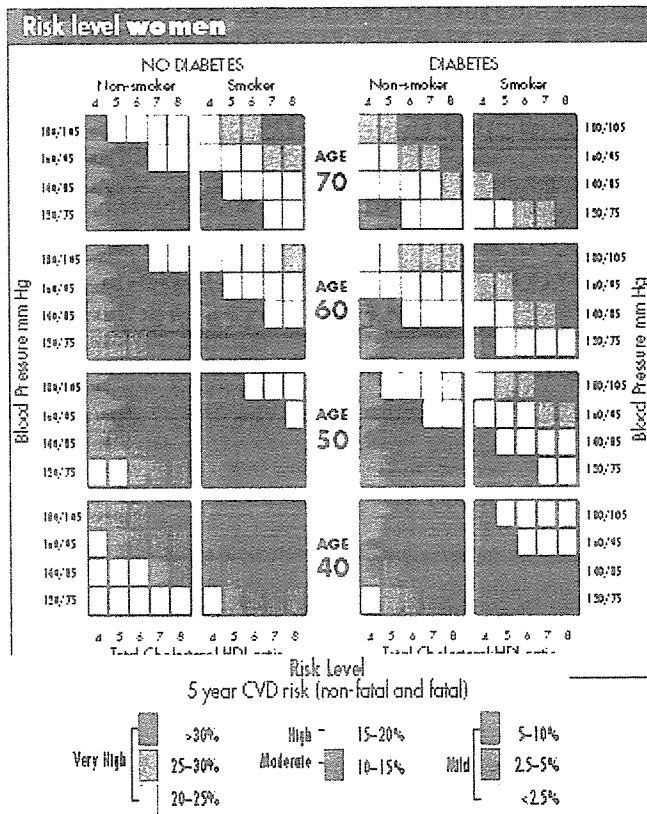


のように共有していくべきか、開かれたディスカッションを積み重ねていく必要がある。診療ガイドラインに関する世界的な動きとして2002年にはWHOを含む24カ国46組織よりなる”Guideline International Net (GIN)”が発足し、2004年には推奨度の決定に関するGRADEの提案などが発表された(16)。ガイドラインへのアクセスビリティを高めるため国内では財団法人日本医療機能評価機構の「医療情報サービス事業」、通称・”Minds” (Medical Information Network Distribution Service)の整備が進められている(図3)(17)。

医療における患者と医師の関係の変化は加速されるであろう。EBM、そして診療ガイドラインを通じて、医療は社会に対するそのアカウンタビリティを問われている(18)。EBMの基盤としての疫学研究の整備とその成果の診療ガイドラインへの連携のあり方をさらに追求していく必要がある。

図1 リスクファクター別5年間の循環器疾患リスク（絶対リスク）の一覧表



Your doctor, nurse or health professional will assess what level of risk you have, which will be shown as a percentage. For example, if your risk is high, e.g. 20%, and there were 100 people like you (without treatment), then approximately 20 of these people would have a heart attack or stroke in the next

図2. Minds のホームページ（財団法人・日本医療機能評価機構 医療情報サービス事業）

Minds 医療情報サービス
厚生労働科学研究費補助金により試験公開中

Mindsについて
Mindsの使い方
医療提供者について

「診療ガイドライン(医療提供者向け・一般向け)」をご覧になるにはこの下のボタンをクリック!!

Mindsユーザー
MindsユーザーのSTART

ゲスト
MindsユーザーのSTART

Mindsユーザーになるには

「Mindsユーザー」「ゲスト」ボタンをクリックしても反応しない場合はお読みください

お知らせ
「クニ下出前」の「Mindsアプリ」を公開しました
2005.11.06 開催「第4回LUM研究フォーラム」のお知らせ
「Mindsユーザー」「ゲスト」ボタンをクリックしても反応しない場合はお読みください

利用条件
Mindsをご利用になる際には、上記の利用条件についてご確認ください。

表 1. 推奨度の決め方（以下の要素を勘案して総合的に判断）

1. エビデンスのレベル
2. エビデンスの数と結論のバラツキ
（同じ結論のエビデンスが多ければ多いほど、そして結論のバラツキが小さければ小さいほど勧告は強いものとなる。必要に応じてメタアナリシスを行う）
3. 臨床的有効性の大きさ
4. 臨床上の適用性
5. 害やコストに関するエビデンス

表 2. 推奨度の形式

1. 米国 Agency for Health Care Policy and Research (1994 年)

- A : 行うよう強く勧められる
- B : 行うよう勧められる
- C : 行うよう勧めるだけの根拠が明確でない
- D : 行わないよう勧められる

2. 脳卒中合同ガイドライン委員会 (2004 年)

- A : 行うよう強く勧められる（少なくとも 1 つのレベル I の結果）
- B : 行うよう勧められる（少なくとも 1 つのレベル II の結果）
- C1 : 行うことを考慮してもよいが、十分な科学的根拠はない
- C2 : 科学的根拠がないので、勧められない
- D : 行わないよう勧められる

3. 米国 臨床予防医学タスクフォース (2001 年)

- A : 臨床医は、対象となる患者に定期的に介入（ここでは予防サービス。以下介入とする）を提供することを強く推奨する。[USPSTF は、介入が重要な健康アウトカムを改善するとする強い科学的根拠を得て、利益は害よりはるかに大きいとの結論を下している。
- B : 臨床医は定期的に介入を対象となっている患者に提供しよう推奨する。
（USPSTF は少なくとも、介入が重要な健康アウトカムを改善するとする十分な科学的根拠を得て、利益は害より大きいとの結論を下している。
- C : 定期的な介入の提供についての是非に対する推奨は行わない。（USPSTF は、少なくとも介入が健康アウトカムを改善するとするかなりの科学的根拠を得たが、利益と

害のバランスが、一般的な勧告を正当化するにはあまりにも接近しているとの結論を下している。)

D：介入を定期的に非症候患者に提供しないよう推奨する。(USPSTFは少なくとも、介入が効果的でない、あるいは害が利益より大きいとするかなりの科学的根拠を得た。

I：介入を定期的に提供することの是非についての科学的根拠が不十分であるとする結論を下す。(介入が効果的であるとする科学的根拠の欠如、その質の低さや矛盾、また利益と害のバランスが決定できない。)

4. カナダ 予防医学に関するタスクフォース (2003年)

A：予防的介入を推奨し得る十分な (Good) エビデンスがある

B：予防的介入を推奨し得る適正な (Fair) エビデンスがある。

C：既存のエビデンスは一致しておらず、予防的介入を肯定・否定する推奨は決定できない。しかし、他の要因が意思決定に影響し得るであろう。

D：予防的介入を行なわないことを推奨する適正な (Fair) エビデンスがある。

E：予防的介入を行なわないことを推奨する十分な (Good) エビデンスがある。

I：推奨を決定するにはエビデンスが量的・質的に不十分である。しかし、他の要因が意思決定に影響し得るであろう。

表3 エビデンスが明確でない場合の意思決定の基準 (カナダ・タスクフォース)

- 意思決定における患者自身の参加を促すこと
- 害を最小化すること
- 強い必要性が明らかな場合に関してのみ、大きな変化を主張すること
- 不要な「ラベリング」を避けること
- 益の不確かな高価な手技を避けること
- 疾病負担が大きい状況に焦点を当てること
- ハイリスクグループの特別のニーズに配慮すること

表4 診療ガイドライン作成・報告のチェックポイント (COGS と AGREE の対応)

COGSの項目	説明	対応するAGREE評価ポイント
1. 要約 (Overview material)	ガイドラインのリリース日、ステータス(初版、改訂版、最新版)、活字情報及び電子情報の出典などの情報を含む構造化抄録を提出すること。	
2. 焦点(Focus)	ガイドラインの対象となる基本的疾患・健康状態や介入・サービス・技術に関する記載を加えること。ガイドラインで採択された以外の予防上、診断上、治療上の介入で、ガイドライン作成段階で検討にのぼった介入についても明示すること。	2. ガイドラインが取り扱う臨床の課題が明確に記載されている。 16. 管理に関する異なるオプションが明確に示されている。
3. 目標 (Goal)	ガイドライン遵守によって達成されるべき目標を記載すること。なぜこのトピックに関するガイドラインを作成するのかについての理由説明も加えること。	1. ガイドライン全体の目的が明確に記載されている。
4. 利用者・利用場面 (Users/setting)	ガイドラインの利用対象者(医療提供者の種類、患者)及びガイドラインが利用されるべき場面について記載すること	6. そのガイドラインの対象となる利用者が明確に示されている。
5. 対象集団 (Target population)	ガイドライン推奨の対象となる対象患者集団について記載し、除外基準があれば全てこれを記載すること。	3. ガイドラインの対象として想定された患者が明確に記載されている。
6. 作成者 (Developer)	ガイドライン作成に責任を持つ組織、及びガイドライン作成に携わる人員の氏名、資格、利害関係の衝突の可能性を明示すること。	4. ガイドライン開発グループの個人個人は、該当する全ての専門家集団からなっている。 23. ガイドライン開発メンバーの利害関係が記録されていた。
7. 資金提供者・スポンサー (Funding source/sponsor)	資金提供者・スポンサーを明示し、ガイドライン作成・報告における役割を記載すること。利害関係の衝突が考えられる場合は、これを開示すること。	22. ガイドラインは編集に関してその資金提供元から独立している。 23. ガイドライン開発メンバーの利害関係が記録されていた。
8. エビデンスの収集 (Evidence collection)	学術文献を検索するのに用いた手法を記載し、検索の期間、検索したデータベース、そして収集した文献を絞り込むのに用いた基準についての記載も含めること。	8. 根拠の検索には系統だった手法が用いられた。 9. 根拠の選択基準が明確に記載されている。
9. 推奨度決定基準 (Recommendation grading criteria)	推奨の根拠となるエビデンスの質を評価するのに用いた基準、そして推奨度を記述するためのシステムについて記載すること。推奨度は、推奨への遵守がどれだけ重要かを示すものであり、エビデンスの質、そして予測される利益と害の程度に基づいて評価される。	10. 推奨を導き出した方法は明確に記載されている。 11. 推奨を導き出すにあたって、健康上の利点、副作用、危険性が考慮された。
10. エビデンスの統合のための手法 (Method for synthesizing evidence)	推奨作成に際してエビデンスをどう用いたかを記載すること。たとえば、エビデンスの表、メタ分析、判断分析など。	
11. リリース前のレビュー (Prerelease review)	リリース前にガイドライン作成者がどのようにガイドラインをレビューし、及び/又は、審査したかについて記載すること。	13. そのガイドラインは出版前に専門家による外部評価を受けた。

12. 更新の計画 (Update plan)	ガイドライン更新の計画があるか否かを記し、適用可能な場合は現在出ている版の有効期限を記すこと。	14. ガイドラインの更新手順は示されている。
13. 定義 (Definitions)	一般的でない用語を定義すること。また、間違っ て解釈される可能性のあるガイドライン については、これが正しく適用されるた めに確実に理解されなければならない用語 を定義すること。	
14. 推奨及び理由説明 (Recommendations and rationale)	推奨される対処法及びこれが実行されるべ き具体的な状況を正確に記すこと。推奨と、 推奨の根拠となるエビデンスとの関連性を 説明することにより、各推奨の正当性を示す こと。9に記載されている基準に基づき、エ ビデンスの質、推奨の推奨度を示すこと。	12. その推奨とその基となった根 拠との関連が明確である。 15. その推奨は具体的で、あいま いさのないものである。 17. 容易に重要な推奨を見分けら れる。
15. 考えられる利益と 害 (Potential benefits and harms)	ガイドラインの推奨を実施することによっ て予測される利益、そして考えられる害につ いて記載すること。	20. その推奨の適用に伴う費用に 関して考慮された。
16. 患者の希望 (Patient preferences)	推奨が患者の選択や価値観に大きく関わる ものであった場合の、患者の希望の扱い方 について記載すること。	5. 患者の視点や選好は考慮され た。
17. アルゴリズム (algorithm)	適切であれば、ガイドラインで取り上げられ ている診療における段階や意思決定を図表 に表し提供すること。	18. ガイドラインの実施手段が示 されている。
18. 実施における検討 事項 (Implementation considerations)	推奨の適用において予測される障壁につい て記載すること。医療提供者と患者のため に、推奨実施を円滑にするような関連文書に ついては全て参考指示すること。ガイドライ ン実施による医療内容の変化を計測するた めの評価基準を提案すること。	18. ガイドラインの実施手段が示 されている。 19. 推奨を適用するにあたって考 えられる組織の障害が検討され た。 20. その推奨の適用に伴う費用に 関して考慮された。 21. ガイドラインには監視・監査 のための主要な評価基準が示され ている。
		7. ガイドラインはエンドユーザ ー利用者によって試験的に用いら れた。

文献

1. Institute of Medicine. Guidelines for clinical practice: from development to use. Washington DC, National Academy Press, 1992.
2. Nakayama T, Budgell B, Tsutani K. Confusion about the concept of clinical practice guidelines in Japan: On the way to a social consensus. *Int J Qual Health Care* 15: 359-360, 2003.
3. 平成10年度医療技術評価推進検討会報告書. 厚生省健康政策局研究開発振興課医療技術情報推進室(監修). わかりやすいEBM講座. 厚生科学研究所, 東京, 8-28, 2000
4. 中山健夫. EBMを用いた診療ガイドライン:作成・活用ガイド. 金原出版, 東京, 78-80, 2004
5. 福井次矢、丹後俊郎. 診療ガイドラインの作成の手順. *EBMジャーナル* 4:28-37, 2003.
6. 棚橋紀夫. EBMに基づく治療ガイドラインの作成と応用の問題点. 第44回日本神経学会総会プログラム・抄録集, 61, 2003
7. 祖父江友孝、濱島ちさと、齊藤博、中山健夫. 有効性評価に基づくがん検診ガイドライン作成手順. *癌と化学療法*. 2005;32:893-900.
8. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet* 342: 1317-1322, 1993.
9. Koyama H, Fukui T: A review of research on the clinical effectiveness of therapeutic practice guidelines, 1991 - 2000: the need for standardization of reporting style. *General Medicine* 1:1-7, 2002
10. 福井次矢、中山健夫. 日本国内で作成された診療ガイドラインの質評価の試み. 厚生労働科学 EBMを指向した「診療ガイドライン」と医学データベースに利用される「構造化抄録」作成の方法論の開発とそれらの受容性に関する研究(主任研究者・中山健夫) 2001年度報告書, p41-45
11. Appraisal of Guidelines, Research, and Evaluation in Europe (AGREE) Collaborative Group. Guideline development in Europe: An international comparison. *Int J Technol Assess Health Care* 16: 1039-1049, 2000.
12. 長谷川友紀. 診療ガイドラインを取り巻く状況: AGREE Collaborationの動向. *EBMジャーナル* 4:294-7, 2003
13. Shiffman RN, Shekelle P, Overhage JM, Slutsky J, Grimshaw J, Deshpande AM. Standardized reporting of clinical practice guidelines: a proposal from the Conference on Guideline Standardization. *Ann Intern Med* 139:493-8, 2003.
14. 中山健夫. EBMを用いたガイドラインの作成と応用: EBMの考え方と問題点を踏まえて. *Progress in Medicine* 23:3143-51, 2003

15. Nakayama T, Fukui T, Fukuhara S, Tsutani K, Yamazaki S. Comparison between impact factors and citations in evidence-based practice guidelines. JAMA. 2003;290:755-6.
16. GRADE Working Group. Grading quality of evidence and strength of recommendations. BMJ. 2004;328:1490.
17. Satoh T, Nakayama T, Sato Y, et al. Physicians' awareness regarding evidence-based medicine, practice guidelines and clinical information resources in Japan: Needs assessment prior to the initiation of "Medical Information Network Distribution Service (Minds)". General Medicine. 2004;5:13-20
18. Nomura H, Nakayama T. The Japanese healthcare system. BMJ. 2005; 24:648-9

* 発表

1. Nakayama T, Sakai M, Slingsby BT. Japan's ethical guidelines for epidemiologic research: A history of their development. J Epidemiol 2005;15:107-12
2. Nomura H, Nakayama T. The Japanese healthcare system: The issue is to solve the "tragedy of the commons" without making another. BMJ 2005;331:648-9.
3. Nakayama T. Evidence-based healthcare and health informatics: Derivations and extension of epidemiology. Epidemiol (in press)
4. 内藤真理子, 中山健夫. 健康政策への応用. 青山英康 (監修), 川上憲人, 甲田茂樹編. 今日の疫学 第2版. 東京: 医学書院, 2005
5. 中山健夫 (分担翻訳). ミュア・グレイ著, 根拠に基づく保健医療 (津谷喜一郎・高原亮治 監訳). 東京: エルセビア・ジャパン, 2005

1.1. 複数のコホートデータの統合に関する統計学的手法について —JALS の実践を通して—

分担研究者

大橋 靖男 東京大学大学院医学系研究科 生物統計学/疫学・予防保健学・教授

研究協力者

吉田 征太郎 東京大学大学院医学系研究科 生物統計学/疫学・予防保健学・大学院生

1. はじめに

疫学研究は地域などからコホートを構成し、多くの参加者を募る観察型の研究である。これら前向きな疫学研究では、得られる結論の一般化可能性を高めるために様々な地域を研究対象集団に含めることがある。一般に、多数の地域を対象としたデータにおいては、地域（コホート）ごとに性・年齢などの背景因子が異なるだけでなく、統計解析では明示的に考慮することが困難な気候、風土、生活習慣などの環境要因も異なることが予想される。したがって、そのような疫学研究においては、地域ごとにリスク因子の効果を評価すべきであり、観察されたリスク因子の効果の大きさが地域間で一様かどうか、すなわち、得られた結論は地域の選択によらず一般化することが可能かについて「得られた結論の地域選択に関する一般化可能性」を検討しておくことが重要であるといえる。

そこで、今回は近年計算機の性能の向上とともに適用することが可能になったマルコフ連鎖モンテカルロ法を用いたコホートデータ統合に関する方法論を述べ、日本動脈硬化縦断研究（Japan Arteriosclerosis Longitudinal Study；以下、JALS）に適用した結果を報告する。

1.1 地域を考慮した解析

上記と同様の議論が、多施設臨床試験における施設の取り扱いにおいてなされている。1991年から活動が開始された日米 EU 医薬品規制調和会議（International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use; 以下、ICH）の成果のひとつである「臨床試験のための統計的原則」を記載したガイドライン（ICH-E9）では、多施設臨床試験が実施される理由のひとつとして、「得られた結果を一般化するためのより適切な根拠を与えるため」を挙げている（厚生省医薬安全局,1998）。観察された治療効果の差が施設間で一様であれば、すなわち治療効果に関する施設間差が小さければ、得られた結論の施設選択に関する一般化可能性は高いと推察されるが、施設によって治療効果が異なり、不安定な結果となる場合には、多施設試験全体としての結果の信憑性に疑問がもたれることになる。

疫学研究において観察されたリスク因子の効果が、地域（コホート）ごとに同じような傾向を示すかどうかの検討は、統計的には、リスク因子と地域との交互作用の評価に相当し、大別して次の3つのアプローチが考えられる。

1. 地域を層として考慮した層別（あるいは層化調整）解析
2. 地域を表す変数を固定効果とした回帰型モデルによる解析
3. 地域を変量効果として考慮したモデルによる解析

地域を1つの「層」とみなせば、疫学研究で広く用いられている「層別解析」の考え方が基本的には適用可能である(Rothman and Greenland,1998)。地域ごとにリスク因子の効果の大きさを求め、それらを重み付けて併合し、同時に効果の共通性（交互作用）の検定やグラフ表示などを行なうことにより、そのリスク因子の効果に関する地域間差の有無やその程度を慎重に検討することになる。

対象地域数が例えば5地域以下というような少数であり、各地域において十分な数のイベントが観察されていれば地域を固定効果とした「回帰型モデルによる解析」を適用することが可能である。しかし、これらの方法が適用できるほど十分な情報が得られない場合、地域の層別や固定効果モデルを用いた方法に代わる手法が必要となり、実際、多くの疫学研究ではこのような解析ができるほど十分な情報量をもったデータが得られることは稀である。イベント数に比べ地域の数が増えていくと最尤推定量の一致性が保証されなくなる。このような場合には、地域を変量効果として考慮した混合効果モデル(mixed effects model)による解析が方法として適切である (Diggle and Zeger,1994)。

固定効果の他に地域を変量効果とみなしモデルに含む解析モデルは、混合効果モデルと呼ばれ、これまで多施設臨床試験における施設間差の検討においてその適用例がいくつか見られる (Gray 1994, Matsuyama et al. 1998, Yamaguchi et al. 1999)。しかし、イベント数の少ない疫学研究データにそのようなモデルを適用し、地域間差を検討している研究は少なくともわが国には存在せず、欧米における大規模疫学研究においてもほとんど見られない。

1.2 地域を変量効果とした解析の問題点

そのような検討を困難にしている統計解析上の問題として、そのモデルに対するパラメータ推定方法が確立されていないことが挙げられる。ベースラインにおける地域効果だけでなく、地域とリスク因子の交互作用も変量効果と仮定しモデルに含めた場合には、そのモデルパラメータの推定に必要な尤度関数を明示的に表現することができず、何らかの近似計算法を用いざるを得ない (Breslow and Clayton 1993, Davidian and Giltinan 1995)。統計解析パッケージ SAS においても、NLMIXED プロシジャの中にいくつかの近似推定法が用意されているが、健診データに基づく疫学研究データのようにイベント数が少ないデ

ータや変量効果のバラツキが小さなデータ、すなわち地域間差が小さいデータに対しては、パラメータ推定のための最適化計算が収束しない可能性が高くなることが知られている (Wolfinger and Lin, 1997)。変量効果モデルとベイズモデルの類似性に加え、非常に多くのパラメータを含んだ複雑なモデルに対しては一般にベイズ推定法が適していることなどから、近年一般化線型混合モデルのパラメータ推定においてもベイズ流の推定法が提案されている (Zeger and Karim, 1991)。

そこで2節においてベイズ流の推定法の1つであり近年の計算技術の発展とともに使用されるようになってきたマルコフ連鎖モンテカルロ法に基づく Gibbs sampling (Gilks et al. 1995) によるパラメータ推定法を紹介し、3節において実際に本邦における大規模疫学研究である JALSO 次研究データ (大橋ら 2003) に適用した結果を提示する。解析例に即して以後イベントを脳卒中発症、関心のあるリスク因子を喫煙とする。

2. 地域間差を考慮したモデル

2.1 ポアソン混合効果モデル

地域間差を考慮した解析を行なうために、地域 i ($i=1, \dots, I$) における対象者 j ($j=1, \dots, n_i$) に対する脳卒中発症の有無を表す変数を y_{ij} 、追跡人年数を t_{ij} 、喫煙の有無を表す変数を x_{ij} 、喫煙以外の年齢などのリスク因子をまとめて表した変数を z_{ij} とすると、エンドポイントにポアソン分布を仮定したポアソン混合効果モデルは以下のように表現できる。

$$\log E(y_{ij} | b_i) = \log t_{ij} + \beta_0 + \beta_1 x_{ij} + \gamma z_{ij} + b_{0i} + b_{1i} x_{ij} \quad (1)$$

ただし、(1)式において、 β_1 と γ はそれぞれ変数 x_{ij} 、 z_{ij} に対する固定効果パラメータであり、 b_{0i} と b_{1i} は地域 i の効果を表す変量効果パラメータである。ここで、変量効果は互いに独立に平均 0 の二変量正規分布に従うと仮定する。

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \underset{iid}{\sim} N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} \right) \quad (2)$$

稀なイベントの発生数はポアソン分布で近似できることが知られており、疫学研究データにおいては結果変数にポアソン分布を仮定することが一般的である。

(1)式で、変量効果 b_{0i} は脳卒中発症に対する地域 i の全体平均からのズレ (ベースライン地域間差) を表しており、 b_{1i} が脳卒中発症に対する喫煙の平均的な効果 β_1 からの地域 i のズレ (地域と喫煙の交互作用効果) を表す地域間差パラメータである。このモデルにおいて関心のあるパラメータは、固定効果 β_1 と γ だけでなく、地域効果を表すパラメータ b_{0i} と b_{1i} 、およびそれらのバラツキ (地域間差の大きさ) を表す変量効果に対する分散パラメータ d_{kl} (k, l は分散の k, l 成分を表す) である。分散パラメータの値が大きければその地域効果

はバラつきが大きいと判断することができ、小さいときは地域効果のバラつきはほとんどない、すなわち地域によってそのリスク因子の効果は異ならないと考えることが出来る。

2.2 パラメータ推定方法

(1)式のモデルパラメータを推定する際に必要となる尤度関数は(3)式で表される。

$$\prod_{i=1}^I \prod_{j=1}^{n_i} f(y_{ij} | b_i) |D|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} b_i' D^{-1} b_i\right) db_i \quad (3)$$

ここで関数 $f(y_{ij}|b_i)$ は、変量効果を与えた下で平均値が(1)式で規定されるポアソン尤度関数、 D は(2)式の分散共分散行列である。(3)式の尤度関数に含まれる積分計算は、誤差に正規分布を仮定した線形モデルを除いては解析的に解くことができず、未知パラメータ推定のための目的関数を明示的に表現することが出来ないため、この積分に対処するための何らかのアプローチが必要となる。

2.2.1 尤度近似によるアプローチ

(3)式の尤度関数に含まれる積分計算をある種の近似を用い近似計算で対処する方法がいくつか提案されている。Laplace 近似に基づく Lindstrom and Bates 法(Lindstrom and Bates, 1990)や、求積法の一つであるニュートン・コーツ公式を改良した適応的ガウス求積法(Pinheiro and Bates, 1995)などがその例として挙げられる。一般に、これらの尤度近似を用いた方法は、イベント数が少ないデータや変量効果の分散が小さいときには、パラメータ推定のための反復計算が収束せず、推定結果が安定しないという可能性が高くなることが知られている (Breslow and Clayton, 1993)。そのため、健診データに基づく疫学研究データのようにイベント数が少ないときや、地域間差は小さいことがあらかじめ予想される状況においては尤度近似によるアプローチは有用な方法と言うことはできず、実際の疫学研究データにおいても尤度近似によるアプローチは計算が収束せず推定値が得られないことが確認されている。

そこで、尤度近似に代わるアプローチとして、モンテカルロ積分によって複雑な尤度関数を評価するアプローチであるマルコフ連鎖モンテカルロ法 (Markov Chain Monte Carlo method; 以下 MCMC 法)に基づく Gibbs sampling を用いてパラメータ推定を行うことが考えられる。

2.2.2 MCMC 法によるアプローチ

MCMC 法とはマルコフ連鎖の性質に基づき、同時確率分布を直接計算することなく、その分布に従うサンプルを何らかの分布から「繰り返し」サンプリングすることによって生成するモンテカルロシミュレーションである(Gilks et al. 1995)。中でも、Gibbs sampling は画像復元のアルゴリズムとして知られるが、Gelfand and Smith(1990)がベイズ統計学の枠組みで適用を拡張し、計算機の性能の向上も相まって統計学の分野でも広まるようになってきた。Gibbs sampling の本質は複数のパラメータを含む複雑な同時分布における各パラメータの完全条件付き分布（関心のあるパラメータ以外の全てのパラメータを条件付けた分布）からの繰り返しサンプリングによって、各パラメータの周辺分布を構成する方法といえることができる。

未知のパラメータを $\theta = (\theta_1, \theta_2, \dots, \theta_m)'$ 、観測されたデータを Y とする。事後分布 $\pi(\theta|Y)$ を求めることが困難であるとき、次のアルゴリズムに従い反復的に θ を発生させることによって事後分布からの標本を得て、パラメータの推定を行なうのが Gibbs sampling のアルゴリズムである。具体的には以下の手順に従う。

- (1) 初期値 $(\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_m^{(0)})$ を適当な分布から発生、あるいは適当に指定する。
- (2) θ_1 に関する 1 回目のサンプル「 $\theta_1^{(1)}$ 」を θ_1 以外のパラメータで条件付けた θ_1 に関する完全条件付き事後分布 $\pi(\theta_1 | \theta_2^{(0)}, \dots, \theta_m^{(0)}, Y)$ から発生させる。
- (3) $\theta_2^{(1)}$ を $\pi(\theta_2 | \theta_1^{(1)}, \theta_3^{(0)}, \dots, \theta_m^{(0)}, Y)$ から発生させる。
- (4) $\theta_3^{(1)}$ を $\pi(\theta_3 | \theta_1^{(1)}, \theta_2^{(1)}, \theta_4^{(0)}, \dots, \theta_m^{(0)}, Y)$ から発生させる。

同様に $\theta_4^{(1)}, \dots, \theta_m^{(1)}$ を順次発生させる。

- (5) $\theta^{(i)} = (\theta_1^{(i)}, \theta_2^{(i)}, \dots, \theta_m^{(i)})$ が得られたら、
 - (i) $\theta_1^{(i+1)}$ を $\pi(\theta_1 | \theta_2^{(i)}, \dots, \theta_m^{(i)}, Y)$ から発生させる。
 - (ii) $\theta_2^{(i+1)}$ を $\pi(\theta_2 | \theta_1^{(i+1)}, \theta_3^{(i)}, \dots, \theta_m^{(i)}, Y)$ から発生させる。
 - (iii) $\theta_3^{(i+1)}$ を $\pi(\theta_3 | \theta_1^{(i+1)}, \theta_2^{(i+1)}, \theta_4^{(i)}, \dots, \theta_m^{(i)}, Y)$ から発生させる。

同様に $\theta_4^{(i+1)}, \dots, \theta_m^{(i+1)}$ を順次発生させていく。

このように $\theta_1^{(i+1)} = (\theta_1^{(i+1)}, \theta_2^{(i+1)}, \dots, \theta_m^{(i+1)})$ を得る作業を $i = 1, 2, 3, \dots$ と繰り返す。

このとき $\theta^{(0)}, \theta^{(1)}, \theta^{(2)}, \dots$ は

$$K(\theta^{(i+1)}, \theta^{(i)}) = \pi(\theta_1 | \theta_2^{(i)}, \dots, \theta_m^{(i)}, Y) \prod_{j=2}^{m-1} \pi(\theta_j | \theta_1^{(i+1)}, \dots, \theta_{j-1}^{(i+1)}, \theta_{j+1}^{(i)}, \dots, \theta_m^{(i)}, Y) \pi(\theta_m | \theta_1^{(i+1)}, \dots, \theta_{m-1}^{(i+1)}, Y)$$

のように $K(\theta^{(i+1)}, \theta^{(i)})$ を推移核とするマルコフ連鎖になっている。 $N \rightarrow \infty$ のとき $\theta^{(N)}$ が事後分布 $p(\theta | Y)$ の標本になることが知られており、この過程を独立に t 回繰り返して $\theta^{(N), j=1, 2, \dots, t}$ が得られた時、個々のパラメータの算術平均

$$\frac{1}{t} \sum_{j=1}^t f(\theta^{(N), j})$$

は繰り返し数 t が十分大きく、目標とする分布(定常分布)からのサンプルに収束したと判断されたとき $E_\theta[f(\theta) | Y]$ に収束する、すなわち、個々の興味あるパラメータの事後分布の期待値に収束することが知られている。この性質を利用し、興味あるパラメータの推定を行うことが出来る。

2.3 サンプルの収束判定方法

定常分布からのサンプルに収束しているかの収束判定方法は

- ・ 図による方法
 - 標本系列(サンプリングパス)や標本平均の時系列プロット
 - 得られたサンプルの変動が初期値に依存せずに安定的であるかどうかを見て収束したかどうかを判定
 - 標本自己相関関数のプロット
 - 自己相関が高いときには状態空間の一部に偏ったサンプルがなされており、効率よく定常分布からサンプルが行なわれていないことを意味する
 - 標本平均・百分位数の時系列プロット
 - 標本平均・百分位数が安定してきているかを見る
- ・ 計算による方法
 - Geweke (1992)
 - 得られた標本系列の前半 10%と後半 50%でパラメータ θ の関数 $g(\theta)$ の期待値が同じかどうかを仮説検定
 - Gelfand and Smith(1990)、Gelman and Rubin(1992)、Liu and Liu(1993)
 - 異なる初期値から、いくつかの標本系列を別々にサンプリングし、標本系列間の安定性に基づいて収束判定を行なう
 - Raftery and Lewis(1992)

- 標本百分位数の収束を用いる
- Heidelberger and Welch(1983)
- マルコフ連鎖の一様エルゴード性の検定
- Mykland et al. (1995)
- 確率過程の再生時間の概念を用いるもの

などが挙げられる。どの方法が適切かについては様々な議論があるが、JALSO 次研究データの解析では Gelman and Rubin(1992)の収束判定法を採用した。この方法は異なる初期値からの標本系列が同様の値に収束しているかによって定常分布からの収束判定を行なう方法である。

2.3.1 Gelman and Rubin(1992)の収束判定法

サンプリングで m 個の独立した連鎖を発生させた下で、それらの連鎖の前半半分を捨てた残りの長さを n とする。連鎖内の各要素 φ_{ij} $i=1,2,\dots,n$ $j=1,2,\dots,m$ で表す。この要素から連鎖間分散 B と連鎖内分散 W を以下のようにして求める。

$$B = \frac{n}{m-1} \sum_{j=1}^m (\bar{\varphi}_{\cdot j} - \bar{\varphi}_{\cdot\cdot})^2$$

$$W = \frac{1}{m} \sum_{j=1}^m s_j^2$$

ここで、

$$\bar{\varphi}_{\cdot j} = \sum_{i=1}^n \varphi_{ij} / n$$

$$\bar{\varphi}_{\cdot\cdot} = \sum_{j=1}^m \bar{\varphi}_{\cdot j}$$

$$s_j^2 = \sum_{i=1}^n (\varphi_{ij} - \bar{\varphi}_{\cdot j})^2$$

である。以下の B, W の重み付け平均を計算することで φ の周辺事後分散を計算する。

$$\hat{\text{var}}(\varphi | y) = \frac{n-1}{n} W + \frac{1}{n} B$$

上記の式は周辺事後分散を過大評価したものであるがサンプルが増加するにつれて不偏推定量となる。連鎖内分散 W は過小評価されているが同じくサンプルが増加するにつれ不偏推定量となる。得られたサンプルが定常分布に収束しているかは次の Potential Scale Reduction の \hat{R} を計算し判断する。

$$\hat{R} = \sqrt{\frac{\hat{\text{var}}(\varphi | y)}{W}}$$

\hat{R} はサンプルが増加するにつれて 1 に近づく。 $\hat{R} \cong 1$ となったところで収束したと判断する。

2.4 各パラメータに対する完全条件付き事後分布

(1)式と(2)式のモデルにおいて推定すべきパラメータは、固定効果 β （ただし、これ以降では、(1)式の β_0 、 β_1 、 γ を全てまとめて β と表現する）、変量効果 b_i 、変量効果の分散共分散行列 D である。各パラメータの完全条件付き事後分布を、それぞれ $[\beta | b, D]$ 、 $[b | \beta, D]$ 、 $[D | \beta, b]$ と表現すると、 β と D は独立なので、その条件付き分布はそれぞれ $[\beta | b]$ 、 $[D | b]$ に帰着する。 $[b | \beta, D]$ についてはこれ以上簡略化することはできない。

・固定効果パラメータ β のサンプリング： $[\beta | b^{(k)}]$

変量効果 b_i を条件付けた下では、(1)式のポアソン混合効果モデルは、各対象者に対して変量効果に関する部分($b_{0i}^{(k)} + x_{ij} b_{1i}^{(k)}$)をオフセット変数とした通常のポアソン回帰となる。

β の事前分布に一様分布を仮定すると β の完全条件付き事後分布は尤度関数

$$\prod_{i=1}^I \prod_{j=1}^{n_i} f(y_{ij} | b_i^{(k)})$$

に比例する。この尤度は、平均が最尤推定量 $\hat{\beta}^{(k)}$ 、分散がフィッシャー情報量の逆数 $V_{\beta}^{(k)}$ の多変量正規分布で近似することが出来る。つまり、固定効果パラメータに関する完全条件付き事後分布からのサンプルは、多変量正規分布 $N(\hat{\beta}^{(k)}, V_{\beta}^{(k)})$ から乱数行列 $\beta^{(k+1)}$ を発生させることで得ることが出来る。

・変量効果 b_i のサンプリング： $[b_i | \beta^{(k)}, D^{(k)}]$

変量効果 b_i の事後分布は、以下の(4)式に示すように、(3)式の尤度関数に含まれる厄介な積分と同じ積分がその分母に含まれるため、明示的にその分布を表現することができないという問題がある。

$$[b_i | \beta^{(k)}, D^{(k)}] = \frac{f(y_i | b_i, \beta)g(b_i | D)p(\beta, D)}{\int f(y_i | b_i, \beta)g(b_i | D)p(\beta, D)db_i} \quad (4)$$

ただし、 $g(b_i | D)$ は変量効果に対する二変量正規分布、 $p(\beta, D)$ はパラメータ β 、 D に対する同時事前分布である。(4)式の分母には解析的に解けない積分が含まれているものの、分子は評価することが可能である。Gibbs sampling では完全条件付き事後分布からのサンプルが必要になるだけなので、必ずしも(4)式を解析的に解く必要はなく、繰り返し計算を必要としない棄却サンプリング (rejection sampling) などのモンテカルロ積分を用いて(4)式の分子からのサンプルを発生させることができる。

2.4.1 棄却サンプリング (Rejection sampling)

密度関数が $\pi(x)$ であるような確率分布から乱数を発生させることが困難であるような場合、正規分布など乱数発生が比較的容易であるような確率分布(以下、密度関数を $g(x)$ とする)からの乱数を利用して、以下の手順で間接的に $\pi(x)$ からの乱数発生を行なうサンプリング手法が棄却サンプリング(Rejection sampling)である。

もし $g(x)$ が $\pi(x) > 0$ であるような全ての x について $\pi(x) \leq cg(x)$ であるならば(ただし c は定数)

ステップ 1

$g(x)$ から乱数 x^* を発生させる。同時に、(0,1)区間上の一様乱数 u を発生させる。

ステップ 2

もし以下の式が成り立てば

$$u \leq \frac{\pi(x^*)}{cg(x^*)}$$

x^* を乱数を発生させたい目標分布 $\pi(x)$ からのサンプルと見なして受容し、そうでない場合は棄却してステップ 1に戻る。

とする。この手順で得られる確率変数 x^* は $\pi(x)$ に従う(Ripley,1987)。これはベイズの定理により以下の式が成り立つことから示すことができる。

$$\begin{aligned}
p(x|u \leq \frac{\pi(x)}{cg(x)}) &= \frac{pr(u \leq \pi(x)/cg(x) | x)g(x)}{\int pr(u \leq \pi(x)/cg(x) | x)g(x)dx} \\
&= \frac{\{\pi(x)/cg(x)\}g(x)}{\int \{\pi(x)/cg(x)\}g(x)dx} \\
&= \frac{\pi(x)}{\int \pi(x)dx} \\
&= \pi(x)
\end{aligned}$$

- ・ 変量効果分散 D のサンプリング : $[D|b_i]$

変量効果分散パラメータ D に対して、標準的な無情報事前分布 $p(D) \propto |D|^{-(q+1)/2}$ を仮定すると、変量効果の分散の逆行列 D^{-1} の条件付き事後分布は、次式で表されるパラメータ行列 $S^{(k)}$

$$S^{(k)} = \sum_{i=1}^I b_i^{(k)} b_i^{(k)'}$$

をもつ自由度 $(I-q+1)$ のウィッシュャート分布に従う。ここで、 q は変量効果の次元である。ランダムサンプル D^* は、自由度 $(I-q+1)$ の標準ウィッシュャート分布から標準ウィッシュャート乱数行列 W^* を発生させ、次式を計算することで求めることができる (Odell and Feiveson, 1966)。

$$D^* = (H^{(k)'} W^* H^{(k)})^{-1}$$

ここで、行列 $H^{(k)}$ はウィッシュャート分布のパラメータ行列 $S^{(k)}$ のコレスキー分解根である。

3. JALS0 次研究データにおける適応例

3.1 JALS0 次研究データ

JALS0 次研究は、JALS 参加研究者を中心とした日本各地の研究者が 1985 年から 2001 年までに各地域・職域で動脈硬化性疾患を対象として集積した既存の前向きコホート研究データを統合し、メタアナリシスを行うことを目的としており、現在までに地域・職域を含め、21 コホート 65,435 人の対象者が登録されている。今回解析対象とするエンドポイントは、脳梗塞、脳出血、クモ膜下出血を含む脳卒中発症である。検討するリスク因子は、年齢、収縮期血圧 (SBP)、Body Mass Index (BMI)、喫煙の有無、飲酒の有無、糖尿病既往の有無である。年齢、収縮期血圧、BMI はそれぞれ 65 歳未満・以上、140mmHg 未満・

以上、25 未満・以上の 2 値で考慮した。本解析では「脳卒中発症に対する喫煙効果に地域差があるか」の検討を行なう。本解析における対象地域は、解析に用いる変数を全て測定している北海道から福岡まで全国各 10 地域コホート(北海道 1～2 [地域番号：1]、秋田 1～3 [2]、秋田 4 [3]、岩手[4]、茨城[5]、滋賀 1 [6]、大阪[7]、愛媛[8]、福岡 1 [9]、福岡 2 [10])である。表 1、2 に男女の地域ごとにおける平均追跡期間、解析対象者数、性・年齢の分布、喫煙・飲酒状況、脳卒中発症者数などの基本特性を示す。解析対象者数は、男性 9,086 人 (75,414 人年)、女性 14,770 人(127,022 人年)で、脳卒中発症率は、男性で 4.51×10^{-3} (/年)、女性で 2.83×10^{-3} (/年)であった。

表 1 解析対象集団の基本特性 (男性)

地域	平均 追跡期間	男性		年齢		喫煙		飲酒		脳卒中発症	
		人数	%	平均	SD	人数	%	人数	%	人数	発症率
1	6.2	670	42.4	59.9	10.6	341	50.9	448	66.9	27	6.50E-03
2	11.3	466	32.7	56.3	13.3	229	49.1	289	62.0	22	4.18E-03
3	10.4	1086	43	55.8	10.8	679	62.5	839	77.3	55	4.87E-03
4	9.9	634	34.6	58.7	11.7	318	50.2	410	64.7	58	9.24E-03
5	9.8	1840	42.3	54.3	9.3	1150	62.5	1254	68.2	68	3.77E-03
6	9.2	1261	34.1	55.9	11.1	733	58.1	874	69.3	38	3.28E-03
7	7.1	1182	41.6	56.3	13.6	688	58.2	860	72.8	31	3.69E-03
8	5	977	30.8	58.3	12.3	391	40.0	652	66.7	15	3.07E-03
9	9.4	292	39.1	60.8	10.3	138	47.3	173	59.2	14	5.10E-03
10	3.7	678	40.4	63.2	10.8	250	36.9	499	73.6	12	4.78E-03
合計	8.3	9086	38.1	57.1	11.6	4917	54.1	6298	69.3	340	4.51E-03

表 2 解析対象集団の基本特性 (女性)

地域	平均 追跡期間	女性		年齢		喫煙		飲酒		脳卒中発症	
		人数	%	平均	SD	人数	%	人数	%	人数	発症率
1	6.4	909	57.6	59.4	9.8	66	7.3	153	16.8	27	4.64E-03
2	12.1	958	67.3	55.7	11.9	15	1.6	13	1.4	24	2.07E-03
3	10.8	1439	57	56	10.7	31	2.2	41	2.8	58	3.73E-03
4	10.8	1201	65.4	57.2	10.4	29	2.4	90	7.5	63	4.86E-03
5	10.1	2510	57.7	54.7	9.5	165	6.6	201	8.0	69	2.72E-03
6	9.8	2438	65.9	52.7	10.6	259	10.6	372	15.3	34	1.42E-03
7	7.3	1661	58.4	56.4	14.1	151	9.1	402	24.2	38	3.13E-03
8	5.3	2200	69.2	52.3	13.6	93	4.2	318	14.5	12	1.03E-03
9	9.8	455	60.9	60.6	10.1	19	4.2	42	9.2	31	6.95E-03
10	3.8	999	59.6	61.9	10.9	20	2.0	259	25.9	4	1.05E-03
合計	8.6	14770	61.9	55.6	11.7	848	5.7	1891	12.8	360	2.83E-03

3.2 Gibbs sampling の実行

Gibbs sampling を行なうにあたり、繰り返し数は 4,000 回とし初期値は 3 通りに変えてサンプリングを行った。収束判定を行なった後、各サンプリングにおける繰り返し 2,000 回日以降のデータをターゲットとする分布への収束後のデータと考え、各パラメータの分布は、 $2,000 \times 3 = 6,000$ 個のサンプルに基づいて記述した。以上の解析プログラムを SAS/IML Ver 9.1 により作成し、実行した。

3.3 結果

(1)式のポアソン混合効果モデルの解析は男女別に行い、いずれの場合も調整変数として年齢、収縮期血圧 (SBP)、Body Mass Index (BMI)、飲酒の有無、糖尿病既往の有無の 5 変数を用いた。Gibbs sampling によるパラメータ推定における収束状況を視覚的に確認するためのサンプリングパス (各繰り返しにおいて得られたサンプリング値の推移) を図 1~3 に示す。ただし、男女で同様の傾向を示したので、男性における喫煙に関する固定効果(図 1)、切片に関する変量効果分散(図 2)、喫煙に関する変量効果分散の結果(図 3)のみを示した。各パラメータとも、サンプリングの繰り返し回数が初期の段階から安定した範囲内で推移していることがわかる。他のパラメータに関しても同様の推移を示していた。

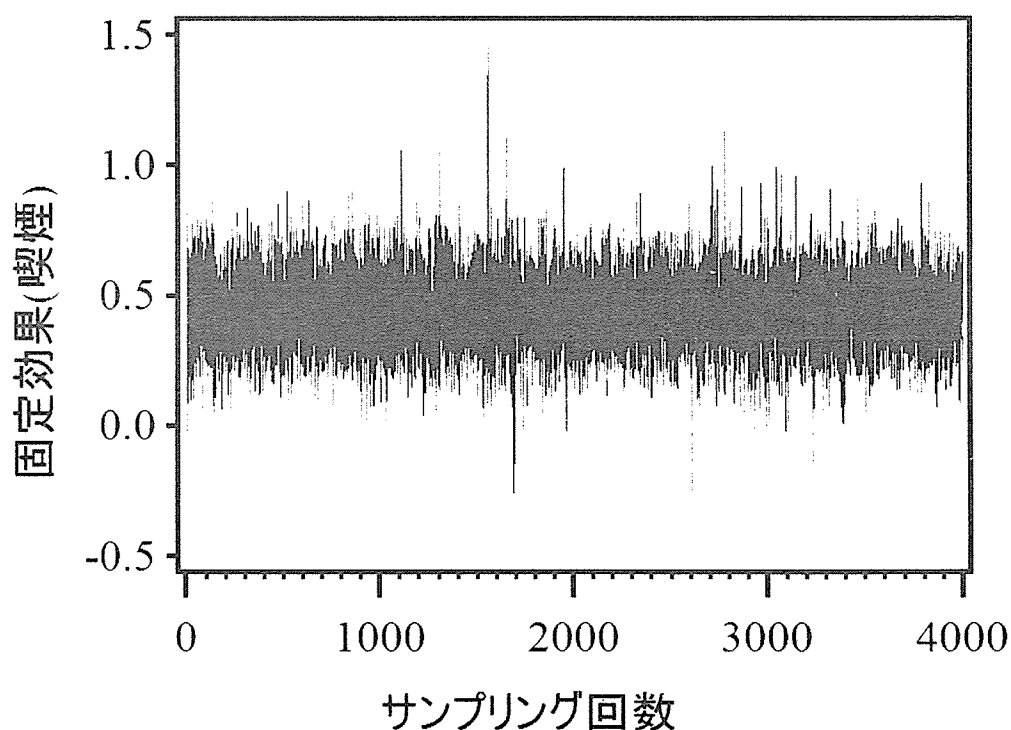


図 1 固定効果(喫煙)のサンプリングパス