよう依頼した<sup>1</sup>。

1. 報告書の目的が有病率の決定であることを考慮した上で、この分析における当局の結論を検討すること。当局は適切なデータを利用していたか？ また、データは妥当であったか？ 分析方法の適合性、アプローチの透明性、結果の頑健性に的を絞ること。

2. 文書中で述べられている分析の目的は以下のとおりである。

米国は 1990 年以来、BSE サーベイランスを鋭意実施してきており、2003 年にカナダの牛 1 頭が陽性と確認されてからはさらに取り組みが強化されてきた。米国における強化 BSE サーベイランス計画の概要は別の報告書( USDA、APHIS、CEAH( Centers for Epidemiology and Animal Health : 疫学・家畜衛生センター) 『BSE 強化サーベイランスの概要』( Summary of Enhanced BSE Surveillance ) 4-6-06 ) にまとめられている。本分析では、2006 年 3 月 17 日以前の 7 年間にわたって収集されたサーベイランスのデータを用いて、米国における BSE の有病率を推定している。

OIE ( World Organization for Animal Health : 国際獣疫事務局 ) が採択している国際ガイドラインに準拠した米国の BSE 状況全般の考察が今後要求される際、この情報は手引きおよび支援となるであろう。

こういった分析目的との関連において、モデルは完成しているか？ また、理にかなっているか？ 現実社会の状況において有病率の推定に影響する重要な要因やこうした要因間の相互作用をモデルは踏まえているか？ 踏まえていないとすれば、抜けている(または正しく特定されていない)パラメータはどのようなものか？ 研究する価値があったであろうモデルは他に存在するか？ モデルが現実から乖離する場合は明確に定義されているか？

3. 入力パラメータに関する仮定事項を検討すること。分析目的との関連において仮定は合理的であるか？ 仮定にはもっとも蓋然性の高い合理的な推定パラメータ値が反映されているか？

注：使用した WinBUGS ソフトウェアの一部はモンテカルロモデリングであり、分布にはポアソン分布を選定して用いた。このことは WINBUGS のコード(文書に含まれている)において明示されている。

---

<sup>1</sup> レビューアが述べた全体的なコメントの中には、6 項目の個別質問に当てはめうるものもあった。しかし RTI の要請により、レビューアは全体的なコメントを個別の質問事項から分けたままとした。

4. モデルおよび入力データの不確実性に対して結論は頑健性を示すか？不確実な事項は完全に記述されているか？ 結論の頑健性はどの程度か？ これらの関係は明解に記述されているか？

5. 用いたモデルの下で、分析結果は正しいか？ プログラムを実行してコードおよび数式を点検し、モデルが意図どおりに機能していることを確認すること。用いたモデルの下で、当局が正しい答えを公表したのかどうかを判定すること。

5 参考文献

## 6 ピアレビュー報告書

『米国における BSE 有病率の推定』のレビュー

ポール・S・モーリー (Paul S. Morley) DVM (獣医学博士) PhD、DACVIM (米国獣医内科学会専門医)

コロラド州立大学獣医学センター 疫学・バイオセキュリティ担当准教授、バイオセキュリティ管理者

コロラド州立大学動物集団衛生研究所 アソシエートディレクター

モーリー博士は CSU (Colorado State University : コロラド州立大学) において、臨床科学部門の准教授であるとともに、獣医学・生物医科学部に属する動物集団衛生研究所のアソシエートディレクターである。また、同大学のジェームズ・L・ボス記念獣医学教育病院におけるバイオセキュリティ管理者でもある。加えて、2005 年の研究休暇中には、カナダ・アルバータ州オコトクスにある飼養場保健管理サービスにおいて飼養場および研究に関する顧問として勤務した。博士は CSU の獣医学部プログラムおよび大学院プログラムにおいて疫学および感染管理を教えるとともに、活発な研究活動を維持している。専門および研究活動に含まれる主なテーマには、家畜感染症に関する分析疫学、農業および獣医学において重要なバイオセキュリティ関連事項、牛肉生産に関するデータの分析、動物における抗菌剤耐性の生態学に関する研究、食品の安全に有効な収穫前アプローチの探究などがある。博士は獣医疫学予防医学会の次期会長であり、また、獣医内科学雑誌 (Journal of Veterinary Internal Medicine) のアソシエートエディターとして、特に大動物の内科学を中心に務めを果たしている。また、米国獣医内科学会、獣医疫学予防医学会、獣医感染管理学会、獣医コンサルタント協会といった専門家団体の活発なメンバーでもある。モーリー博士はオハイオ州立大学で 3 年間教員を務めた後、1998 年にコロラド州立大学の教員となった。生まれはネバダ州で、ワシントン州立大学で学士号および獣医学の学位を取得した。ワシントン州立大学卒業後、カナダ・サスカチュワン州サスカトゥーンにあるサスカチュワン大学獣医学部 (Western College of Veterinary Medicine) で大動物内科分野のインターンおよびレジデントとして勤務した。ウマの呼吸器感染症の疫学に関する研究で、同大学から博士号を授与された。博士は米国獣医内科学会認定の専門医である。モーリー博士は動物集団における疫学および感染管理の権威であり、北米の複数の大学獣医学部や集約的動物生産施設で感染管理およびバイオセキュリティ関連事項に関する顧問を務めている。査読を受けた論文の発表数は 75 以上で、その他に共著が数編ある。

**緒言 :** RTI から送られてきた指示書、および USDA・APHIS が立案したレビューアに対する課題において要請されているとおり、この文書では、『米国における BSE 有病率の推定』と題された報告書 (2006 年 4 月 27 日付、以下「本報告書」とする) ならびに報告書の作成

に用いられた 2 種類の数学的モデルに関する全体的コメントおよび個別コメントを述べている。この本来のレビュー対象文書その他、USDA のウェブサイトで見いだした裏付けとなる文書（2006 年 4 月 27 日付の『米国における BSE 強化サーベイランスの概要』など）についてもレビューを行った。また、USDA の BSE サーベイランス結果および BSurvE 推計モデルを含むスプレッドシート（BSurvE\_0603\_USData.4.26.06.xls） および報告書の付録 B に示されたコードを用いて得られる WinBUGS モデルの出力結果についてもレビューを行った。さらに、BSurvE モデルのスプレッドシートとともに提供された文書、すなわち EFSA（European Food Safety Authority：欧州食品安全機関）による BSE サーベイランスモデル（BSurvE）の評価に関する科学的報告書（2004 年、17 号、1-6）ならびに OIE の陸生動物衛生規約のうち BSE およびサーベイランスに関連する情報（特に第 2.3.13 条「牛海綿状脳症」、附則第 3.8.4 条「牛海綿状脳症に関するサーベイランス」、附則第 3.8.1 条「動物衛生サーベイランスに関する一般的ガイドライン」）についてもレビューを行った。

**全体的コメント：**ほとんどの部分で、本報告書ならびにその裏付けとなる文書類およびモデルの記述は、USDA の結論の導出に用いられた仮定および方法に関して明確かつ透明性の高いものであった。科学的な「主張」および結論の真価を十分に評価するための説明に際して読者が要求するであろう事項のほとんどを、USDA は考慮および対処してきたと思われる。しかしながら、このレビューで考察および検討するに値する問題点もいくつか存在した。そうした問題点を以下に示す。

1. サーベイランスの標本は参照集団を真に代表するものであったか？
2. 上記の疑問は、この問題に関する OIE のガイドラインが真の代表標本を要求しているかどうか、集団を代表する確率標本抽出が行われなかった場合でも適切な結論に達しうるほどに推奨サンプリング方法が慎重なものであるかどうか、という見地から解釈しなければならない。
3. 他のデータ源から USDA のデータに当てはめる際の推定は、次の事項に関して適切であっただろうか。
  - a. 1999～2003 年の標本の一部で年齢が記録されていなかったが、動物の年齢分布の仮定は適切であったか？
  - b. 飼料規制の結果として期待される BSE 発生の減少率は適切であったか？
4. 標本について記録されている年齢データの精度は、モデル中で示されているとおり本当に 1 年以内であったか？ 本報告書の不確定性に関する節でこの仮定の評価が行われなかったのはなぜだろうか？
5. BSurvE モデルにおいて「擬似症状」サーベイランス区分に動物を割り付ける際に用いる、臨床症状に関する敏感度および特異度の推定値は適切か？
6. 上記の各疑問に関して、記載の研究方法を用いた中で偶然に発生したかもしれない誤り

は、正しくない結論に達するような偏りをモデルにもたらしたでしょうか？

**個別コメント：**

**USDA の課題 1)** 「報告書の目的が有病率の決定であることを考慮した上で、この分析における当局の結論を検討すること。当局は適切なデータを利用していたか？ また、データは妥当であったか？分析方法の適合性、アプローチの透明性、結果の頑健性に的を絞ること。」

・ **一般的アプローチ：**明らかに、BSE やその他の疾患の有病率を求めるにはさまざまな方法がある。したがって、OIE の陸生動物衛生規約（第 2.3.13 条「牛海綿状脳症」および附則第 3.8.4 条「牛海綿状脳症に関するサーベイランス」）で概説されている一般的アプローチを USDA が用いたことは当然かつ合理的である。さらに、WTO( World Trade Organization : 世界貿易機関 )の SPS 協定( agreement on the application of sanitary and phytosanitary measures : 衛生植物検疫措置の適用に関する協定 )の促進および利用の背後にある重要な原則の 1 つは、疾患の管理およびサーベイランスに関する方策が、貿易パートナー間で調和し、互いに同等で、透明性を有しているということである。また、BSurvE が特に各国牛群における BSE 感染率の推定および各国のサーベイランス計画のために欧州委員会の後援を受けて開発されたモデルであることを考慮すれば、BSurvE モデルを今回の分析の主軸として利用したことも当然であった。さらに、EFSA ( European Food Safety Authority : 欧州食品安全機関 )は科学専門家ワーキンググループを任命して、BSurvE モデルのレビューを実施した。その報告書の要旨には次のように記されている。「……このモデルは優れた所産であり、適切な各国 BSE サーベイランス計画の開発における大きな進歩であろう。これは非常に強力なツールで、技術的な面でうまく設計されている。操作性も比較的良好な上、極めて入手しやすく、高い透明性を有する。さらに、モデルの構造も理にかなっているようだと結論がなされた……」BSurvE モデルおよび文書のレビューでは、仮定および方法の透明性が高いことを見いだした。加えて、このモデルを今回の分析に適用することに関して USDA が提供した文書は非常に明解であった。

BSurvE モデルで得られた結果を BBC モデルと BSurvE モデルの併用によって得られた結果と比較したことも適切であった。米国における BSE 感染率に関して BSurvE モデルでは考慮されなかった別の情報（初期有病率にかかわらず、反芻動物に対する飼料規制導入後には BSE 発生率が低下した可能性が高いという情報）が利用できるため、この情報を考慮するようにベイズ型のアプローチを利用することは理にかなっている。

・ **USDA の結論の検討：**本報告書のエグゼクティブ・サマリーに述べられているとおり、USDA の結論は、記載の方法を用いて今日まで分析されてきたデータに基づく適切な結論で

あると思われる。他の課題質問でこれらの研究の完全性について尋ねられているが、この点に関しては考慮されるべき事項がまだ他にある。しかし上述のとおり、本報告書の結論は適切であると思われる。

本報告書の別の箇所で述べられている結論には適切でないと思われるものがある。本報告書の「不確定性分析の結果」の節（24 ページ）に、「..... [ BSurvE モデルと BBC モデルからそれぞれ得られた ] 結果の類似性は、分析に用いた方法、仮定、入力が合理的であり、米国における BSE 有病率を正確に反映しているということを示唆している。」との記述がある。反対の見解をとる人なら、その「結果の類似性」は 2 つのモデルが同じように偏っていて、同じように不正確で、仮定もまた同じように不適切であることを示していると結論することも同様に妥当だと主張するかもしれない。BBC モデルから得られた有病率の点推定値および信頼区間はベースラインである BSurvE モデルで得られたそれらよりも低かったが大きな相違はない、という結論の方が、より合理的で反論を受けにくいと思われる。飼料規制が重大な意味を持つ疾患管理・予防策であることは明らかだが、2 つのモデルの結果が類似していたということは、この方策の有効性について強い仮定を立てなくとも、1996 年の時点での BSE 有病率が牛 100 万頭あたり 1 頭未満である可能性が極めて高いという結論は得られるということを示している。

BSurvE モデルで得られた推定値を BBC 分析に用いていること、および、BSurvE の結果は BBC モデルの結果ほど慎重ではないということとを考慮すると、USDA は BSurvE の結果の強調を控えすぎるべきでなく、また、BBC モデルの結果を強調しすぎるべきでもないと思われる。そういった理由で、BSurvE モデルに関する入力事項および仮定のバリデーションが厳密かつ完全なものでなければならない（おそらく BBC モデルの場合よりも厳密である）ということは、絶対不可欠であると思われる。

**USDA の課題 2)** 「文書中で述べられている分析の目的は以下のとおりである。米国は 1990 年以來、BSE サーベイランスを鋭意実施してきており、2003 年にカナダの牛 1 頭が陽性と確認されてからはさらに取り組みが強化されてきた。米国における強化 BSE サーベイランス計画の概要は別の報告書 (USDA、APHIS、CEAH (Centers for Epidemiology and Animal Health: 疫学・家畜衛生センター) 『BSE 強化サーベイランスの概要』 (Summary of Enhanced BSE Surveillance) 4-6-06) にまとめられている。本分析では、2006 年 3 月 17 日以前の 7 年間にわたって収集されたサーベイランスのデータを用いて、米国における BSE の有病率を推定している。OIE (World Organization for Animal Health: 国際獣疫事務局) が採択している国際ガイドラインに準拠した米国の BSE 状況全般の考察が今後要求される際、この情報は手引きおよび支援となるであろう。」

・ 上記の通り、この項目は本報告書の背景および目的を簡単に記述したものである。次以降の質問項目で扱われていない個別の質問またはレビューに対する指示はこの項目中で見いだせなかった。尋ねられていたのは、目的が合理的および適切であったかどうか、また、モデルはこの問題に明確に取り組んでいたかどうかということだろうか？ 質問がこの2つだとすれば、答えはいずれも「はい」である。

**USDA の課題 3)** 「こういった分析目的との関連において、モデルは完成しているか？ また、理にかなっているか？ 現実社会の状況において有病率の推定に影響する重要な要因やこうした要因間の相互作用をモデルは踏まえているか？ 踏まえていないとすれば、抜けている（または正しく特定されていない）パラメータはどのようなものか？ 研究する価値があったであろうモデルは他に存在するか？ モデルが現実から乖離する場合は明確に定義されているか？」

・ この質問を検討するにあたって、OIE 陸生動物衛生規約の第 3.8.4.1 条の記述を読み直すことが有用かつ妥当であることを見いだした。

「国、地域、またはコンパートメントの BSE ステータスは、サーベイランス計画のみに基づいて決定できるものではなく、第 2.3.13.2 条に挙げられたすべての要因に従って決定されるべきである。サーベイランス計画は、上述のセクターに関連した診断上の限界、およびセクター間での感染牛の相対的な分布を考慮に入れるべきである。本附則における目標ポイント (target point) およびサーベイランス評価ポイント (surveillance point value) は、下記の要因を統計学的モデルに適用することによって得られた。

- 成牛 10 万頭あたり 1 頭という有病率
- 95% という信頼性
- BSE の病因、ならびに病理学および臨床的な BSE の発現
- + 使用された診断方法の感度
- + 年齢による相対的な発現頻度
- + 各部分集団内における相対的な発現頻度
- + 臨床上の病理学的変化と臨床的発現の間の期間
- 牛群の統計値 (年齢分布を含む)
- 4 種類の部分集団を通じた牛群からの牛の淘汰または頭数削減に BSE が及ぼす影響
- 摘発されていない牛群における感染牛のパーセンテージ」

・ BSurvE モデルおよび BBC モデルの利用に際し、OIE が綿密に記述している上記の重要な要因を USDA は適切に考慮したと思われ、また、モデルのパラメータも適切なようである。BSurvE モデルがこれらの同じパラメータに焦点を合わせていること、および、EU と



EFSA が任命した専門家委員会が BSurvE モデルのレビューを実施してこのモデルが適切かつ頑健であるとしたことは、このモデルの妥当性をさらに裏付けている。しかし、BSurvE のテンプレートが実験データの規格に関して幅広い柔軟性および特異性を許している限り、一部のサーベイランスデータの制約を逸脱していたかどうかという点を USDA は慎重に検討する必要があると思われる。このことは、本当はモデルの記述およびパラメータ表示が適切かどうかという問題ではなく、むしろ仮定および適用に関連する問題である。そういった理由で、生じた疑問および留保の考察は次以降の質問項目の中で詳しく行うこととする。

- ・ BBC モデルの追加が必要であったかどうか、または分析に対し有効に寄与したかどうかを検討するにあたって、OIE 陸生動物衛生規約の第 3.8.4.1 条の最初の文（上記に引用済み）を考慮することが重要と思われる。この文の意味するところの 1 つは、BSE 有病率を推定するには孤立したサーベイランス結果だけでなく、過去および現在のリスク因子のほか、適切な疾患管理対策が当該期間に実施されていたかどうかをも考慮する必要があるということである。また、世界の科学界が一致して BSE の管理および予防に関してもっとも重要であるとしている 3 つの要因、すなわち 1) 生きた牛の輸入の適切な管理、2) 飼料の輸入の適切な管理、3) 反芻動物または哺乳類に由来する蛋白質のフィードバンを厳密に実施することについても考慮する必要がある。まとめるとこれらは、今回の有病数推定の試みから適切な結論を引き出そうとするならば BBC モデルで得られた結果を考慮する必要があることを示していると思われる。重要性は低いことであるが、このことは USDA が BSurvE の B 型サーベイランス用ワークシートの使用を決定したことの裏付けでもあると思われる。

- ・ **分析モデルに内在する仮定(本報告書の 10~11 ページにて考察):** 2 つの仮定を除いて、当該箇所述べられている仮定に対する反対意見やコメントはない。

- 米国における飼料規制は英国の飼料規制と同程度かそれ以上に有効であるとの仮定に関連して、この仮定はおおむね容認できるものであると考える。しかし、米国における相対的な発生減少率は飼料規制導入後の英国において観察された相対的な発生減少率に正比例するとは限らないと主張することができよう。例えば、英国では BSE 有病率が非常に高かったこと、および汚染飼料の量も非常に多かったであろうことを考慮した上で、米国では規制開始時点ですでに汚染飼料が少なかったと仮定すれば、米国の飼料規制よりも英国の飼料規制の方が大きな影響をもたらしたと主張できるだろう。言い換えれば、米国における発生減少率は英国における流行の後期段階で見られる発生減少傾向に近いのかもしれない。同時に、前述のとおり、BBC モデルの結果はベースラインである BSurvE モデルの結果と比べてそれほど大きく異なっていないということは、USDA の結論はこの仮定に関して頑

健であるということを示唆している。

- 標本集合の代表の程度についてもっと詳細な考察を行い、このことが分析の結果や結論に影響した可能性があるかどうかを USDA は示す必要があると思われる。本報告書の 11 ページには、BSurvE は標本間の独立性を必要とするが「無作為」抽出されたものである必要はないとの記述がある一方で、標本は参照集団を「代表」しているべきであるとも強調している。このことは OIE の陸生動物衛生規約に記された一般的ガイドラインと一致している。このことに関して、「代表」という言葉の正しい定義あるいは解釈はどのようなものだろうか？ この言葉は、「もう一度標本抽出を行ってもほぼ同じ結果が得られるだろう」ということを意味していると思われる。今回の場合、標本抽出はサーベイランス計画が行われたすべての年で等しく有効であったという内在的な仮定があるため、「代表」という言葉は、「もう一度標本抽出を行い、その際に標本抽出比に変動があっても、本質的に同じ結果が得られるだろう」ということも意味している。このことは、適用される定義の標本抽出の時期による違いも含むだろう。このことに関して、標本抽出期間全体を通して標本抽出が真に「代表」的であったかどうかを USDA が十分に考察したとは思われないし、記録にあるこうした変化がこの調査の結論に影響しなかったとは思われない。例えば、7 年間にわたる標本抽出期間の間に標本抽出方法が大きく変化したこと、およびサーベイランス評価ポイントが強化 BSE サーベイランス計画に示された後期の標本抽出期間に強く重み付けしていることを考慮すると、標本抽出作業がもっと均一に再配分されていたら、あるいは前期の標本抽出期間に重きを置いていたとしたら、どうなっていただろうか？このことに関する USDA の仮定が誤っているとの確信はないが、USDA が正しいと強く確信させてくれるような情報を示しているわけでもない。USDA は他の仮定に関してはもっと強いデータを示している。こうした不安は、標本集合（基礎となる牛群ではない）の年齢に関する仮定の妥当性についての懸念によって強められている。このことは次の課題質問に関連して考察する。

USDA は、この問題に関してより詳細な考察を行うとともに、不確定性分析で検討されるパラメータにこのパラメータを追加することでこの問題に取り組むことができるだろう。

**USDA の課題 4)** 「入力パラメータに関する仮定事項を検討すること。分析目的との関連において仮定は合理的であるか？ 仮定にはもっとも蓋然性の高い合理的な推定パラメータ値が反映されているか？」

・ **牛群統計値および年齢分布について：**

- **参照牛群：**米国の一般的な牛群統計値に関するパラメータについて、USDA は非常に良いデータをもっていると思われる。しかし、USDA が示しているとおり、出生コホートの実

際の年齢統計値および年齢に伴う退出確率に則してモデル中の理想的牛群に関する値を見積もるには仮定が必要であった。分析の目的から、仮定は適切であったと納得した。不確定性分析の結果もこのことに関する安心材料であった。

- **サーベイランス群**：サーベイランス群に用いられた年齢に関するいくつかの問題に関心を抱いた。第1に、『強化サーベイランス報告書』に示されているとおり、前期（1999～2003年）に標本抽出された動物の年齢データはカテゴリー化して回収されていたが、このことに関する USDA の仮定は、この期間中の真の年齢分布は強化サーベイランス期間中に得られた標本の年齢分布によって代表されるとというものであった。サーベイランス期間の前期と後期で標本抽出方法が大きく異なっていたことを考慮に入れば、さらなる正当化が保証されると思われる。

第2に、サーベイランス群の牛の年齢を1歳刻みで最大17歳まで推定できる（見積もることができる）と仮定することは妥当でないかもしれないと思われる。『強化サーベイランス報告書』の表1は、この期間中に収集された標本の約半分はレンダリング施設から得られたということを示している。レンダリング施設は無症状死亡牛として記録される牛（したがってサーベイランス区分は「死亡牛」に割り付けられる）を集めていた可能性が高い。

年齢特異的なサーベイランス評価ポイント割り付け方法について調べると、すべてのサーベイランス区分において5歳の牛がもっとも高いサーベイランス評価ポイントを獲得することがわかる。強化サーベイランスの対象期間中は牛の実年齢が要求されていたが、およそ5歳に達した後の牛の年齢を外見上の身体的パラメータだけを用いて正確に知ることは本質的に不可能であることから、これらのデータは偏っていると思われる。牛の年齢推定に用いられるもっとも一般的な方法は、歯列を調べることである。8本すべての永久切歯が萌出して摩耗し始めたら（通常は約5歳でそうなると考えられる）、ある程度の年齢に達したと考えられる時期になって歯が抜け始めるか摩耗しきるまで、牛の年齢を正確に知るとは極めて困難になる。そういった理由で、1歳刻みで年齢を報告することが妥当であったと仮定するよりも、サーベイランスのデータを年齢カテゴリーに分類する方が合理的であったかもしれないと思われる。この潜在的な誤分類による影響はどのようなものでありえたかについて、USDA が再検討するよう勧めたい。（おそらくは OIE の陸生動物衛生規約に記された手法を用いて）牛の年齢をモデル分析用に集約することができるし、また、サーベイランス評価ポイントを再割り付けした後に、こうしたモデルから得られた結果をベースラインのモデルと比較することで、この情報を不確定性分析に織り込むことができるだろう。

\*\*\*\*\*

表 1. 2004 年 6 月 1 日から 2006 年 3 月 17 日までの期間に各種収集施設において検査された対象標本の数

収集施設
と殺場*
レンダリング施設
農場
公衆衛生検査室
診断検査室
3D-4D 施設
その他
合計

対象標本数全体に対する比率 (%)

\*標本採取のために用地外の施設 (3D-4D 収集施設) に輸送されたと殺前の廃用決定牛は含まない。

\*\*\*\*\*

・ 標本集団に含まれる牛を「擬似症状」サーベイランス区分に割り付ける際に補助として用いられる尤度比の値に関して、USDA の文書『米国における強化 BSE サーベイランスの概要』にある主張には同意しない<sup>2</sup>。USDA は、BSE の診断に関して複数の臨床症状がそれぞれ独立であることを仮定しているが、それにもかかわらず、同じ文書の中でこの仮定と矛盾するデータが存在すると述べている。直観的に言えば、USDA が仮定したように、患畜が何らかの感染症状を示す場合に複数の症状を示すことはないだろうと仮定することは無意味と思われる。しかし、この問題は不確定性分析の中で扱われている。この仮定は不正確であると思われ、そのために敏感度および特異度の推定値や尤度比の推定にに影響すると考えられるが、不確定性分析の結果、モデルがこの仮定に対して頑健であることが示されていることから、この仮定が結果に影響して全体の結論に有意な偏りをもたらしているわけではないということに納得している。

・ 特に不確定性分析の結果を考慮すると、その他の入力パラメータに関する仮定は合理的であると思われた。

USDA の課題 5) 「モデルおよび入力データの不確定性に対して結論は頑健性を示すか？ 不確定な事項は完全に記述されているか？ 結論の頑健性はどの程度か？ これらの関係

<sup>2</sup> RTI による注：レビューは、主たる報告書で参照されているが今回のレビューの対象ではない裏付け文書についてコメントしている。こうしたレビューは特に要請されていないが、裏付け文書についての結果はこの報告書に示されている分析に意味をもっている。

は明解に記述されているか？」

- ・ 不確定性分析に追加する必要があると思われる 2 つの問題点（サーベイランス標本の代表性、およびサーベイランス群の年齢に関する仮定）については上記で詳しく述べてある。

- ・ モデルの仮定事項に対してモデルが頑健であるかどうかについて、不確定性分析の結果と記述は極めて安心できるものであった。また、方法は非常に明解に詳述されており、透明性の高いものであった。

- ・ 結論への影響について：今回の結論（有病率は成牛 100 万頭あたり 1 頭未満）に影響するような結果の変化を伴わずに新たに 3 件の例を分析に含めることができるということが不確定性分析によって示唆されていること、および、モデルの入力および仮定の変動に関するモデルの頑健性が高いということを考慮すると、このことは本報告書およびその結論の信用性を大きく高めていると思われる。しかしそれでもなお、前述の追加項目には留意したいと思う。

**USDA の課題 6)** 「用いたモデルの下で、分析結果は正しいか？ プログラムを実行してコードおよび数式を点検し、モデルが意図どおりに機能していることを確認すること。用いたモデルの下で、当局が正しい答えを公表したのかどうかを判定すること。」

- ・ 提供されたデータおよびコーディングを用いてモデルを実行してみたところ、その結果と本報告書の記述事項との間に矛盾は見られなかった。しかし、RTI からの説明書で指示されているとおり（「……モデルの純粋なコーディング面よりも方法論的 / 技術的側面に力を入れてください……」）モデルの基本的な仮定が計算において一般にどのように用いられているかを理解する目的で 2 つの数学的モデルを再検討することによって多くの時間を費やした。これらのモデルのコーディングや数式、プログラム特有の機能のすべてを詳細に再検討したわけではない。繰り返しになるが、このことは RTI からの指示の意図と一致していると考えている。

- ・ この結論が「正しい解答」であるかどうかについて最終的な判断を下す前に、ここで挙げた問題点に対する USDA の対応を知りたいと思う。しかし、提供されたものの限界内で、また、さらなる分析でも標本抽出の代表性に関する仮定やサーベイランス群の年齢に関する仮定に対してモデルが頑健であることが示されるということを前提とすれば、今回の結論（有病率は成牛 100 万頭あたり 1 頭未満）は適切であったと思われる。

## 『米国における BSE 有病率の推定』のレビュー

ラリー・G・ペーズリー (Larry G. Paisley) D.V.M., M.S. (理学修士) Ph.D.  
デンマーク獣医学研究所 疫学・リスク評価部門 主席研究員 (デンマーク)

ペーズリー博士はモンタナ州の生まれで、現在はデンマーク・コペンハーゲンにあるデンマーク食品獣医学研究所の主席研究員である。博士は、ワシントン州立大学で獣医学博士号を取得し、ミネソタ大学では家畜繁殖学で理学修士号を得て、さらに王立獣医学農業大学 (デンマーク・コペンハーゲン) からは疫学の博士号を授与されている。

博士は若い頃、家畜繁殖学者として、ワシントン州立大学 (ワシントン州プルマン) およびロス大学 (西インド諸島セントキッツ島バステール) で教鞭をとっていた。

1991年に博士号を授与されてから、ペーズリー博士は USDA : APHIS の獣医学部門 (VS) の地域疫学者としてプエルトリコで勤務し、また、北部区域疫学者としてニューヨーク州オールバニーで勤務したことがある。1996年から博士はノルウェー・オスロの国立獣医学研究所において疫学者 / リスク分析官として勤務し、その後、2000年1月にはデンマーク食品獣医学研究所で、TSE (transmissible spongiform encephalopathy : 伝達性海綿状脳症) のサーベイランス、疾患モデリング、リスク評価を活動の中心とする疫学者としての勤務を開始した。ペーズリー博士はデンマークの獣医学関連当局、植物局、外務省のために数々の TSE / BSE に関するリスク評価を実施してきた。博士は、TSE/BSE や FMD (foot and mouth disease : 口蹄疫)、CSF (classical swine fever : 豚コレラ) に関する複数の EU プロジェクトに参加しているとともに、欧州食品安全機関 (EFSA) のワーキンググループに専門家として招かれたこともある。

### 緒言<sup>3</sup>

私、ラリー・G・ペーズリーは、牛海綿状脳症 (BSE) 有病率を推定するモデル、およびこれに随伴する報告書 (USDA の CEAH (Centers for Epidemiology and Animal Health : 疫学・家畜衛生センター) に属する全米サーベイランスユニットが作成) に関するピアレビューを実施するよう、リサーチトライアングル研究所 (RTI) の食品・農業政策プログラムによって任命された。

レビューの課題には次の事項の評価が含まれる。

- ・ 用いたモデルの適合性、透明性、頑健性

<sup>3</sup> RTI による注 : このレビューは他の 2 編とは構成が異なっている。ペーズリー博士は最初に APHIS のアプローチを概説し、その文脈の中で関連するすべてのコメントを行うという形で、5 つの課題質問に答えている。その後、末尾の「レビューの結論」と題された節においてコメントのまとめと要約を行っている。

- ・ データの妥当性
- ・ モデルに入力する仮定
- ・ モデルの結果の不確定性
- ・ 結論

これらの各項目が考察され、米国における BSE 有病率の推定値に関する全体的な結論が導かれている。

## モデル

BSE 有病率の推定に用いられたモデルは、BSurvE モデル (Wilesmith et al., 2004)、BBC (Bayesian Birth Cohort: Bayesian 出生コホート) モデル、外挿モデル (Cohen et al., 2001, 2003) であった。

### BSurvE モデル

2003 年に欧州委員会は、共同委託研究所 (Community Reference Laboratory) に対し、各国の BSE ステータスを評価するための疫学的に妥当な総合的アプローチを開発するよう要請した。そして、牛群統計値およびサーベイランスデータを用いて現存牛群における BSE 有病率を推定する統計的モデル (BSurvE) が開発された (Wilesmith et al., 2004)。BSurvE モデルは EFSA の科学専門家ワーキンググループからおおむね好意的な評価を受けた。ワーキンググループの結論は、このモデルはうまく設計されており、操作性も比較的良好な上に極めて入手しやすく、しかも高い透明性を有するというものであった。ワーキンググループは、同一のデータおよび仮説を用いる条件で BSurvE の結果を他の逆算法による結果と比較することにより、モデルの完全性を分析するよう推奨した。データや仮定が不正確であった場合のモデルの頑健性を評価するため、一連の感度分析を行うことが求められた (EFSA, 2004)。

主な仮定事項は次のとおり。

- ・ その国の全牛群の頭数および年齢分布は年によって変動しない。
- ・ 臨床的発症年齢は対数正規分布に従い、平均値は 5 歳である。
- ・ 臨床症状を示して牛群から退出する感染牛のサーベイランス区分間における分布は時とともに変化しない。
- ・ 1 歳未満の牛だけが感染に対して感受性を有する。
- ・ 臨床症状発症の 0~3 カ月前に実施した場合の検査の感度は 100% である (改訂版による)。
- ・ BSE の自然発症例は存在しない (EFSA, 2004)。

必要なデータは次のとおり。

- ・ 牛群の頭数
- ・ 牛群の年齢分布
- ・ BSE 検査データ
- 年齢ごと
- サーベイランス区分ごと
  - ・ 正常な屠殺
  - ・ 切迫と殺
  - ・ 死亡牛
  - ・ 臨床疑似患畜

主な結果は次のとおり。

- ・ BSE サーベイランス評価
- ・ BSE ステータス評価
- 常在国の場合
  - ・ 現存牛群における BSE 有病率（および信頼限界）
  - ・ 出生コホートにおける BSE 有病率（および信頼限界）
- 非常在国の場合
  - ・ サーベイランスリソースの割り当て

### **Bayesian 出生コホートモデル**

Bayesian 出生コホートモデルは BSurvE モデルの結果を利用する。「Bayesian 出生コホート法は、BSurvE モデルの根底をなす疫学的な理論と、飼料規制が有病率に及ぼす影響に関する情報とを組み合わせることにより、米国における有病率のより正確な推定値を与える。この方法では出発点として、0%から 100%の間の任意の値を有病率として仮定し、次に標本抽出された米国の各出生コホートに関するサーベイランス標本抽出ポイントの合計を用いて初期値を更新する」(CEAH, 2006)。また、米国で 1997 年に実施された飼料規制は英国で 1988 年に制定された飼料規制と同程度かそれ以上に有効であったと仮定している。英国の 1988 年以降の出生コホートにおける BSE 発生率減少に基づくデータを用いて、米国の飼料規制の効果をモデリングした。

### **外挿モデル (CEAH, 2006)**

外挿モデルではまず、ベータ分布に従う標本抽出データから直接的に対象牛群における BSE 有病率を推定する。非対象牛群における有病率は、対象牛群における有病率と比例するものとして推定する。サーベイランス感度の推定値（約 40%）を用いて、みかけの有病率を補正する。EU における検査のデータは、高リスク牛群では BSE の可能性が 28 倍高いことを示唆している。全牛群の有病数は、高リスク牛群は全牛群の 3%を占めるとの仮定を用い



て、部分集団の加重平均から推定される。

「見かけの有病率分布を感度（40％）で割り、生じたベータ分布を解くことによって、高リスク群における真の有病率分布の推定値が得られる。高リスク群における「真」の BSE 有病率分布の推定値に 1 対 28 という比率を適用し、生じたベータ分布を解くと、健常群における「真」の有病率の推定値が得られる。高リスク群および健常群のベータ分布の加重平均をシミュレートすると、有病率分布の平均値、5 パーセンタイル値、95 パーセンタイル値の推定値が得られる」(CEAH, 2006)。

## 米国における BSE 有病率の推定（草案）に対するコメント

### モデルへのデータ入力

#### 標本

本報告書は、分析対象である BSE サーベイランス標本は 2006 年 3 月 17 日までの 7 年間にわたって収集されたと述べている。抽出された標本の総数は 735,213 頭とのことである。「正常な屠殺」区分は 23,322 頭、「死亡牛」区分は 571,888 頭、「切迫と殺」区分は 134,232 頭、「臨床疑似患畜」区分は 5,771 頭であった。有病率 B のスプレッドシート（表 17a および表 17b）において、陰性が 735,054 頭、陽性が 2 頭となっている点にはやや問題がある。ただしこれは軽微なひっかかりに過ぎない。表 17 のデータは検査結果（表 13）から導かれている。このデータは 5 年分の検査結果である。7 年分の標本抽出データをどのようにして 5 年分の検査データに組み入れているのかが不明である。

### サーベイランス区分

強化サーベイランス制度の下で抽出された標本は、BSurvE モデルの要件に従ってサーベイランス区分に割り付けられた。強化サーベイランス制度の実施前に収集された標本は、検査室の提出書類に記された臨床歴に基づき、疫学者によってサーベイランス区分に割り付けられた。強化サーベイランス実施中に検査された牛の多くでは、BSE と共通する症状が記録されていたが、臨床疑似患畜カテゴリーへの割り付けはなされなかった。こうした牛のうち、米国の対象牛群と比べて BSE 症例で観察される可能性が 807 倍以上高い症状を示したものは、臨床疑似患畜カテゴリーに再割り付けされた。この方法が正当であるとする根拠は APHIS, 2006 に見いだせる。

### 牛群の年齢分布によって影響されるパラメータ

牛群の年齢分布は、非感染牛および感染牛の退出確率、および各サーベイランス区分から退出する非感染牛と感染牛の割合に影響する。米国の牛群の正確な年齢分布は不明であるが、詳しく記述された方法（Appendix A in CEAH, 2006）によって推定された。

## **モデル方法における仮定**

### **症状発現前に検出できる割合**

BSE 感染牛の 40% は臨床症状を呈する前の 1 年の間に検出可能になるとの仮定がなされた。BSurvE モデルにおいてこの値を変更することは可能である。現在の情報は、臨床症状発症の約 3 ヶ月前には BSE 感染を迅速検査で検出できるので、値は 25% とする方が合理的かもしれないということを示唆している。しかし本報告書の感度分析の結果は、この推定値が有病数の推定値にほとんど影響しないことを示唆している。

### **生後 1 年以内の感染**

これは合理的な仮定である。

### **有病率が一定 / 減少**

BSurvE モデルでは時間が経っても発生率は一定であると仮定し、一方、BBC モデルでは飼料規制の後に発生率は減少すると仮定する。しかしこうした仮定の違いは、有病率の推定値にほとんど影響しなかった。

### **米国の飼料規制は 1988 年の英国における規制と同程度かそれ以上に有効である**

これは合理的な仮定である。

### **検査の感度**

BSurvE モデルでは、BSE の臨床段階にある動物に関する BSE 検査の感度は 100% であると仮定している。症状発現前の段階で検査した場合の感度は 40% である。上記で考察したとおり、この推定値は 25% としたほうがよいかもしれない。

### **疾患に伴う挙動**

コメントなし。

### **標本分布の影響**

コメントなし。

### **BBC モデルに関する事前分布**

コメントなし。

## BSurvE 有病率 B 方式による有病率の推定

BSurvE は BSE サーベイランス計画の複数の側面を分析するよう設計されているが、本報告書では有病率 B 方式による推定方法だけを用いている。有病率 B 方式は、サーベイランス計画を実施しているが BSE 症例がゼロまたは極めて少ない国で使用されることを意図している。有病率 B 方式を選択したことは米国に関しては適切である。この方法では、ある国における感染率が比較的一定のままである期間の存在を仮定し、全サーベイランスデータにわたって有病率を推定する。このモデルの記述は明解であり、BSurvE の文書 (Wilesmith et al., 2004) と一致しているようである。

## Bayesian 出生コホートによる有病率の推定

Bayesian 出生コホートモデルは BSurvE モデルを拡張したもので、1997 年実施の飼料規制に関する情報を組み入れて現存牛群における有病率を推定する。このモデルでは感染率について、飼料規制の前は一定で、その後の 5 年間は減少すると仮定していた。この減少は、英国において 1988 年の飼料規制後に観察された減少と同じ速度で起こると仮定された。

モデルの記述は明解であり、手法は妥当である。

BBC モデルの結果は BSurvE モデルから得られた結果に大きく左右されるということは指摘されるべきである。BBC モデルでは、BSurvE モデルによって割り当てられたサーベイランス評価ポイントを年ごとに集計したものをを用いて有病率を推定する。

## 不確定性分析

不確定性分析は、BSE 有病率の推定値に関するモデルの入力および仮定の影響を評価するための仕組みである。有病率の推定値にもっとも大きな影響を及ぼすような不確定性を生じうる原因として 3 項目が挙げられた。項目は次のとおり。

- ・ BSurvE アルゴリズムとその仮定に対する有病率推定値の感度
- ・ 同じ量の陰性サーベイランス結果とともに症例を追加すること (例えばカナダで発生した症例など) に対する有病率推定値の感度
- ・ BSurvE モデルへの入力パラメータに関するさまざまな変動に対する有病率推定値の感度

不確定性分析には 2 種類の方法が用いられた。

- ・ BSens (BSurvE モデルについての感度分析を行うよう設計されたソフトウェア)
- ・ 外挿モデル (CEAH, 2006)

### **有病率推定の結果**

3種類の方法（BSurvE 有病率 B 方式、BBC、外挿モデル）によって得られた有病率推定の結果が示されている。得られた結果はモデルの入力事項と首尾一貫しており、また各モデルは正しく実行されている。予想どおり、BBC モデルによる有病率推定値は有病率 B 方式による値よりも小さかったが、これは BBC モデルが飼料規制に関する情報を組み入れているためである。しかし、どちらの方法でも、米国の牛群における BSE 有病率は極めて低く、100 万頭あたり 1 頭未満であるとしている。

### **不確定性分析の結果**

これはおそらく、本報告書の中でもっとも情報価値があり有用な部分であろう。

### **BSurvE アルゴリズムおよび仮定に対する有病率推定値の感度**

BSurvE アルゴリズムおよび仮定に対する有病率推定値の感度を外挿法によって推定した。外挿法では有病率 B 方式および BBC モデルよりもやや高い推定値が得られた。しかし外挿法においても BSE 有病率は 100 万頭あたり 1 頭未満であることが示唆されている。

### **陽性牛を 1~3 頭追加することに対する有病率推定値の感度**

1~3 件の症例を追加することは、有病率 B 方式および BBC モデルによる有病率推定値にわずかな影響しか及ぼさなかった。米国の牛群の大きさを考えれば、これは驚くことではない。

### **BSurvE モデルの仮定および入力パラメータに関する変動に対する有病率推定値の感度 尤度比**

BSE サーベイランス標本の一部を臨床疑似患畜カテゴリーに正しく割り付けるためにかなりの数が費やされた。これが行われたのは明らかである。その理由は、サーベイランス評価ポイント（これを元にして BSurvE モデルおよび BBC モデルによる有病率の推定が行われる）のほとんどが臨床疑似患畜に由来することにある。驚くことではないが、この方法は有病率推定値に関する不確定性をもたらした。しかし、807 というカットオフ値を増減しても有病率推定値にはほとんど影響しなかった。

不確定性の第 2 の原因は、尤度比の分母にすべての陰性サーベイランスデータを含めることであった。「原因不明死亡牛」に分類される標本を分母から除外すると、臨床疑似患畜に分類される標本数が少なくなるが、有病率推定値にはほとんど影響しなかった。

年齢分布（BSurvE の変数  $D_{j,t}$  および  $C_{j,t}$ ）、発症前検出、非感染牛および感染牛が各サー

## ベイランス区分から退出する確率（割合）に対する有病率推定値の感度

非感染牛および感染牛の退出に関する定数 ( $D_{j,t}$  および  $C_{j,t}$ )

牛群の年齢分布は非感染牛および感染牛の退出確率に影響する。プロジェクトグループはかなりの労力を費やして、米国の牛群の年齢分布を推定する方法を詳述している。「我々はこの入力値が米国の牛群に関して正確であることに大きな自信を持っており、また、不確定性の原因となりうるこの事項が分析結果に実質的な影響を及ぼすことはありそうにないと考えている」との記述には同意する。

米国に特定されたデータが存在しないため、BSurvE モデルのデフォルト値が用いられた。BSurvE モデルで用いられている退出確率のデフォルト推定値にはやや問題がある。第 1 に、この文書や BSurvE モデルの説明書 (Wilesmith et al., 2004; Wilesmith et al., 2004) の中に推定値に関する明確な説明または正当とする根拠がなく、ただ英国や EU のデータに由来するとあるだけである。理論的には、EU 諸国では 24~30 カ月齢を超えた牛の大半が検査されている。2004 年の場合、「正常な屠殺」、「死亡牛」、「切迫と殺」、「擬似症状」の各サーベイランス区分で BSE 検査が実施された割合は、それぞれ約 88%、10%、0.7%、0.03%であった (EC, 2005)。BSurvE では各サーベイランス区分に 89%、7%、4%、0.09%と入力するよう示唆している。この場合、BSurvE の推定値は合理的であるように見受けられる。

2004 年の EU において、「正常な屠殺」、「死亡牛」、「切迫と殺」、「擬似症状」の各サーベイランス区分での BSE 症例の割合は、それぞれ約 29%、49%、23%、18%であった。しかし、BSurvE モデルにおける BSE 症例に関するデフォルト値はそれぞれ約 17%、10%、10%、64%とされている（「パラメータ」ワークシートの表 9 を参照）。これは何かが間違っているように思われる。全サーベイランスにおいて実際に見つかった BSE 症例がどのサーベイランス区分に属するかという割合の方が、BSurvE モデルで示されている値よりも良い退出確率の推定値であると思われる。これほどの規模の差異が有病率推定値に有意な影響を及ぼさなかったということはやや驚きである。退出確率の変動が米国における有病率推定値にどのような影響を及ぼすかは、時間がないために判定できなかった。この点に関してもっと検討してもらいたいと思う。EU の BSE 検査データを徹底的に分析することが役立つかもしれない。

## 現存する症状発現前の牛が検出できる割合

現存する症状発現前の牛が検出できる割合が変動しても有病率推定値に起こる変化はわずかであるということに同意する。

時間  $t$  において退出するとしたときに非感染牛 (BSurvE の変数  $d_{j,t}$ ) および感染牛 (BSurvE の変数  $c_{j,t}$ ) が各サーベイランス区分から退出する確率（割合、 $j$ ) (BSurvE の表 4 および

## 表 5 より)

各退出確率の変動による影響については徹底した考察がなされており、事実上、退出確率の変動に対する有病率推定値の感度は比較的低いということが示唆されている。しかし、米国のサーベイランスデータを  $d_{jt}$  の割合に含めながらも「正常な屠殺」区分に関する BSurvE の割合をそのままにしたことは、有病率推定値を増加させる影響があった。EU の検査データに基づく退出確率を用いたときに有病率推定値に現れる影響を知ることは興味深いことであろう。また、いくつかの入力値が同時に変化したときの影響がわかれば情報価値があることだろう。

## 不確定性分析の結論

表 6 に示されているとおり、有病率 B 方式モデルおよび BBC モデルは入力値の変動に対して極めて頑健である。入力値にどのような変化があっても、有病率推定値は極めて低いままであった。3 件の症例を追加したとしても、有病率の推定値は 100 万頭あたり 1 頭未満であった。

## レビューにもとづく結論

レビューの対象は以下のとおり。

- 採用されたモデルの適性・透明性・確実性
- データの妥当性
- モデルへの入力データの前提事項
- モデルの結果にともなう不確実性
- 結論

## 採用されたモデルの適正・透明性・確実性

BSurvEモデルは、統計的・疫学的に妥当なモデルとして欧州食品安全機関（EFSA）が好評価しており、BSEの有病率を推定する上で有効と考えられている（EFSA 2004）。このモデルの透明性は概ね良好であり、CEAHの報告書も同様に透明性が高い。したがって、BSurvEモデルの確実性はきわめて高い。

1997年の飼料禁輸措置に関する情報を利用するベイズのBayesian出生コホート（BBC）モデルもまた適性・透明性・確実性が高いが、BSurvEモデルによるサーベイランス・ポイントをもとに有病率が計算されているため、両者の結果は相互に密接に関連している。

不確実性の分析に利用された「推定モデル」は前提事項の少ない比較的単純なモデルである。このモデルにもとづく報告書の記述は透明性が高く、その意味では不確実性の分析に適している。また、同モデルにもとづく推定有病率は、有病率BおよびBBCモデルによる推定値より若干高いものの、米国におけるBSE感染牛の個体数がきわめて低いことを示している。

## データの妥当性

BSurvEモデルには、4つのサーベイランス・ストリームに配分された「個体数・年齢構成・BSE試験に関するデータ」が必要となるが、こうしたデータは一部の国では入手困難である。米国の牛の個体数と年齢構成の推定値は分析対象として妥当と考えられる。一方、健全牛の解体のストリームにおいて実施された試験は限られており、各サーベイランス・ストリームにおける試験の配分に不備がある。そのため、サーベイランス・データの質はそれほど高くないが、有病率の推定には十分と考えられる。

## モデルへの入力データの前提事項

モデルにおける前提事項は概ね妥当である。しかし感染牛の死亡確率については若干の疑問が残るため、その信憑性をさらに分析する必要がある。この死亡確率により、健全牛と感染牛がどのサーベイランス・ストリームで死亡するかがわかる。また、感染牛の死亡確率は、それぞれのサーベイランス・ストリームで特定されるBSE症例の比率に大きく影響するが、同確率が変動しても「米国におけるBSEの感染率はきわめて低い」という結論に大きな影響が及ぶとは考えられない。

## モデルの結果にともなう不確実性

モデルの結果にともなう不確実性の原因の多くはすでに特定・分析されている。それぞれの入力パラメータにおける変動の大半は推定有病率に影響するほど重大とは考えられないが、同パラメータのいくつかは同時に変動した場合の影響は未知である。また、感染牛の年齢分布や潜伏期間の分布等、まだ特定・分析されていない不確実要素もある。感染を想定する際の牛齢の選択や潜伏期間の分布の変化も推定有病率に影響する可能性がある（Ferguson 他《1997年》、ArnoldとWilesmith《2004年》）。

## 結論

推定有病率には不確実性がともなうため、報告書が提示する実数とその信頼区間を過信してはならないが、「米国におけるBSEの有病率は成牛100万頭につき1頭未満」という結論には十分な根拠があると考えられる。また、死亡牛や健康な老牛を対象に今後実施されるBSE検査により、推定有病率の信頼性はさらに高まるものと思われる。

## 参考文献

APHIS ( 2006 ). USDA:APHIS: VS, CEAH. Summary of Enhanced BSE Surveillance in the United States, April 28, 2006. ( 米国におけるBSE強化サーベイランスの概要 )

Arnold, M. & Wilesmith, J.( 2004 ). Estimation of the age dependent BSE infection of dairy animals in Great Britain. Preventive Veterinary Medicine, 66:35-47.( 英国の乳畜における牛齢別BSE感染率の推定 )

Centers for Epidemiology and Animal Health ( CEAH ) ( 2006 ) An Estimate of the Prevalence of BSE in the United States ( DRAFT ) ( 米国におけるBSE有病率の推定《草稿》 )

EC ( European Commission ) . ( 2005 ) Report on the Monitoring and Testing of Ruminants for the Presence of Transmissible Spongiform Encephalopathy ( TSE ) in the EU in 2004.

[http://europa.eu.int/comm/food/food/biosafety/bse/annual\\_report\\_tse2004\\_en.pdf](http://europa.eu.int/comm/food/food/biosafety/bse/annual_report_tse2004_en.pdf)

( EUにおける反芻動物のTSEに関する監視と試験《2004年》 )

EFSA ( European Food Safety Authority ) . ( 2004 ) Scientific Report of the European Food Safety Authority on the BSE surveillance model ( BSurvE ) established by the Community Reference Laboratory for TSE.

[http://www.efsa.eu.int/science/tse\\_assessments/bse\\_tse/661/sr17\\_bsetse\\_bsurve\\_en1.pdf](http://www.efsa.eu.int/science/tse_assessments/bse_tse/661/sr17_bsetse_bsurve_en1.pdf)

( TSEを対象とした共同委託研究所作成のBSEサーベイランスモデル《BSurvE》に関する欧州食品安全機関の報告書 )

Ferguson, N. M., Donnelly, C. A., Woolhouse, M.E.J. & Anderson, R. M.( 1997 ). The epidemiology of BSE in cattle herds in Great Britain .2. Model construction and analysis of transmission dynamics. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 352:803-838. ( 英国の牛におけるBSEの疫学、2 . 伝播動態のモデル構築とその分析 )



Wilesmith J, Morris R, Stevenson M, Prattley D, Cannon R, Benard H( 2005 ). BSurvE: A Model for Evaluation of National BSE Prevalence and Surveillance. User instructions for the BSurvE model, Version 05.04b, April 2005. Weybridge, England. Veterinary Laboratories Agency.

( BSurvE:英国におけるBSE有病率とそのサーベイランスを評価するためのモデル )

Wilesmith J, Morris R, Stevenson M, Cannon R, Prattley D, and Benard H( 2004 ) Development of a Method for the Evaluation of National Surveillance Data and Optimization of National Surveillance Strategies for Bovine Spongiform Encephalopathy. Weybridge, England, European Union TSE Community Reference Laboratory, Veterinary Laboratories Agency.

( 英国におけるサーベイランス・データの評価方法の構築とBSEサーベイランス戦略の最適化 )

## 米国におけるBSE推定有病率のレビュー

カリフォルニア州立大学統計学部教授Eric A. Suess

Eric A. Suess教授は、家畜の疾病モニタリングに関する生物統計学・疫学の分野でいくつかの研究プロジェクトに従事している。同分野におけるSuess教授の研究の目的は、ベイズの階層モデリングおよびMCMC法にもとづいて感染率を群・地域・国全体レベルで推定する点にある。統計の応用に際しては一般にRおよびWin Bugsが利用されている。Suess教授の研究対象には、回帰分析・時系列分析法や計算統計の応用、コンサルティング、統計教育等も含まれている。

- 1) 報告書の目的が有病率の算定にあることを踏まえ、当分析に関するUSDAの結論を検証する。USDAが採用したモデルとそのデータは果たして妥当か？ 各種方法の適性、採用した手法の透明性、最終結果の確実性を以下で検証する。

### モデル

疾病の不在に関する研究分野では、採用されるモデル（BSurvEとBBC）は文献を反映している。モンテカルロ・シミュレーションまたは「ベイズのモデルとMCMC法のみを利用したモデル」等、選択肢はいくつかある（参考文献の一覧が必要であれば文献の一部は入手可能）。採用された方法は、シミュレーションより計算に重点を置いているため、答えをより迅速に計算できる。

13ページに示したBSurvEモデルの結果には多くの誤りがある。モデルを提示する際、指標（indices）を取り入れることが重要である。モデルではデータが $x$ ではなく、 $ijx$ で表示されており、同様に $n$ ではなく、 $ijn$ がひとまとめでパラメータとなっている。また、 $f$ と $a$ もモデルには含まれていないため、適当な符号が必要とされる。モデルを提示する際は理解しやすいよう、特別に配慮する必要がある。符号も別途定義しなければならない。

BSurvEモデルはアルゴリズムではない。あくまでも「モデル」と解釈すべきである。

BBCモデルも数学モデルというよりはベイズの統計モデルと解釈すべきである。ベイズの統計は物理学と同様、数学の一部門である。これらのモデルはデータを駆使することから、いずれも統計モデルといえる。

一方、統計の提示方法には問題があり、その結果の意味についての理解も不足している。統計を取り扱う際の言葉使いの問題点について以下に述べることにする。特に、信頼区間

に関する手順が誤って論じられている。BSurvEモデルとソフトウェアを文書化する際に使用している用語を統一しなければならない。全体として見ると、報告書が提示する統計はすべて妥当（表中の数字はいずれも正確である）と考えられるが、各データを表現する用語は必ずしも正確でない。

BSurveEモデルのカテゴリに割り当てられる「point values（ポイント値）」を表す用語は不正確である。そのため、言葉の選択に注意し、表現を統一する必要がある。これはすべての専門用語について検討すべき問題である。

「mathematical（数学的）」、「statistical（統計的）」、「analytical（分析的）」が同義的に使われているが、これらを「統計的」に統一すべきである。また、「probability（確率）」と「proportion（比率）」も「probability（確率）」に統一すべきである。さらに「likelihood（可能性）」が「probability（確率）」の意味で少なくとも1回使われているが、モデルの一部を決定する要素として「likelihood function（尤度関数）」が使われているため、この用法は好ましくない。

10ページ最初の節も意味が曖昧である。「analytic models（分析的モデル）」は「mathematical models（数学的モデル）」または「statistical models（統計的モデル）」とすべきである。データがモデルに組み込まれており、パラメータにともなう不確実性が評価され、それがモデルに含まれていることを踏まえ、「statistical（統計的）」が望ましい。この節では「numbers（数）」という単語が「data（データ）」の意味で使われている。調査においては、「number（数）」と「data value（データ値）」に互換性がない点に留意すること。「number」は数直線上の「number」であり、「data value」は度数分布から抽出されるため、エラーを含み得る測定値である。「number（数）」という言葉をここで使うと確実性について誤解が生じる。これはモデルが意図するところではない（少なくとも筆者はそう解釈する）。データの収集にともない、「sensitivity（感度）」と「specificity（特異性）」が論じられているが、「statistically large numbers（統計的の大きな数）」の意味が筆者には理解できない。統計的モデルを論じる場合、無作為に抽出されるべきデータがその基本となる。こうしたモデルに利用されるデータは比較的「large（多数）」だが、それらは対象の個体群を代表する1サンプルに過ぎない。

## データ

筆者の知る限り、データは妥当なソースに由来しており、正しく解釈されている。もちろんデータの妥当性は誰が収集するかによって変わってくる。感染の可能性が高い牛を対象

にサンプリングした場合、データの分析に多少問題が生じるが、採用された「points（ポイント）」法で対象可能と考えられる。

## 方法

前述の2つのモデル（BSurvEとBBC）はいずれも透明かつ平易な適正モデルである。BSurvEモデルはExcelの表計算ソフトを使っているため、扱いやすく、質の高いマニュアルも備えている。一方、BBCモデルはさらにマニュアル化する必要がある。BBCモデルは部分的にベイズのモデルと解釈できる。パラメータの大部分は定数として扱われ、いずれもBSurvEモデルから引用されている。BBCモデルがベイズのモデルとするならば、BSurvEモデルから引用される数は、ベイズのモデルのパラメータにおける事前事項を決める要素となり得る。たとえば、「values（数値）」は、パラメータにおける適正な事前分布の平均値または最頻値となり得る。同モデルを完全に理解していないかぎり、これがモデルの改善につながるかどうかは判断できないが、定数を導入すればベイズのモデル的要素は少なくなる。

## 結論

論文の主眼はその結果部分ではなく、モデルの感度分析にあるようである。全体的に見ると、専門用語が統一されておらず、統計用語の使い方にも誤りが多い。

特に統計用語の不適切な使い方が問題だが、そのほとんどは単純な誤用である。また、専門用語を使うべきところに一般用語が使われている。一番の問題は、結果部分における「信頼区間」の解釈である。

当報告書が2つのモデルをもとにしていると思われる点も問題である。それらを「mathematical（数学的）」または「analytical（分析的）」モデルと呼ぶのは不適切である。

BSurvEモデルについて「アルゴリズム」という言葉を使うのは不適切であり、矛盾している。BSurvEモデルはアルゴリズムではなく、表計算を基本としている。また、感度分析においてモデルが変化しているような印象を受けるが、これは事実と反する。モデルに入力する数値やデータが変わっても、モデル自体は変わらない。感度分析の目的は、入力値やデータの変化がパラメータの最終推定値（この場合は有病率）に及ぼす影響を見きわめる点にある。表6ではBSurvEモデルが「アルゴリズム」および「モデル」となっているが、すべて「モデル」に統一すべきである。

「Uncertainty Analysis（不確実性分析）」という用語も不正確である。入力値やデータの変化にともなう影響を分析する作業に対し、「Uncertainty Analysis」という用語は一般的でない。この場合、統計学における標準用語は「Sensitivity Analysis（感度分析）」である。専門用語の言い換えは混乱を招く。

## 「信頼区間」と「信用区間」の使い分け

報告書の一番の問題は統計用語の不適切な使い方である。採用された方法はすべて統計モデル（データとモデルの不確実性を部分的に利用するモデル）にもとづいているようである。まず、誤用されている専門用語を訂正しなければならない。特に「Confidence Intervals（信頼区間）」と「Credible Intervals（信用区間）」（またはBayesian Confidence Intervals, ベイズの信頼区間）を正しく使い分けることが重要である。信頼区間は、95%といった「confidence level（信頼水準）」をもとに計算される。したがって「信頼水準」は定数であり、計算されるものではない。計算の対象は「信頼区間」である。有病率分析の結果部分（23ページ）、さらに表4と表6を見ると、「信頼区間」の上限と下限の計算値が「信頼水準」とされていない。

信頼区間が95%と計算された場合、下限値は2.5パーセントイル、上限値は97.5パーセントイルとなり、**95%**が信頼水準となる。

従来の統計では、信頼水準はベイズの統計における反復抽出の概念にもとづいて解釈され、信用区間またはベイズの信頼区間は確率区間（probability interval）となる。したがって、従来の設定によると、信頼区間について、「反復抽出で算出された区間に真のパラメータ値が含まれる確率は95%」という解釈が可能となる。また、ベイズの設定によると、「後信用区間にパラメータ値が含まれる確率は95%」である。

したがって、上記の2つの方法はまったく異なる。つまり、BSurvE法は従来の統計、BBC法はベイズの統計の概念を取り入れたものである。

こうした誤りは、「不確実性分析の結果」の章（24ページ）にも見受けられ、「95%の信頼区間」の範囲を示す言葉として「第5パーセントイル」と「第95パーセントイル」が使われている。こうした不正確な用語への変換も訂正しなければならない。また、「5%と95%の信頼水準」という表現も以前に使われているが、これも不正確である。図2と図3中の「第5」と「第95」も間違っている。これらは95%の信頼区間における信頼限界の上限と下限で

ある。それぞれ、2.5パーセンタイル、97.5パーセンタイルと解釈される可能性があるが、これらは信頼区間を示す表現ではない。この場合、一般的な表現は「95%の信頼区間の上限と下限」または「95%のCIの下限」と「95%のCIの上限」である。

図2と図3のプロットの仕方も正確ではない。X-Yプロットの場合、Y軸に対してX軸をプロットするのが一般的だが、図2と図3ではその逆となっている。

## 2) 分析の目標は以下のように明示されている。

米国当局は1990年以降、BSEのサーベイランスを強化している。その一環として、2003年に陽性と判定されたカナダの牛を契機にいくつかのプログラムが強化されている。米国におけるBSEの強化サーベイランス・プログラムは報告書（USDA, APHIS, CEAHによるBSE強化サーベイランス・プログラム、4-6-06）の中で別途要約されている。この分析では、2006年3月17日に至る7年間で収集されたデータにもとづき、米国におけるBSEの有病率が推定されている。

こうした情報は、国際獣疫事務局（OIE）の国際ガイドラインにもとづき、米国におけるBSEの全体状況を検討する上で有効である。

採用されたモデルは分析の目標を適切にとらえている。また、各モデルの長所と短所に加え、それらを既存のデータに適用した結果も提示されている。

3つのモデルの比較対照は不要であり、外挿法の利用は不適當と考えられる。ひとつのモデルが他の2つのモデル（BSurvEとBBC）を比較する上で適當とすれば、そのモデルは最初から採用されているはずである。

**3) 調査目標を達成する上で当該のモデルは十分かつ妥当か？ そのモデルは、有病率の推定結果に影響する重要な要素および実際の状況におけるそれらの相互作用を的確にとらえているか？ そのとらえ方が不十分である場合、欠落しているパラメータ、または誤って既定されているパラメータはなにか？ 検討に値するモデルはほかにあるか？ モデルに関し、現実と乖離している点が明示されているか？**

以上の点を考察すると、BSurvEモデルは十分かつ既存のデータに対して妥当と考えられる。すなわち、有病率の推定に有効なパラメータは網羅されている。感度Pはモデルに含まれているが（陽性、感染）、特異性P（陰性、非感染）は考慮されていないようである。一方、

特異性を1と想定することに問題はないと思われる。

BBCモデルもまた妥当と考えられる。可能性から導き出されたポイント値にもとづく加重方法も加重を算出する上で妥当である。

筆者が取り組んだモデルも代替モデルのひとつである（Suess, Gardner, Johnson, Prev. Vet. 2002を参照）。提案したそのモデルは完全なベイズのモデルである。検討に値するMCシミュレーションに着目したわれわれの論文では他の論文をいくつか参照している。

4) 入力パラメータの前提事項を検討すること。それらは分析目標を達成する上で妥当か？  
また、前提事項はパラメータについての妥当な推定結果を反映しているか？

注：モンテカルロ・モデルは採用されたWin Bugsのソフトウェアの一部であり、選択された分布はWin Bugsのコード（文書の一部）に明示されたポアソン分布を利用している。

#### **BSurvEモデル**

入力に関するBSurvEモデルの前提事項は既存のデータを的確に利用している。一方、疫学モデルでは、データ収集に関する関連要素のすべてが考慮されているようである。年齢グループの情報を得ることはきわめて重要だが、それが含まれている点は妥当である。

BSurvEの表計算は非常に簡潔かつ整理されている。扱いは簡単であり、毎年繰り返して使用可能である。モデルへの入力はもちろん、それらと計算との関連も一目瞭然である。

Excelによるモデルのコーディングは確認していない。

#### **BBC モデル**

BBC モデルの考え方は概念的に妥当である。疫学に明るい人間であれば、このモデルのポイント値が正しいスケールにもとづいて測定されているかどうかを正確に判断できるだろう。論文は「ポイント」の意味を明示しているが、Win Bugs のプログラムの場合、感染率推定の根拠となる値の出典が明示されていない。計算の方法は理解できても、どこで計算されたかが不明である。年間の有病率をダウンウェイトする概念は妥当である。英国の値を利用することも妥当だが、おそらく米国のポイント値はまったく同等でないため、英国の値から事前事項を導き出し、年間のファクターをモデルのパラメータとする方法もある。

厳密にいうと、米国についてはこれらの値が不明であるため、ポイントをパラメータとしてモデルをテストすることは次の段階の作業（すなわち現実をより反映させてモデルを改良する作業）と考えられる。なお、ポアソンモデルの利用は妥当である。

報告書では「 $\text{beta}(0,1)$ 」とされているが、これは「 $\text{Uniform}(0,1)$ 」に相当する「 $\text{beta}(1,1)$ 」とすべきである。

全体として見ると、採用された値が正確であれば、Win Bugsによるシミュレーションは適正である。

シミュレーションにはパラメータがほとんどなく、お互いに相関していないため、パーインは不要と思われる。また、プログラムは速やかに収束する。

**5) モデルと入力データ双方の不確実性を踏まえ、結論自体は確実か？ そうした不確実性はすべて明示されているか？ 結論はどの程度確実か？ それらは明示されているか？**

感度分析をさまざまな面から検討した結果、モデルとデータの双方について結論の確実性は高い。より一般的な用語である「Sensitivity Analysis(感度分析)」の代わりに「Uncertainty Analysis(不確実性分析)」を使うことは不適である。報告書の主題はモデルの変更ではなく、入力値と感染牛の頭数を変え、それがモデルの推定結果に及ぼす影響を見きわめる点にある。いずれにせよ、論文が示すところによると、結果の確実性は高い。

信頼区間を意味する用語に誤りがあるため、結果が明確に伝わっていないが、これを訂正すれば分析の意味は自ずと明確になる。

**6) 採用したモデルを踏まえた上で、分析の結果は適正か？ プログラムを実施してコードとフォーミュラをチェックし、モデルがその目的どおりに機能するかどうかを確認すること。採用されたモデルを踏まえ、USDAが正しい答えを導き出しているかどうかをチェックすること。**

BSurvEとBBCモデルの数値結果から有病率を比較対照できるものと思われる。モデルが疫学的に正しく、重要な変数やデータを網羅しているとすれば、そのモデルは適正な結果を導き出す。

同じく、重要な変数やデータを網羅しているとすれば、全体的な有病率に対してBBCモデ



ルは適正な答えを導き出す。したがって、モデルに残る唯一の問題は、利用しているデータがテスト・データに由来している点である。対象の牛のテスト方法、および臨床的徴候が現われる前にBSE陽性と判定する根拠の詳細は知らないが、非感染牛が誤って陽性と判定される可能性はある。本件は徴候発現テストの特異性と関連している。報告書によると、BSE発症前の牛に対して感度は「0.4」となっているが、特異性についての記述がない。この場合、特異性は「1」と想定できる。これは、対象の牛が健康である特異性を「1」と想定した場合、その牛が陰性と判定される可能性に相当する。

## 望ましい表現とその他の変更点

8ページの「likelihood ratio(尤度比)」に関するの節を見ると、最初の文の「many(多数の)」の意味が曖昧である。テストした牛についてデータがどのように収集されたかを明示すべきである。「many(多数の)」のデータでは、それぞれの牛について全データが収集されたようには受け取られない。

8ページ後半の節では、「sensitivity(感度)」の代わりに「uncertainty(不確実性)」という言葉が使われている。

8ページでは、「BSurvEモデル」が「ダウンロードされたモデル」とされている。この記述は正しいかもしれないが、当該のウェブサイトが明示されていない。ここは単に「BSurvEモデル」とすべきである。

9ページの第2節を見ると、7ページの記述にしたがい、「points(ポイント)」の代わりに「value(値)」という言葉が使われている。混乱を防ぐため、一般的な用語を採用すべきである。「likelihood value(尤度値)」を意味する言葉として「point values(ポイント値)」を採用し、表現を統一すべきである。

9ページ最後の節でベイズの分析が論じられているが、この場合、「データとパラメータに関する事前情報を網羅したモデル」と表現した方が明確である。事前情報は参考のためのものであり、テストに由来する測定データではない。一方、BBCモデルで利用する事前事項は一定であるため、これらについて事前情報は想定されない。

10ページのタイトルについては、2つのモデルが取り上げられ、比較対照のために第3のモデル(推定モデル)も取り上げられているため、「model(モデル)」を「modeling(モデリング)」とすべきである。

10ページの第1節の表現は、前述のとおり、好ましくない。

10ページの「sensitivity(感度)」に関する記述を見ると、BSE発症前の牛に関する「sensitivity」が取り上げられていない。「sensitivity」はひとつと想定されているが、その妥当性を明示すべきである。

11ページのサンプル分布の影響に関する記述については、まずタイトルを「sampling(サンプリング)」とすべきである。節の内容は「サンプルの分布」よりも「サンプリング方法」に重点を置いている。目標のサンプリングに関する記述はあるが、分布の形態は論じられていない。また、「双方の方法において」の意味を明示する必要がある。この場合、「BSurvEにおけるA法とB法」という表現が好ましい。

11ページのBBCの事前分布に関する節では、文頭の「only(唯一)」が不適である。この言葉を削除しても意味は変わらず、むしろ文がすっきりする。また、最後の文のカンマで区切られたあとの「it(それ)」の意味が曖昧である。「starting value(初期値)」が「prior distribution(事前分布)」を示す「it(それ)」によって誤解されているようである。シミュレーションのもととなるのはWing Bugsのプログラムにおける「starting value」だが、データが優先する対象はPにおける「prior distribution」である。

12ページの「value(値)」は可能性の「point values(ポイント値)」と明示すべきである。

12ページの第4節の最後に「the 95% Confidence Interval(95%の信頼区間)」という表現がはじめて登場するが、これは誤りである。「95th percentile confidence level(第95パーセンタイルの信頼水準)」という表現は正確でない。信頼区間は特定の信頼水準について計算される。したがって、CIが95%と算定されれば、その信頼水準は95%に固定される。このパーセンテージは当該の分布の95%である。CIの計算のもととなるのはこれらの分布の2.5パーセンタイルおよび97.5パーセンタイルである。したがって、いわゆる「95%CIの下限值」と「95%CIの上限值」がある。論文では「パーセンタイル」という言葉が誤用されている。CIを意味する用語の正確さを期すため、これらをすべて訂正する必要がある。

12ページ最後の節では、抽出したサンプルに対して「random(ランダム)」という言葉が使われている。一方、論文の前半部分では「targeted sampling(対象を絞ったサンプリング)」が論じられているが、当該のサンプリングについて「targeted random sampling(対象を絞ったランダム・サンプリング)」という表現は使われていない。この点を明確にする必要がある。モデルが前提としているのは「targeted random sampling」と思われるが、実際にそうで

あるとすれば、その旨明示すべきである。なお、時間の関係でデータ・サンプリングの方法は確かめていない。

13ページでは、すでに述べたように、BSurvEモデルの説明に留意すべきである。曖昧な表現では同モデルの内容が正確に伝わらない。正確さを期するため、BSurvEモデルのマニュアルを利用すべきである。

13ページのBBCモデルに関する記述を見ると、「Bayesian mathematical methods (ベイズの数学的方法)」とあるが、「Bayesian statistical methods (ベイズの統計的方法)」とした方が正確である。

13ページの以下の節も修正を要する。

第1行 「knowledge about ( ~に関する情報 )」を「knowledge of ( ~の情報 )」に変更

第4行 「empiric」を「empirical ( 経験による )」に修正。「empiric」は名詞。ここでは形容詞を使うべきである。

第6行 「有病率は飼料禁輸前に増加し、禁輸後に減少」

第7行 「ref.」のあとに「and」が必要。

14ページの「The Bayesian Birth Cohort ( Bayesian出生コホート )」ではじまる節の「epidemiologic ( 疫学 )」は「epidemiological ( 疫学の )」とすべきである。事前事項が一定である点を説明するに当たり、「starting point ( 起点 )」という言葉が使われているが、これが可能性 ( likelihood ) の「point value ( ポイント値 )」を意味する「point」と混同されている。また、次の節に「from this point forward ( この点より先 )」という表現があるが、報告書における「point」には専門的な意味があるため、正確さを期すべく、その意味を持つ「point」を「point value ( ポイント値 )」と明示するか、または一般的な意味での「point」の使用を排除することが望ましい。

BBCモデルにおける有病率は確率であり、0から1の間の数値で示すべきものである。パーセント値ではない。

15ページの図1を見ると、「epidemic curve ( 流行曲線 )」という言葉を使ってプロットが説明されている。これは疫学の専門用語だろうか？また、「cumulative ( 累積 )」プロットとされているが、これは正しくない。パーセンテージを表す棒があるのは棒グラフである。これは通常ヒストグラム的一种であり、累積ヒストグラムではない。したがって、「cumulative ( 累積 )」という表現は正しくない。

15ページの「infectivity (伝染力)」という言葉は妥当だろうか？ここでの主題は「rate of infection (伝染率)」と考えられる。また、後に続く「dose of infectivity (伝染力の程度)」という表現からは、「rate (率)」がある程度抑制されているような印象を受ける。

15ページ最後の節の第3行にある「禁輸措置以降の暴露も英国の場合より低い」上記のように米国での使用は、英国にも一貫性をもたらすであろう。**The use of U.S. above would lead to U.K. also for consistency.**

16ページの図1では、パーセンテージについて「cumulative (累積)」という言葉が使われている。

16ページの表1の表題を見ると、表中には「proportion (割合)」とあるのに「percent」が使われている。また、星印を付した脚注があるが、これに相当する箇所が見当たらない。

16ページ最後の節の「p」はモデルの「P」であり、一定の密度(0,1)、すなわちbeta(1,1)を持つものと考えられる。

17ページのWin Bugsのソフトウェアについては報告書をとおして常に参照する必要がある。

17ページの第1節を見ると、ポアソンの利用が論じられている。「due to the large sample (多量のサンプルにより)」という理由は、「due to the large sample and the small prevalence (多量のサンプルと低い有病率により)」とするべきである。サンプルの量が多く、確率が低い場合、ポアソンのモデルは2項式モデルに近くなる。

17ページの第1節の最後に、「one BSE point value is equivalent to one randomly sampled animal (BSEポイント値(1)は任意に抽出された1頭の牛に相当する)」という記述がある。モデルを正しく解釈すると、「one randomly sampled animal that has BSE (任意抽出された1頭のBSE感染牛)」という一文を加えるべきだが、100%の確証はないため、チェックを要する。

17ページの「uncertainty(不確実性)」は一般に「sensitivity(感度)」のことである。「sensitivity」という用語は、データや入力パラメータの変化に対するモデルの堅牢性を確かめる際に使われるが、ここでは「uncertainty」の意味が曖昧である。この用語が以前に使用された例は見当たらない。ベイズの文献では「sensitivity」が使われている。また、「sensitivity」は統計学でも一般的な用語である。

17ページ最後の節では、「probability(確率)」または「chance(機会)」の代わりに「likelihood(可能性)」が使われている。「likelihood(可能性)」は「likelihood function(尤度関数)」のほか、BSEモデルでポイント値を論じる際に使われる。混乱を避けるため、ここでは「probability(確率)」を使うことが望ましい。

18ページではBSurvEモデルがアルゴリズムとされているが、これは正しくない。同モデルは表計算モデルである。

最初の節を見ると、BBCモデルの中に「surveillance points」として「points」が新たに登場する。「surveillance points」に統一すべきかもしれないが、用語の変更は混乱を招くおそれがある。

「mechanics」の意味も曖昧であり、文全体の意味が不明である。データまたは事前情報の変化がBSEモデルによる有病率の推定結果にどの程度影響するかを感度分析で調べる方法もある。ほかのモデルを使う場合は感度分析にならない。むしろそれはモデルの比較対照である。この章のタイトルも不適當である。ここで論じられているのは感度分析ではなく、モデルの比較対照である。

18ページの第3節の最後でも「probability(確率)」の代わりに「likelihood(可能性)」が使われている。

19ページの最初の節(ベータ版の適合)には、「the advantage of matching moments is that it provides an unbiased estimate of prevalence(積率の利点は有病率を公平に推定できること)」とあるが、これはおそらく間違いである。密度を適合させればパラメトリック密度を推定できるため、有病率自体ではなく、有病率の分布を公平に推定できるかもしれない。積率法をベータ版に使用すれば公平な推定ができるため、有病率も公平に推定される可能性が高い。結局のところ、密度を適合されることで推定値が得られることはない。

20ページでは、推定が行われる前に「true(真の)」という言葉が使われている。「true」は「true population prevalence(真の個体群有病率)」のように使われる場合もある。推定とは真の個体群有病率の推定を意味する。推定量には分布があるため、推定量を表す際に「true」は一般に使われない。

21ページ最後の節を見ると、「probabilities(確率)」のあとに「proportions(割合)」が続

いているが、意味が曖昧になるため、これは不要である。報告書を通して「probabilities」で統一し、「proportions」を省いて「likelihood (可能性)」を残すことが望ましい(ポイント値の計算方法を定義する「likelihood function (尤度関数)」)

22ページの「download (ダウンロード)」は削除すべきである。

22ページの第2節に「large coefficient suggests inputs are... (大きな係数は入力データが~であることを示している)」とあるが、感度分析の場合、入力データ面で検討されているのはモデルである。したがって、モデルは入力データに反応する(その逆ではない)。この一文の意味を明示すべきである。

22ページでは、モデルがデータに合わされている。一般的にデータがモデルに合わされるようには書かれていない。したがって、ガンマ分布に関する文には「BSens fits the cattle age distribution with a gamma model (ガンマ・モデルによってBSensが牛の齢分布に適合する)」とある。次ページの冒頭の「to」は「by」にすべきである。

23ページの結果の章を見ると、前述のとおり、信頼区間についての記述が不正確で紛らわしい。最初の一文は、「the expected prevalence and the 95% confidence intervals are shown in Table 4 (予想される有病率と95%の信頼区間を表4に示し)」のように書くべきだろう。節の最後の文も同様である。

24ページの表4については、「5 and 95 percent confidence levels (5%と95%の信頼水準)」を「95% confidence intervals (95%の信頼区間)」とすべきである。

「level (水準)」という言葉も報告書では一切使わないことが望ましい。「信頼区間はすべて95%」と明示されれば「level (水準)」が定まるため、それ以上の記述はない。「level (水準)」は「upper and lower limits (上限・下限値)」とすべきである。

24ページ最後の節の「5th and 95th (第5および第95)」という記述は不正確である。「the 95% confidence interval (95%の信頼区間)」とすべきである。

25ページにも信頼区間について同様の誤りがある。適切に表示するため、図においてマニュアルを参照すべきである。「5th and 95th (第5および第95)」も不正確かつ曖昧。「the upper and lower 95% CI limits (95%のCIの上限・下限値)」とすべきである。

25ページの図2と26ページの図3の「versus (対)」は順番を取り違えている。本来なら「y versus x (y対x)」とすべきである。

27ページの図4は「distribution (分布)」という言葉で始まっているが、これは「distribution」ではなく「likelihood ratio values (尤度比の値)」に関するプロットである。また、xのスケールが一定でないため、非常に紛らわしい。正しいスケールで表示すればプロットはまったく違って見えるであろう。x軸のスケールを一定にするか、または対数スケールを使うべきである。

第1節の最後の「unremarkable (目立たない)」は、「a sensitivity of an estimator from a model (モデルによる推定量の感度)」に対して一般的な形容詞ではない。「was low (低かった)」とする方が適切である。

最後の節の「confidence (信頼)」は「a high degree of certainty that (高い確実性)」とした方が妥当かもしれない。統計学では、「confidence」は「confidence interval (信頼区間)」に使われるため、この言葉を使いまわすと混乱を招く。

28ページも同様に、「sensitivity (感度)」に対して「unremarkable」を「low」とすべきである。また、最後の節の「insensitive (感度の低い)」も「low sensitivity (低感度)」とし、表現の統一を図るべきである。

29ページの第1節の文末にピリオドが2つ打たれている。これはひとつでよい。

29ページの表5が分断されている。脚注にあるモデルのインデックス(j と t)はイタリック体で表示すべきである。報告書においてモデルを参照する場合は常にイタリック体を使用すべきである。

30ページの節の最後にある「94パーセント」の対象が曖昧である。

32ページの「5 and 95 percent confidence levels (5%および95%の信頼水準)」を「95%の信頼区間」とすべきである。

33ページの第3節の「work sheet」は報告書の前半部分で「worksheet」と1語になっている。

第5節の「estimations」は「estimates」とすべきである。

また、「uncertainty analysis（不確実性分析）」は「sensitivity analysis（感度分析）」とすることが望ましい。

不確実性分析のガイドラインに関するEPAの論文がリンク先に挙げられているが、ベイズの統計に関わる立場から言わせてもらおうと、これは「sensitivity analysis（感度分析）」とする方が一般的である。

<http://risk.lsd.ornl.gov/homepage/tm35r1.pdf>

34ページの信頼区間に関する記述も訂正が必要である。

35ページの「Exp」は「exp.」とすべきである（function exp.=関数式）。「Equation Editor」を使って正しく表示する方法もある。また、モデルを参照する際はイタリック体を使って表示を統一すべきである。

「データはWeibull（ワイブル分布）によって適合した」

「number of 1-year olds, minus the probability of culling for the first year heifers（1歳牛の頭数から1歳の若雌牛が淘汰される可能性を引いたもの）」という一文があるが、後半部分は「minus the expected number（予想頭数を引いたもの）」とすべきかもしれない。

37ページの「BSurvE mechanics of point calculation（ポイント算定のBSurvEメカニクス）」は「BSurvE point estimation（BSurvEによるポイントの推定）」とすべきである。