

〈教育報告〉

平成19年度専門課程 II
地域保健福祉分野

中堅保健師の仕事ストレスに関連する要因

北山明子

The Factors of Job Stress by Mid-Career Public Health Nurses

Akiko KITAYAMA

Abstract

Objectives The objectives of this study are to identify the factors of job stress that mid-career Public Health Nurses (PHNs) experience in their work of Hokkaido Public Health Center, and to consider possible strategy to prevent burnout. **Methods** The research participants target of the study were 114 of mid-career PHNs who work in Hokkaido Public Health Center, and the research was conducted by mailed self-administered questionnaires. PHNs' response to stress in job was measure by Burnout scale (Japanese version of MBI), and the questionnaire was structured into three frameworks such as "working environment", "buffer factor" and "personal factor", which were examined by relevancy to Burnout. **Results** Eighty-nine responses (response rate 78.1 percent) were collected, and effective 84 responses were analyzed. After reviewing the responses according to years of their experience, self accomplishment felt off significantly among PHNs in the first term (between 6 and 10 years). The second terms of PHNs (between 11 and 15 years) showed the highest degree of emotional exhaustion and depersonalization. Burnout of the first term was significantly related to a burden to a job, support to others, and professional self-efficacy. Burnout in the second term was related to job conflict and job satisfaction. In the third term (between 16 and 20 years), it was related to communication problems with others and insufficient job control. The fourth term (more than 20 years) of PHNs showed lower score of burnout. **Conclusion** It was indicated that supervisor's support and assistance from people in the same section were the most effective to ease emotional exhaustion and support from colleagues was the most effective for depersonalization and decreased self-accomplishment. To prevent burnout of mid-career PHNs, it is necessary to establish working environment in which the first and second term of PHNs have enough support.

Keywords: mid-career Public health nurses, Burnout, Job stress, organizational reformation of public health center, Social support

Thesis Advisor: Hirohisa IMAI

I. 目的

機構改革による保健師の分散配置，少人数体制によるマンパワー不足，仕事内容の複雑多様化などにより，業務の中心を担う中堅保健師はストレスの多い職場にさらされている¹⁾。本研究では，北海道保健福祉事務所で働く中堅保健師が感じる仕事のストレス反応である Burnout の関連要因を明らかにし，保健師の Burnout 予防に向けた職場環境について検討することを目的とした。

本研究では，中堅保健師を「行政保健師としての経験年数が6年目以上であり，かつ，係長・主査などの役職に

就いていない人」と定義する。

II. 研究方法

1. 調査方法

北海道保健福祉事務所で就労している中堅保健師114名を対象に，郵送法による無記名自記式質問調査を実施した。調査期間は2007年11月12日から11月30日である。なお，本調査は，国立保健医療科学院倫理委員会の審査を受けている。

2. 調査項目

仕事上のストレスによって生じる反応を Burnout 尺度（日本語版 MBI²⁾）を用いて測定し，Burnout に関連すると考えられる要因を，①職場要因（仕事の負担感，業務量に

指導教官： 今井博久（疫学部）

見合った報酬，不十分な技術活用，職場のコンセンサス，地域と関わる仕事)②緩衝要因(上司・係内・保健師同士のサポート，休息)③個人要因(職務満足，自己肯定感，キャリア志向性)の3つの枠組みで構成し，4件法で回答を求めた。Burnout尺度はいずれも得点が高いほどBurnoutが高いことを示す。統計ソフトはSPSS Ver.15.0 for windowsを用い，有意水準は5% (両側)とした。

III. 結果

1. 対象者の概要

回収数は89名(回収率78.1%)，有効回答数は87名であり，女性84名を解析対象とした。対象者の平均年齢は38.7歳であった。

2. 中堅保健師全体の Burnout の傾向

Burnout 得点結果は，「情緒的消耗感」は14.58±4.39，「脱人格化」は12.52±4.37，「個人的達成感低下」は16.46±3.70であった(平均±SD)。

3. 経験年数別の Burnout の比較(図1)

中堅保健師を，経験段階別に「I期(6-10年目)」「II期(11-15年目)」「III期(16-20年目)」「IV期(21年目以上)」に分類し，Burnout 得点の比較を行った。最も情緒的消耗感と脱人格化が高かったのはII期であり，個人的達成感が低下していたのはI期であった。

4. Burnout に関連する要因

I期の情緒的消耗感には，「仕事の負担」が多く「係内サポート」がないことが関連していた。脱人格化と個人的達成感低下には「地域と関わる仕事」が出来ていないこと，「保健師としての成長感」がないこと，「保健師同士のサポート」がないことが関連していた。II期では，3つの下位尺度すべてに「職務満足」が関連しており，「悩む仕事」が脱人格化に，「保健師からの助言」がないことが個人的達成感低下と関連していた。III期では「自分のペースで出来ない仕事」が情緒的消耗感と関連しており，脱人格化や個人的達成感低下には「係内の意思疎通」がなく「助け合い」がないこと，「保健師同士のコンセンサス」が得られないことが関連していた。IV期は Burnout 得点が低かった。

5. 上司の職種別にみた Burnout 緩衝要因

他職種の係長は，係員である保健師に対するサポートが少なかった。他職種の係長の下で働く保健師の情緒的消耗感と脱人格化の緩衝要因は，「上司のサポート」「係内サポート」「休息」であった。

IV. 考察

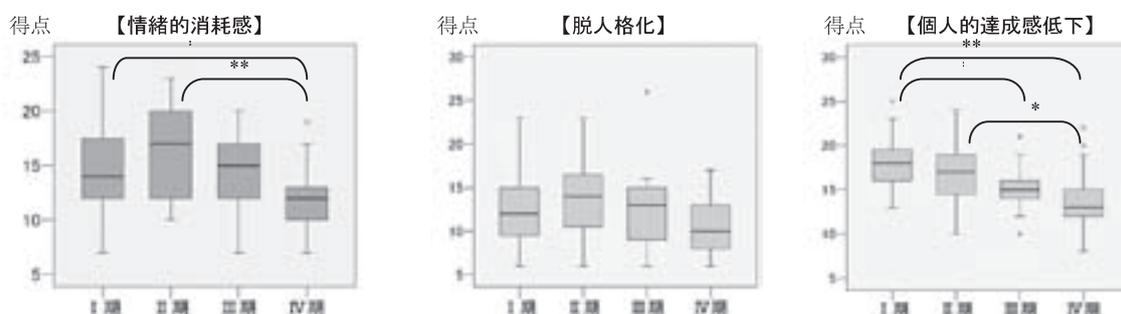
本研究では中堅保健師の Burnout 関連要因が明らかになった。I期は経験が少ないため，未経験の仕事や複数の仕事を抱えると負担を感じやすい。係内で業務の遂行状況を理解しあい，必要な時に手助けを求めやすい体制づくりが必要である。達成感低下の予防には，地域と関わる仕事を増やし，保健師としての成長を実感できる機会を増やすことが必要であろう。II期は10年目を過ぎた保健師として新たな役割や責任が加わり始める時期であり，中間的立場でもあるため仕事の進め方や周囲との調整に悩むことが多い。他系の保健師係長や同じ中堅保健師同士で相談・助言しあえる体制があると，経験を交えた具体的助言が得られるので悩みが解決され，満足度の高い仕事ができるであろう。中堅保健師の Burnout 予防には，特にI期とII期に対するサポートが必要である。

V. 結論

1. 中堅保健師の Burnout 関連要因が経験期別に明らかにされた。
2. Burnout 予防対策はI期には専門職としての成長感を獲得できるような職場内教育，II期には保健師同士の助言体制整備が必要である。

文献

- 1) Imai H, Nakao H, Tsuchiya M, Kuroda Y, Katoh T. Burnout and work environments of public health nurses involved in mental health care. *Occupational & Environmental Medicine* 2004;61(9):764-768.
- 2) 久保真人, 田尾雅夫. 看護師におけるバーンアウト～ストレスとバーンアウトとの関係. *実験社会心理学研究* 1994; 34: 33-43.



【図1 経験年数別の Burnout 比較】 **p<0.01 *p<0.05

〈教育報告〉

平成19年度専門課程 II
地域保健福祉分野

学校保健と地域保健の連携プロセスにおける保健所保健師の役割

徳永瑞希

A Study of Public Health Nurses' Roles in the Process of Collaboration with School Health and Community Health

Mizuki TOKUNAGA

Abstract

Object and Method The purpose of the study is to understand process in which Public Health Nurses (PHN) sustained and developed the solidarity between school health and community health, and their roles in the process. This qualitative study employed the interview method, including the nine PHNs' within seven public health centers in six local governments. As the framework of analysis, the study adopted the model of 8 steps with 13 categories of organization collaboration process by the Japan Nursing Association. **Result and Conclusion** A new level "level-0; the groundwork for collaboration with school health and community health emerged in addition to the previous eight -leveled model. This new model resulted in nine steps. **Conclusion** The identified PHNs' roles in creating and sustaining solidarity between school health and community health were 1) to include schools as a target that requires public health activity and to establish mutual goals through collaboration, 2) set up same goal with school, 3) everyday collaboration and relationship in workplace in order to develop to establish the systematic solidarity base in public health center, 4) to identify a key person and approach him / her based on the stage of collaboration process, 5) to promote multi-sector approach, 6) to make and share the benefit of collaboration outcome, and 7) share the result of program evaluation,

Keywords: roles of public health nurse, school health, community health, collaboration process

Thesis Advisor: Hiroko OKUDA

I. 目的

学校保健と地域保健の連携を実現し継続させていくプロセスにおいて保健所保健師が果たしている役割を明らかにする。

II. 研究デザインと方法

1. 対象者：学会誌などで得られた過去5年間における保健所保健師の学校保健への実践事例を候補とし、同意の得られた6自治体7保健所9名の事業担当保健師。
2. 調査期間：平成19年9月
3. 調査方法：予備調査によって半構成式質問紙を作成し、本調査では①事業概要②事業背景③取り組みの実際④学校保健と地域保健の連携意義について面接調査を行なった。許可を得て会話を録音するとともに関連資料の提供を受け

た。

4. 分析方法：語りを逐語記録に起こして連携をすすめるための視点や働きかけ、判断などに関する部分を抽出してコード化して、さらにサブカテゴリーに分類し、「組織的な連携を促進するモデル」¹⁾(以下、モデル)と照合してカテゴリー化を行った。なお、分析では指導教官や保健師経験をもつ複数の研究者と検討を繰り返し、信頼性・妥当性を高める努力をした。

5. 倫理的配慮

国立保健医療科学院研究倫理審査委員会の承認を受けた。(承認番号 NIPH - IBRA#07014)

III. 結果及び考察

1. 研究対象者の特性：9名の保健師から7事業の聞き取りを行なった。(表1)

指導教官：奥田博子（公衆衛生看護部）

表1：対象保健師の属性及び主な事業名

事例	経験年数(年)	担当部署	主な事業名
A	13	企画部門	青少年健康被害予防活動事業
B	24	母子保健	子どもの心の健康を考える事例検討会
C	13	感染症	性感染症予防啓発プログラム作成会議
D	22	母子保健	思春期保健事業
E	23	企画部門	思春期健康支援プロジェクトチーム
	24	母子保健	
F	38	総括(課長)	思春期保健教育事業
G	8~9	母子保健	小児のぜん息予防教室プログラムの検討
	10~11		

2. 組織的な連携を促進するプロセス

714ワード、51サブカテゴリー、22カテゴリーと新たに0段階が抽出され、9段階に整理された。(図1) また、全ての事例がモデルの1~8段階を辿っていた。新たな部分を中心に以下に示す。文章中の段階は【】、カテゴリーは《》、サブカテゴリーは『』、留意点は〔〕で示す。

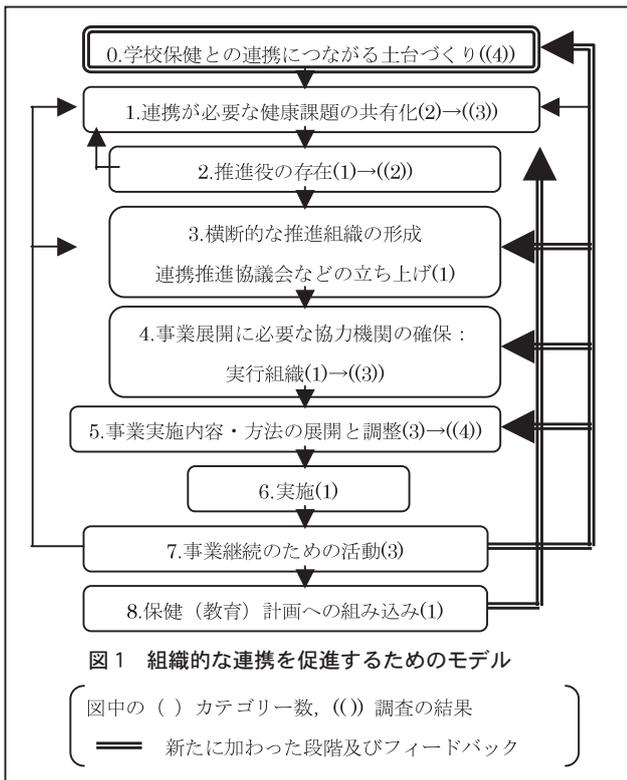


図1 組織的な連携を促進するためのモデル

〔 図中の () カテゴリー数, (()) 調査の結果
 ━━━━━━━━━━ 新たに加わった段階及びフィードバック

2.1 0段階【学校保健との連携につながる土台づくり】：《学校保健に係わる地域保健事業の位置づけ》《平常時の学校保健とのかかわり》《子どもを取り巻く地域実態を掴む》《子どもを取り巻く地域課題への対応》の4つで構成された。日常から地域のあるべき姿から学校保健との連携によってめざす姿を捉えて連携の必要性を認識しており、日

常に課題や連携目的の共有を貫いて学校保健へ意図的にかかわっていたことが土台となって、1~8段階のプロセスにつながると考える。

2.2 1段階【連携が必要な健康課題の共有化】：《必要となる事項を明らかにする》《課題を共有化する場所をもつ》というモデルの間に《具体的な事業方針を決定する》が抽出され、さらに事業方針も共有されていたため、《課題及び事業方針を共有する場所をもつ》の3つとした。

2.3 2段階【推進役が存在】：《推進役の明確化》《推進役の役割を明らかにする》の2つが新たに抽出され、保健師は「連携をする組織や個人がメリットを得られる働きかけ」に留意していた。

2.4 4段階【事業展開に必要な協力機関の確保】：《必要な人材の明確化》《人材を確保するための働きかけ》《実行組織の形成と各組織の相互理解》の新たなプロセスを経て実行組織が形成されていた。

2.5 5段階【事業実施内容・方法の展開と調整】：最初のプロセスとして新たに《目的の共有》が抽出され、モデルの《具体的な活動方法、内容、ツールの開発》《関係者を実施者へと変える》《評価方法の決定》と合わせて4つとした。

2.6 7段階【事業継続のための活動】：《評価活動とその共有化》《協力者/協力者の広がり》《新たな資源の動員》の3つで、『実施方法の評価』『評価の共有化(フィードバック)』が図られていた。

2.7 8段階【保健(教育)計画への組み込み】：『事業方針の協議・見直し』を行い次の展開を決定していた。

2.8 プロセス全体

日常から課題や連携目的の共有を図ることを貫いて意図的に学校保健や多機関とかかわって協働していた。また、組織的な連携成果を目的に合わせてフィードバックし発展的に継続させていた。

IV. 結論

組織的な連携を実現し継続させていくプロセスにおいて、保健所保健師が果たしている主な役割は以下の7点である。

1. 連携によってめざす姿を目的として捉える
2. 共に解決すべき課題と連携目的を学校保健と共有する
3. 日常的に所内連携体制を整え組織的連携の基盤をつくる
4. キーパーソンを見極めて段階に沿ってアプローチする
5. 必要な機関を幅広く捉え多機関を巻き込む
6. 組織と個人のメリットが得られるように働きかける
7. 事業成果はフィードバックして共有し次の取り組みへとつなぐ

<引用文献>

1) 日本看護協会. 子どもの健康づくりにおける地域・学校保健連携支援事業報告書, 2005.3.31

〈教育報告〉

平成19年度専門課程 II
生物統計分野

小地域におけるリスクの推定値の評価

水口正彦

The Assessment of Risk Estimator in Small Area

Masahiko MINAGUCHI

Abstract

Background When the mortality risks are compared between regions, SMR (Standardized Mortality Ratio) is assumed to be an index, and it shows in disease map. However, it is pointed out that the accuracy of SMR gets unsteadiness differing by the region in the comparison when the small region of the population scale of each municipal district town and village exists together. The evaluation is not enough though empirical Bayes estimator of SMR and other estimators are proposed from these problems. **Purpose** The feature of estimator that takes the place of SMR is clarified and evaluated from the viewpoint of fitting corresponding to the feature of data. **Method** We make data sets with the feature that suits the real situation, and compare models by the Monte Carlo simulation. The Poisson-Gamma model, the Log-Normal model, the BYM (Besag, York and Mollie) model, and the Mixture model are used for comparing. These models are evaluated by DIC (Deviance information criterion) and the mean square error of estimator. **Conclusion** Regardless of structure, estimator of the Mixture model is good from the viewpoint of DIC and MSE.

Keywords: disease map, risk assessment, Bayesian model, small area, SMR

Thesis Advisor: Kunihiko TAKAHASHI, Kazue YAMAOKA, Toshiro TANGO

I. はじめに

一般に死亡リスクの地域間比較の場合、死亡事象が Poisson 分布に従うことを仮定すると、観測死亡数は期待死亡数、死亡リスクを用いて以下のように記載できる。

$$O_k \sim \text{Poisson}(\theta_k E_k) \quad (1)$$

O_k : 地域kにおける観測死亡数,

E_k : 地域kにおける期待死亡数,

θ_k : 地域kにおける相対リスク

死亡リスクを地域間比較するという事は、この時の θ を推定したいリスクと考え、これらを地域毎で求めて比較するという事である。特に死亡リスクの地域間比較の際には SMR (Standardized Mortality Ratio; 標準化死亡比) がリスク指標とされることが過去において多くなされてきた。しかし市区町村などいわゆる小地域の死亡リスクの比

較検討の場合に SMR を用いることは、その不安定性により不適切であることが今まで指摘されてきた。これらの問題点から、ベイズ推定によるモデルを用いた推定値などリスクを推定するための様々なモデルが提案されている¹⁾。しかし、一方で、提案されてきた各モデルの比較調査は、限られた状況を除いてなされていなかった。

近年になって Lawson 他²⁾による広範囲のモデルについての比較結果、Richardson 他³⁾による人口規模を考慮し、背景リスクより高リスクの地区を設定するシナリオに基づき各モデルに対する推定値の評価結果などが出てきている。しかし、背景リスクより高リスクとなる地区もあれば低リスクとなる地区もあるのがより一般的であると考えられる。

そこで、本研究では低リスクも含めた異なるシナリオを設定し、Lawson 他²⁾に使用されていないものを含めた複数のモデルに対してそのリスクの推定値を求め、その推定値の適合度の観点から比較評価することを目的とする。

II. 方法

1. 相対リスク推定のための各種モデル

指導教官: 高橋邦彦, 山岡和枝, 丹後俊郎 (技術評価部)

本論文で取り扱うベイズ推定によるモデルについて述べる。この中には経験ベイズ法と、フルベイズ法があるが、より一般的なモデルであることから、比較モデルとしては後者を対象とする。

1.1 Poisson-Gamma (PG) モデル

事前分布が Gamma 分布に従うモデルである

1.2 Log-normal (LN) モデル

Poisson-Gamma モデルでは困難であった共変量を用いた拡張が容易なモデルである

1.3 BYM (Besag, York and Mollie) モデル

空間的な相関のない地域差と空間的な相関のある地域差と両者を考慮に入れたモデルである

1.4 Mixture (MIX) モデル

隣接地域リスクの空間的に連続した平滑性を考慮した地域とそうでない地域の両方をあわせもつものとして提案されたモデルである。

2. シミュレーション

2.1 シミュレーション設定

本研究では真の相対リスクを設定し、対象となるモデルに対してリスクの推定値、ならびに DIC を算出する。データでは Clayton⁴⁾ 他による Scotland の口唇癌の数値を用いる。

本研究ではこのデータを用いて地区の特定と高リスク・低リスクを考慮したシナリオを設定する。また、特に Mixture モデルの評価をするため特別なパターンを追加する。これらシナリオ設定に基づき各地区のリスクに対応した観測死亡数を発生させる。これを用いてモンテカルロシミュレーションを行いモデルの比較、および推定値の適合度の評価を行うものとする。評価は DIC、および推定値の平均二乗誤差により行う。

III. 結果

代表例とし、高リスクを設定した DIC, Mixture モデル評価 Pattern での結果を図 1, 図 2 で示す。図 1 で実線、点線 (dashed), 点線 (dotted) の各グループが同じリスクをもつ。また、S,M,L はそれぞれ人口規模を示す。

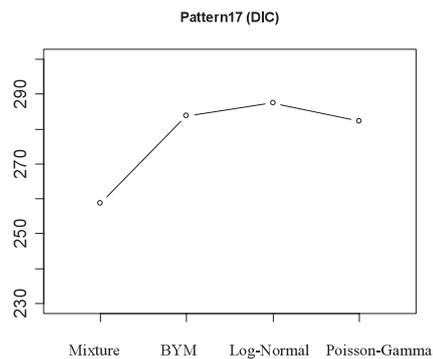


図 2. Mixture モデル評価 Pattern での DIC

IV. 考察

今回のシミュレーションから、DIC, MSE に基づく評価においては Mixture モデルが全ての Pattern において最良、もしくはほぼ最良に近い次善のモデルとなる結果が得られた。

Mixture 評価用にドーナツ状に高リスクを配置した場合には隣接情報を取り込む BYM モデルの場合には不利になることが予想されたが、結果として BYM モデルは Mixture モデルには及ばないが Log-Normal モデル、Poisson-Gamma モデルと比較した場合には良い結果を得た。

V. 結論

Mixture モデルによる推定値は全体的によい結果を示した。また、地域相関のあるデータでは、高リスク、低リスク両方を含むような場合であっても、Log-Normal モデル、Poisson-Gamma モデルより Mixture モデル、BYM モデルの推定値の方がよい結果を示した。

参考文献

- 1) 丹後俊郎, 横山徹爾, 高橋邦彦. 空間疫学への招待: 朝倉書店; 2007
- 2) Disease mapping collaborative group., Disease mapping models :an empirical evaluation. *Statistics in Medicine* 2000;19:2217-2241.
- 3) Richardson S., Thomson A., Best N., Elliott P. Interpreting Posterior Relative Risk Estimates in Disease-Mapping Studies. *Environmental Health Perspectives* 2004;9:1016-1025.
- 4) Clayton D, Kaldor J. Empirical Bayes Estimates of Age-standardized Relative Risks for Use In Disease Mapping. *Biometrics* 1987;43:671-681.

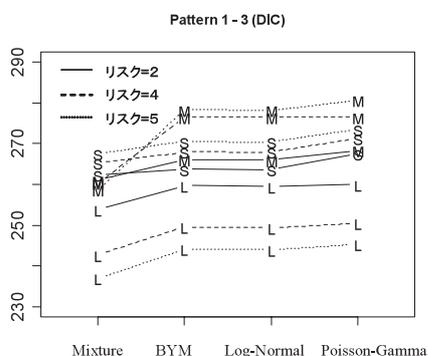


図 1. 特定地区のみに高リスクを設定した場合における各モデルの DIC

〈教育報告〉

平成19年度専門課程 II
生物統計分野

不完全なデータに基づくメタ・アナリシス グループ化されたデータに基づく回帰直線に関する検討

城戸口和成

A Trend Estimation from Grouped Dose-response Data and Its Application to Meta-analysis

Kazushige KIDOGUCHI

Abstract

Objective As study of trend estimation from grouped data of general research report, we examined regression line and slope of the association between dose or exposure (x) and corresponded value (y), approached to meta-analysis. **Methods** We supposed linear single regression model to summarize two variables of x and y , and assumed suitable distribution to grouped data of x , estimated assigned values in each level of x by maximum likelihood and expectation of truncated distribution. And we estimated regression slopes and lines by weighted regression analysis in assigned values of x and corresponded values of y . Moreover we examined estimation of distribution x by Monte Carlo simulation, confirmed precision of regression slopes and compared estimations of slopes in simulation with original regression line. **Result** We could estimate regression slopes and one assigned values of x . Estimation of regression slopes by simulation was well, and comparison with original model. **Conclusion** We thought that we could do meta-analysis by regression slopes obtained by estimations of regression analysis in grouped data.

Keywords: trend estimation, grouped data, regression line, regression slope, meta-analysis

Thesis Advisor: Kunihiko TAKAHASHI, Tetsuji YOKOYAMA, Toshiro TANGO

1. 目的

疫学調査などにおいて、ある因子の曝露と疾病の関連をみる際に、その曝露量の増加に伴って疾病に対するリスクがどの程度大きくなるかをみることがある。そのような研究において、独立変数である曝露量 X とある疾病のリスク指標である従属変数 Y との関係を線形回帰モデルによる解析を行い、傾き β の推定値 $\hat{\beta}$ によってその関連性を検討することがよく行われる。

一方で、複数の研究を統合するメタ・アナリシスを考える場合、発表されている研究結果・論文のデータを用いた統合を考えることになる。しかし、一般に公開されている研究結果・論文では解析に用いた全てのデータが記載されることはなく、特に、連続量に関するデータに関しては、そのデータを任意の範囲ごとにグループ化し、その区分ご

とに対応する従属変数 Y の要約統計量のみを表記されることがある。グループ化されたデータを表 1 に示す。

X の各区間の代表点を選び、それを用いた重み付き回帰による推定から、各研究の傾きを推定することが考えられるが、その場合、代表点の選び方で推定される傾き $\hat{\beta}$ の値に違いが出てしまう。本研究では、高橋・丹後^{1,2)} 同様、独立変数 X がグループ化されている 2 つの連続量の関係を表す回帰直線の傾きの推定を行い、さらにメタ・アナリシスに向け傾きの統合について検討を行う。また、その推定精度の議論のためシミュレーションによる検討を行う。

表 1 アルコール摂取と収縮期血圧 (SBP) の関係

Alcohol 摂取量 (g/week)	データ 数	平均値 (mmHg)	SE
0 ~ 20.4	58	119.320	2.387
20.4 ~ 183.8	57	125.683	2.470
183.8 ~ 367.7	18	127.011	4.336
367.7 ~	4	125.994	9.160

指導教官：高橋邦彦，横山徹爾，丹後俊郎（技術評価部）

II. 方法と結果

1. 回帰係数の推定及び傾きの統合

[回帰係数の推定及び統合に関する手順]

1. 独立変数 X の分布を推定する.
2. その分布から各区間の代表値 d_j ($j=1, 2, \dots, m$) を推定し, (d_j, \bar{y}_j) の値を用いて重み付き回帰推定を行う.
3. 各研究から推定された回帰係数を統合する

[手順1] X の分布の推定

X がある分布 $F(\theta)$ にしたがっているとき, X_j が区間 j に入る確率を P_j ($j=1, 2, \dots, m$) としたとき, X_1, X_2, \dots, X_N の同時確率から各区間のデータ数 n_j ($j=1, 2, \dots, m$) が与えられたときの尤度 L 及び対数尤度 l は, 次のとおりである.

$$L = C * P_1^{n_1} P_2^{n_2} \dots P_m^{n_m}$$

$$l = \log L = \sum_{j=1}^m n_j \log P_j$$

本研究では, 正規分布又は対数正規分布を F として考えた. 表1に適用した結果, 対数尤度の値から, 対数正規分布が選択された.

[手順2] 回帰係数の推定

推定された X の分布から切断分布を考え, その期待値より X の区間の代表値を推定した. X の代表値と平均値 Y より重み付き回帰を行い, 回帰係数を推定することができた. また, 傾き β の漸近分散により95%信頼区間を推定することができた.

[手順3] 傾きに関するメタ・アナリシス

今回は, 複数の研究を想定して異なるグループ化データを作成し, これを用いてメタ・アナリシス³⁾の検討を行った. 各データごとに傾き $\hat{\beta}$ を推定し, 漸近分散と, データ数より重みをつけて, 重み付き平均により共通の傾きを推定し, その95%信頼区間を推定することができた.

2. シミュレーションと結果

真の回帰モデルを設定し, X の分布に正規分布と対数正規分布の場合を考えて検討した. 真のモデルの傾き (1.75) に対して, X の分布を正規分布とした場合, シミュレー

ションによる傾きの推定値は1.7099であった. また, 全体のデータによる傾きは1.7498であった.

X の分布を対数正規分布とした場合, シミュレーションによる傾きの推定値は1.6607であった. また, 全体のデータによる傾きは1.7486であった.

なお, シミュレーションの際, X の分布を正規分布とした場合, 約4割は対数正規分布が推定された.

III. 考察

グループ化データに示される区間を持つ変数 X の分布は未知であるが, 初めから分布を仮定して推定する場合よりも, 本方法により分布を推定し, 回帰係数を推定する方法は, 推定の質を高めると考えられた. シミュレーションの結果から, 本方法による回帰係数の推定精度は良好と考えられたが, 分布の区分に関する問題点はある.

メタ・アナリシスの検討から, 複数のグループ化データから得た傾きを統合し, 共通の傾きの推定が可能と考えられたが, 選択される分布などの課題はある.

IV. まとめ

今回の結果から, 全データが入手できない場合でも, 本方法により, 研究結果・論文等に表記される2変量の要約統計量のみから回帰直線の推定が可能と考えられた. また, メタ・アナリシスを行うことができた. 今後, その推定精度やメタ・アナリシスの方法について, さらに検討, 改良の余地があると考えられた.

文献

- 1) 丹後俊郎, 高橋邦彦. メタ・アナリシスにおける2つの回帰直線の傾きの比の信頼区間, 厚生労働科学研究費補助金医療技術評価総合研究事業「エビデンスを適切に統合するメタ・アナリシスの理論, 応用と普及に関する調査研究」平成16年度 総括・分担研究報告書. p.19-22.
- 2) 高橋邦彦, 丹後俊郎. メタ・アナリシスにおける2つの回帰直線の傾きの比の信頼区間, 厚生労働科学研究費補助金医療技術評価総合研究事業「エビデンスを適切に統合するメタ・アナリシスの理論, 応用と普及に関する調査研究」平成17年度 総括・分担研究報告書. p.31-34.
- 3) 丹後俊郎. メタアナリシス. 東京: 朝倉書店; 2002.

〈教育報告〉

平成19年度専門課程 II
生物統計分野

乳がんの治療における QOL の向上 —アロマターゼ阻害薬の治療に関するメタ・アナリシス—

佐藤恵子

Meta-analysis of the Effects on Quality of Life of Aromatase Inhibitors as Adjuvant Treatment for Postmenopausal Patients with Breast Cancer

Keiko SATO

Abstract

Objective To evaluate the efficacy on the quality of life (QOL) of aromatase inhibitors (AIs) as adjuvant treatment for postmenopausal patients with hormone receptor positive breast cancer, compared with tamoxifen therapy or placebo. **Method** Systematic review and meta-analysis. Electrical databases as well as clinical guidelines and other unpublished literature resources were used for systematic review of randomized controlled trials (RCT). Difference in means of scores of QOL in terms of physical domain and endocrine and postmenopausal domain and rate ratio of QOL improvement rate were the effect sizes. **Results** From the electric search and hand search, 279 studies were identified. Three RCTs met all the inclusion criteria. Because final result was not obtained for one RCT, the results from interim analysis were used in the present study. Compared to tamoxifen, 2-years AIs treatment did not show significant improvement in QOL in terms of physical domain and endocrine and postmenopausal domain. **Conclusions** We could not establish the efficacy on QOL of AIs from the interim analysis. Assessment by the final results of the trials is needed. Focusing on other aspects of QOL such as psychological, social and functional domains is important as well.

Keywords: quality of life, breast cancer, aromatase inhibitors, systematic review, random effects model.

Thesis Advisor: Kazue YAMAOKA, Toshiro TANGO

I. 背景と目的

財団法人がん研究振興財団発行のがんの統計'07によると、女性の乳がんの2001年における罹患率は部位別で1位、2005年における死亡率は、部位別で第5位となっている。毎年のように臨床現場で使用可能になる新規治療薬剤の効果で、乳がんの生存率は向上している。しかし現時点では、QOL 評価の重要性は指摘されているものの、メタ・アナリシスによる乳がん治療薬剤の QOL へ与える影響の評価は行われていない。本研究では、閉経後の早期・一次性乳がんの補助療法におけるアロマターゼ阻害薬（以下、AI）治療に関するメタ・アナリシスを実施し、AI 治療が乳がん患者の QOL におよぼす影響について検証する。

II. 研究デザインと方法

研究仮説：早期・一次性乳がんの術後補助療法において AI 治療を行った患者の QOL は、従来治療の患者の QOL に比べて向上する。

1. 乳がん治療に関する無作為化比較試験のシステムティック・レビュー

データベースでの論文検索及びハンドサーチにて、乳がん治療に関する無作為化比較試験（以下、RCT）を検索し、メタ・アナリシスを行う論文を選定した。統合する指標は、ランダム化後から特定の一時点までの QOL スコアの変化量の平均値（以下、変化量）およびランダム化後の特定の一時点において QOL が悪化または不変である対照群に対する AI 治療群のリスク比をとりあげた。統合する QOL の次元は、身体症状全般および乳がん患者に特有の問題である内分泌・更年期症状とした。統合する RCT ご

指導教官： 山岡和枝、丹後俊郎（技術評価部）

と異なる QOL 尺度が用いられている場合には、QOL スコアの変化量の標準化された平均値を比較することとした。

2. メタ・アナリシスの統計モデル

研究間の均質性の検定は、カイ二乗検定により行った。RCT における介入方法や QOL 測定方法の研究間での相違を勘案し、変量効果モデルを主要な検定として取り扱うこととした¹⁾。統合値の推定には、代表的な変量効果モデルである DerSimonian-Laird (以下、DSL) の方法²⁾を用いた。

III. 結果

279論文がシステマティック・レビューの対象となり、乳がんの RCT 論文91報を詳細に検討した結果、3つの RCT (ATAC, IES, MA.17 study) が抽出された。これらの RCT の QOL サブ・プロトコルの報告^{3) 4) 5)}は中間解析時点までのみ論文化されているものもあったため、本研究では中間解析時点までの結果の統合を行うこととした (表 1)。

表 1. 統合に使用したRCTのQOLサブ・プロトコルの概要

Study (試験開始年)	Participants	Median follow-up (months)	QOL measurement	Timing of measurement	Primary end point	Result
ATAC (1996)	1,105	24	FACT-B ES	Entry, 3,6,12, 18,24 month	TOI score of FACT-B	NS
IES (1998)	582	24	FACT-B ES	Entry, 3,6,9,12, 18,24 month	TOI score of FACT-B	NS
MA.17 (1998)	3,612	30	MENQOL SF-36	Entry, 3,6,12, 24,36 month	Proportion of patients who reported a change in QOL at any time in any of the SF-36 summary measures or domains of MENQOL	Vasomotor domain showed significant difference 22%(placebo), 29%(AI) reported worsening (p<0.001).

TOI: Trial Outcome Index, NS: Not Significant

使用されていた QOL 尺度は RCT 間で異なっていた。本研究では、身体症状全般として FACT-B の Trial Outcome Index (以下、TOI) および SF-36の Physical Component Summary (以下、PCS)、内分泌・更年期症状として Endocrine Subscale (以下、ES) SF-36, Menopause Specific Quality of Life Questionnaire の Vasomotor domain (以下、vaso) のスコアを評価した (図1-3)。

変量効果モデルによる検討結果では、身体症状全般、内分泌・更年期症状のいずれにおいても、AI 群と対照群の QOL スコアの変化量の標準化された平均値に有意差は認められなかった。さらに身体症状全般において対照群と AI 治療群の QOL 悪化・不変のリスクに有意差は認められなかった。

IV. 結論

本研究では、メタ・アナリシスの対象とした QOL サブ・プロトコルの中間解析の時点では、従来治療群に比較して AI 群の有意な QOL 向上は認められなかった。今後、最終結果が報告された際に、最終的な影響を評価したいと考える。

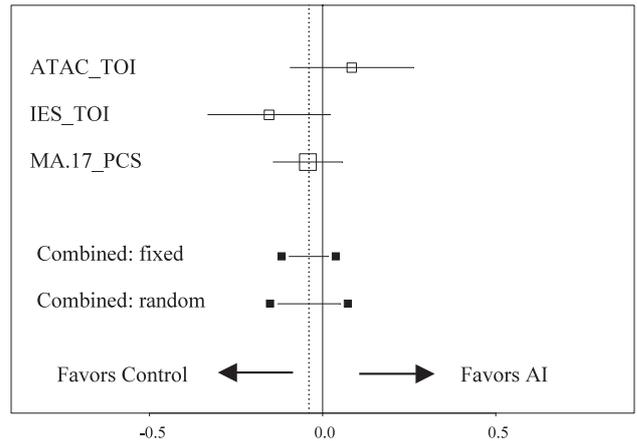


図 1. QOL スコアの変化量の標準化された平均値の差の統合結果 (身体症状全般)

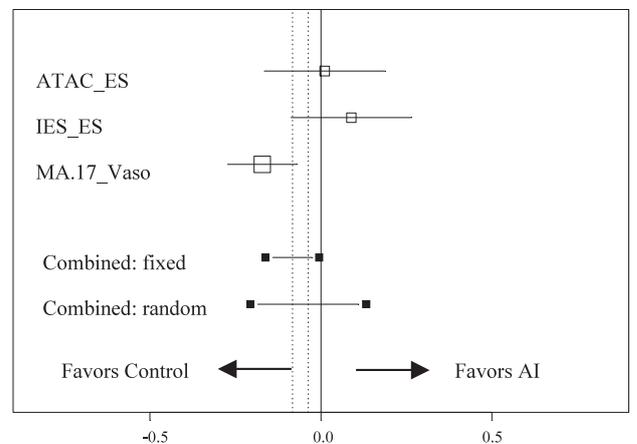


図 2. QOL スコアの変化量の標準化された平均値の差の統合結果 (内分泌・更年期症状)

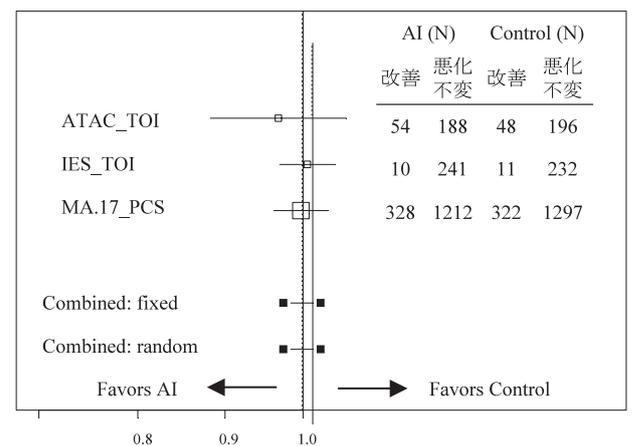


図 3. QOL 悪化・不変のリスク比の統合結果 (身体症状全般)

また、経時的な QOL の評価や、他の心理的、社会的、機能的な次元なども検討する必要があると思われる。

参考文献

1) 丹後俊郎. メタ・アナリシス入門. 東京: 朝倉書店; 2002.

- 2) DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986 Sep;7(3):177-88.
- 3) Cella D, Fallowfield L, Barker P, Cuzick J, Locker G, Howell A; ATAC Trialists's Group. Quality of life of postmenopausal women in the ATAC ("Arimidex", tamoxifen, alone or in combination) trial after completion of 5 years' adjuvant treatment for early breast cancer. *Breast Cancer Res Treat*. 2006 Dec;100(3):273-84.
- 4) Fallowfield LJ, Bliss JM, et al. Quality of life in the intergroup exemestane study: A randomized trial of exemestane versus continued tamoxifen after 2 to 3 years of tamoxifen in postmenopausal women with primary breast cancer. *Journal of Clinical Oncology* 2006;24:910-7.
- 5) Whelan TJ, Gross PE, et al. Assessment of quality of life in MA.17: A randomized, placebo-controlled trial of letrozole after 5 years of tamoxifen in postmenopausal women. *Journal of Clinical Oncology* 2005;23:6931-6940.

〈教育報告〉

平成19年度専門課程 II
生物統計分野

経時的繰り返し測定データに対する 個人プロフィールに基づく評価方法の研究

家中和浩

A Study of the Evaluation Method Based on the Individual Profile for Repeated Measures.

Kazuhiro IENAKA

Abstract

Objective There are two approaches of evaluation methods for repeated measures. One is a method based on the average profile of the treatment group and the other is a method based on the individual profiles. In this study, I clarify the characteristic of the method based on the individual profile by comparison with the one based on an average profile of the treatment group. **Method** Three true models are examined, and four methods of analysis are compared by Monte Carlo simulation, I compare the power among the evaluation methods based on the individual profile, (t-test for the simple linear regression coefficients of every individuals of treatment groups, t-test for AUC of every individuals of treatment groups), and the evaluation methods based on the average profile of the group, (t-test for the change from the baseline at the end, an analysis of variance). **Result** The power of t-test for the simple linear regression coefficients is higher than the others or as high as the best method in the case of all true models. **Conclusion** When a profile is monotonously decreasing (increasing), t-test for the simple linear regression coefficients is suggested as the best analysis

Keywords: clinical trial, repeated measures, a individual profile, an average profile of the group

Thesis Advisor: Masako NISHIKAWA, Kunihiko TAKAHASHI, Toshiro TANGO

I. 目的

繰り返し測定データを対象とした評価方法の比較に関する先行研究としては、Frison¹⁾らが、投与後平均、ベースライン測定値の平均と投与後測定値の平均の差、そして、ベースラインの平均値を共変量とした共分散分析の方法を比較している。Fitzmaurice²⁾らは個人プロフィールの要約指標としてAUC (Area Under the Curve, 曲線下面積) と回帰直線の傾きについてある程度の有用性を認めている。本研究では経時的繰り返し測定データを対象とし、集団の平均プロフィールに基づく評価方法との比較により、個人プロフィールに基づく評価方法の特徴をシミュレーションにより明らかにする。

II. 比較する解析方法

集団の平均プロフィールに基づく解析方法として、最終時点 (シミュレーションでは投与12週後) における投与前からの差を対象としたt検定 (以下、「12週後の差」と分散分析 (以下、「平均的な減少量」) を、個人プロフィールに基づく解析方法として、個人毎の回帰直線の傾きを対象としたt検定 (以下、「回帰直線の傾き」と個人毎のAUCを対象としたt検定 (以下、「AUC」) を比較することとした。

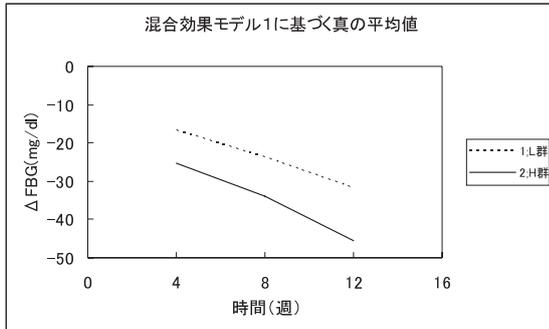
III. シミュレーション

シミュレーションは3つの真のモデルのもとに行った。3つのモデル (混合効果モデル1, 混合効果モデル2, 回帰モデル) いずれにおいても質的な交互作用がないこと、被験者を変量効果とし、各測定時点の誤差項は独立で同一分布を仮定した。3つのモデルのうち回帰モデルでは、被

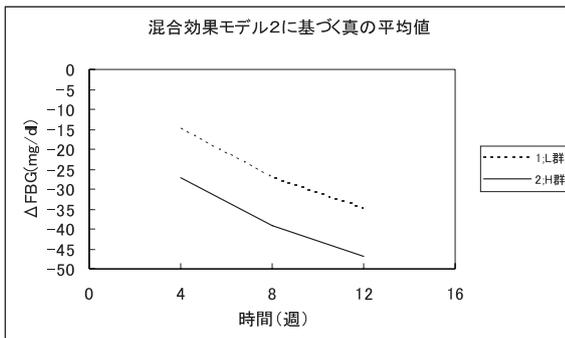
指導教官: 西川正子, 高橋邦彦, 丹後俊郎 (技術評価部)

験者の回帰直線についての変量効果は対数正規分布に従うことと、および各測定時点の誤差項は独立で同一分布を仮定した。従って、データのばらつきは、経時的に大きくなる。残りの2つの混合効果モデルでは被験者の変量効果は正規分布に従い、誤差は等分散であることを仮定した。実際の臨床データを元に真のモデルのパラメータを設定した。

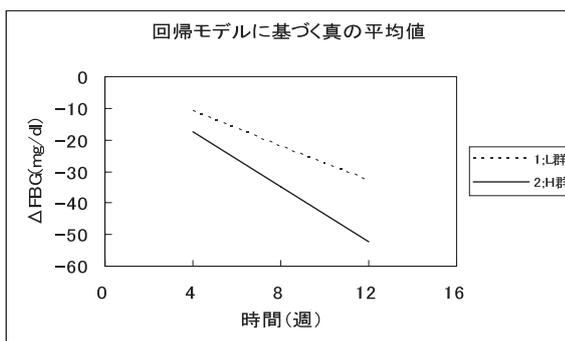
① 混合効果モデル 1



② 混合効果モデル 2



③ 回帰モデル



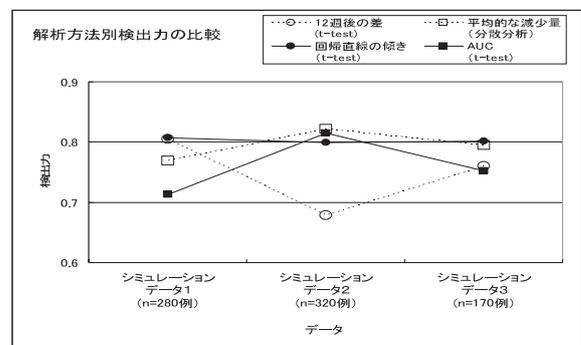
症例数の設計は、真のモデルに対して最適と考えられる解析手法を用いて、シミュレーションにより設定した。シミュレーションには SAS データステップにて正規乱数を使用して、その回数を1000回として、シミュレーションデータごとに設定した最適と考えられる解析手法により有意水準 両側 5 %、検出力80 %以上となる症例数を5例きざみで求めた。

真のモデルごとに正規乱数を使用して SAS データステップにてシミュレーションデータを発生させ、各解析手法による解析を実施した。このシミュレーション1000回において、有意水準 両側 5 %で検出される割合を検出力とした。

解析方法間の比較は、シミュレーションで仮定した真のモデルごとに、症例数算出に用いた解析手法による検出力を基準とし他の解析手法の検出力と比較した。

IV. 結果

真のモデルが混合効果モデル 1 (シミュレーションデータ 1) の場合、基準とする解析方法は「12週後の差」である。「12週後の差」の検出力は0.805であった。「回帰直線の傾き」の検出力は0.807であり、「12週後の差」の検出力との差は +0.002であった。真のモデルが混合効果モデル 2 (シミュレーションデータ 2) では、基準とする解析方法は「平均的な減少量」である。「平均的な減少量」の検出力は0.822であった。「AUC」の検出力は0.815であり、「平均的な減少量」の検出力との差は、-0.007であった。「回帰直線の傾き」の検出力は0.799であり、「平均的な減少量」の検出力との差は、-0.023であった。真のモデルが回帰モデル (シミュレーションデータ 3) の場合、基準とする解析方法は「回帰直線の傾き」である。「回帰直線の傾き」の検出力は0.802であった。「平均的な減少量」の検出力は0.795であり、「回帰直線の傾き」の検出力との差は -0.007であった。



V. 考察

検出力の標準誤差は約0.013であるため、真のモデルが混合効果モデル 1 の場合、「12週後の差」の検出力と「回帰直線の傾き」の検出力の差は有意ではなかった。真のモデルが混合効果モデル 2 の場合、「平均的な減少量」の検出力と「AUC」の検出力の差は有意ではなかった。「平均的な減少量」の検出力と「回帰直線の傾き」の検出力の差は、0.023と標準誤差に近いものであった。真のモデルが回帰モデルの場合、「回帰直線の傾き」の検出力と「平均的な減少量」の検出力の差は有意ではなかった。また、「平均的な減少量」での変量効果の共分散構造を複合対称

性に設定したため検出力は実際の状況よりやや高めになっていることが推測される。以上のことを考慮すると、今回のシミュレーションから「回帰直線の傾き」がもっとも好ましい解析方法という結果となった。これは真の平均値の推移が単調減少、また、時間推移が進むことにより二群間の差が大きくなる場合では、回帰直線が良く当てはまり、個人ごとの回帰直線の傾きのバラツキが小さくなることによるためと考えられる。

VI. まとめ

個人プロファイルに基づく評価方法である個人毎の回帰直線の傾きを対象とした t 検定が、今回設定したシミュ

レーションの状況（プロファイルが単調減少（増加）する、かつ、質的交互作用が無い）の場合、もっとも適切な解析手法であることが示唆された。

文献

- 1) Frison Lars, Pocok, Stuart J. Repeated measures in clinical trials : Analysis using mean summary statistics and its implications for design. *Statistics in Medicine* 1992;11: 1685-1704.
- 2) Fitzmaurice, Garrett M, Laird, Nan M, Ware, James H. *Applied longitudinal analysis*. Hoboken, N.J. : Wiley-Interscience; 2004.

〈教育報告〉

平成19年度専門課程 II
生物統計分野

ベイズの理論を用いた医薬品の評価方法の検討 - すでに提案されているがんの Phase II 試験のデザインの中止基準の検証 -

金子裕一郎

Consideration about Evaluation Method of New Drug Using Bayesian Theorem - Confirmation of Validity about Stopping Criteria Related to Phase II Trials' Design Which is Already Suggested-

Yuichiro KANEKO

Abstract

Background and Objective Phase II clinical trials in cancer is generally executed by single-arm design, and its objective is to evaluate safety and efficacy. Efficacy is evaluated by antitumor reduction effects (ex: binary response rate). Thall et al.¹⁾ (2003) proposed a statistical method including Bayesian posterior probability for conducting phase II trials in settings where the disease is categorized into multiple subtypes. And Thall set the same stopping criteria (d_L) and the same maximum sample size for all subtypes. But this d_L is not validated. In this study, I evaluated a d_L described by Thall by simulation where there are two subtypes. **Method** Using the statistical method proposed by Thall et al. (2003), I conducted simulation studies and evaluated d_L s from the view of false positive probability and false negative probability for various combinations of true response rate in each subtype and d_L . d_L s considered are 0.001, 0.005 and 0.01. **Result and Consideration** Although there is not any clear difference in false negative probability among three d_L s, false positive probability in $d_L=0.01$ is lower than that in $d_L=0.005$ or in $d_L=0.001$. As described so far, $d_L=0.01$ is more reasonable criteria than $d_L=0.001$ or $d_L=0.005$.

Keywords: phase II clinical trial in cancer, posterior probability, early stopping in futility

Thesis Advisor: Masako NISHIKAWA, Toshiro TANGO

I. 目的

がんの Phase II 試験は、他の臨床試験と異なり、通常1群無対照で行い、抗腫瘍縮小効果等の有効性と安全性の評価を目的としている。

Thall, et al.¹⁾は、Imatinib (グリベック) の臨床試験を例として、肉腫 (Sarcoma) 患者に対して10種類のサブタイプで薬効が異なることを考慮したがんの Phase II 試験の Bayes の事後確率を用いた解析手法を提案している。提案された方法では、サブタイプあたりの最大症例数を30例と設定し、中止基準については、各サブタイプごとに別々に、同一の中止基準を設定している。また、階層 Bayes モデルを仮定し、各サブタイプが、階層構造の下独立に同一の分布に従うことを仮定することで、別々に試験

するよりも、効率的であるということを報告している。

しかし、Thall, et al.¹⁾の論文では中止基準、最大症例数の妥当性については検証されていない。

本研究では Thall (2003) がすでに提案したがんの Phase II 試験の解析手法について、サブタイプが2つの場合の、中止基準の妥当性をシミュレーションにより検討する。中止基準の妥当性は真の応答率が有効である場合に、早期無効中止する確率 = 偽陰性率と、真の応答率が無効である場合に早期無効中止しない確率 = 偽陽性率で評価する。

II. 研究デザインと方法

2.1 シミュレーション設定条件

〈がん Phase II 試験〉

- ✓ デザイン：1群無対照試験で抗悪性腫瘍薬投与
- ✓ 主要評価項目：腫瘍縮小率 (以降“応答率”)

指導教官：西川正子、丹後俊郎 (技術評価部)

- ∨ サブタイプ：2 値バイオマーカー 1 変数
- ∨ 試験継続、中止を判断する期待応答率 $p^* : 0.30$
- ∨ 早期無効中止を判定するための最低症例数：2 つのサブタイプを合わせて20例
- ∨ 最大症例数：各サブタイプ30例
- ∨ サブタイプの真の構成比：1 : 1

無効中止か継続するかの判定は、各サブタイプの真の応答率 p_j ($j=1, 2$) が期待応答率 $p^* = 0.30$ を上回る事後確率 $\Pr(p_j > p^* | data)$ を計算し、

$$\Pr(p_j > p^* | data) < d_L$$

であれば無効中止と判定する。この時、Thall et al.¹⁾ の中止基準 $d_L = 0.005$ の根拠は明確ではない。シミュレーションで検討した2つのサブタイプの真の応答率と d_L の組み合わせを表1に示した。

サブタイプごとの中止確率は、サブタイプごとにいずれかの時点で早期無効中止になった回数をシミュレーション回数で除して算出した。サブタイプごとの平均症例数は、1回のシミュレーションごとにサブタイプごとに早期中止時点(30例未満)または終了になった時点(30例)までの例数を記録しその総和をシミュレーション回数で除して算出した。シミュレーションは各設定条件ごとに200回行った。

III. 結果

表1に各条件における早期無効中止確率と平均被験者数の結果を、表2に中止基準ごとの偽陽性率と偽陰性率の結果を示す。

IV. 考察

今回のシミュレーション結果では、3つの中止基準全体でも、偽陰性率が0~0.105の範囲であり、がんのPhase II 試験で設定される Type II error $\beta = 0.10$, $\beta = 0.20$ と比べても劣っておらず、有効な薬を検出する可能性は頻度論的な臨床試験と比べても高いと考えられる。そのため、偽陽性率で性能を比較すると、偽陽性率が0.020~0.185の範囲内にあり、他の2つに比べて最も低く保たれている $d_L = 0.01$ が適切であると考えられた。

V. まとめ

以上の偽陽性率、偽陰性率の結果から、今回のシミュレーションの条件下では、中止基準 $d_L = 0.01$ が、 $d_L = 0.005$

表1：各条件における早期無効中止確率、平均被験者数

サブタイプ1の真の応答率 p_1	サブタイプ2の真の応答率 p_2	中止基準 d_L	サブタイプ1の中止確率	サブタイプ2の中止確率	サブタイプ1の平均症例数(人)	サブタイプ2の平均症例数(人)
0.05	0.3	0.001	0.895	0.025	17.6	29.7
0.05	0.3	0.005	0.975	0.060	14.2	29.0
0.05	0.3	0.01	0.990	0.105	12.9	28.6
0.1	0.3	0.001	0.595	0.060	23.0	28.9
0.1	0.3	0.005	0.770	0.075	18.8	28.6
0.1	0.3	0.01	0.835	0.095	17.0	28.2
0.05	0.4	0.001	0.880	0	18.7	30.0
0.05	0.4	0.005	0.970	0.010	14.5	29.8
0.05	0.4	0.01	0.980	0.015	13.4	29.7
0.1	0.4	0.001	0.585	0	23.4	30.0
0.1	0.4	0.005	0.815	0.010	18.7	29.9
0.1	0.4	0.01	0.815	0.020	17.2	29.7

表2： (p_1, p_2) の組み合わせと偽陽性率、偽陰性率との関係について

中止基準	サブタイプ1の真の応答率 p_1	偽陽性率	サブタイプ2の真の応答率 p_2	偽陰性率
0.001	0.05	0.105	0.3	0.025
	0.1	0.405	0.3	0.060
	0.05	0.120	0.4	0.000
	0.1	0.415	0.4	0.000
0.005	0.05	0.025	0.3	0.060
	0.1	0.230	0.3	0.075
	0.05	0.030	0.4	0.010
	0.1	0.185	0.4	0.010
0.01	0.05	0.010	0.3	0.105
	0.1	0.165	0.3	0.095
	0.05	0.020	0.4	0.015
	0.1	0.185	0.4	0.020

や $d_L = 0.001$ に比べて、より妥当な中止基準であると考えられた。

文献

- 1) Thall PF, Wathen JK. Hierarchical Bayesian approaches to phase II trials in diseases with multiple subtypes. *Statistics in Medicine* 2003;22:763-780.

〈教育報告〉

平成19年度専門課程 II
生物統計分野

経時測定データを用いた治療効果の評価における 共変量を考慮した混合分布モデルの応用

田口奈緒子

A Latent Class Mixture Model Combined with Proportional Odds Model for Repeated Measurements in Clinical Trials

Naoko TAGUCHI

Abstract

For the analysis of longitudinal data in randomized controlled trials, generalized linear mixed-effects models or random coefficient models are known to provide a flexible and powerful tool to deal with heterogeneity among patients' profile. However, they are still based upon the "homogeneous" assumption in a sense that the effects of treatment is evaluated by the difference in mean response profiles. If subject \times time interactions within each treatment arm are not negligible, the problem is not so simple. To this problem, Tango (1989) proposed mixture models in which 1) each treatment arm consists of a mixture of several ordered latent classes of profile of change from baseline, e.g., "improved", "unchanged", "worsened", and so on, 2) a low-degree polynomial can represent the "mean profile" for each of latent classes and 3) the effects of treatment can be characterized by the difference in the mixing proportions of these latent classes. Tango (1998) extended his model so that it can cope with improper longitudinal records with missing values, which is cited as a new approach by Everitt and Pickles (2004). However, Tango did not take both covariate adjustments and the nature of "ordered" classes into account in evaluating the effects of treatment.

In this presentation, we propose a generalized model by incorporating Tango's model with proportional odds model to take both covariate adjustments and the nature of "ordered" classes into account. The proposed model is illustrated with data from some randomized controlled trial.

Keywords: randomized trials, mixture model, covariate, proportional odds model,

Thesis Advisor: Masako NISHIKAWA, Kunihiro TAKAHASHI, Toshiro TANGO

I. 目的

丹後¹⁾で提案された混合分布モデルを用いた薬効評価の方法を拡張し、共変量を調整した場合の混合分布モデルによる薬効評価の推定方法を提案する。

II. 方法

本研究には、丹後¹⁾に掲載されているグリチロン錠二号(当時)の臨床試験のデータを使用した。本試験は19施設で総数224例の慢性肝炎患者の二重盲検の無作為化比較試験である。本試験は、肝機能検査値 GPT (ALT, 単位は

mg/dL) によって、グリチロン錠投与群とプラセボ投与群の治療効果を比較することを目的としている。投与開始時、4週、8週、12週、16週の計5時点でデータが取られている。掲載されていた一群82例、計164例分の対数変換した GPT の変化量を用いた。対数変換した GPT 変化量の群別時点別男女別平均値を見ると男性のほうが女性に比べ、大きく低下していることがわかった。これにより性別による調整を試みることにした。

薬剤 $d (= 1, 2)$ に割り当てられた被験者 $j (= 1, 2, \dots, N_d)$ の測定時点 $t (= 0, 1, \dots, T)$ の検査値を X_{djt} とする。ここで、 X_{dj0} は投与開始時点の検査値である。ここで、投与開始時点からの変化量 $Y_{djt} = X_{djt} - X_{dj0}$ の変動パターンに対して、次のモデルを導入する。

指導教官： 西川正子, 高橋邦彦, 丹後俊郎 (技術評価部)

$$Y_{djt} = \mu(t) + \varepsilon_{djt} \quad \varepsilon_{djt} \sim N(0, \sigma^2) \quad t=1, \dots, T$$

ここで、 $\{\varepsilon_{djt}\}$ は互いに独立で、平均値の経時的変動 $\mu(t)$ は、被験者 j が「不変群」に属するとき $\mu(t) = \mu_0(t) = 0$ 、「改善群」に属すれば

$$\mu(t) = \mu_1(t) = \sum_{k=1}^{R_1} \beta_{1k} t^k, \text{ 「悪化群」に属すれば}$$

$$\mu(t) = \mu_2(t) = \sum_{k=1}^{R_2} \beta_{2k} t^k \text{ としている.}$$

各群の確率密度関数は、多変量正規分布

$$f_j(Y_{dj}) = \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{T}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^T (Y_{djt} - \mu_j(t))^2 \right\} \text{ で与えられる.}$$

被験者 j の経時変動パターンを、以下の混合分布モデルで表現する.

$$h(Y_{dj}|\theta) = p_0 f_0(Y_{dj}) + p_1 f_1(Y_{dj}) + p_2 f_2(Y_{dj})$$

この混合分布モデルに対し $q_0 = p_1$, $q_1 = p_1 + p_0$ とした次の比例オッズモデルを考える.

$$\log \left(\frac{q_0}{1-q_0} \right) = \theta_0 + \gamma_1 x_1 + \gamma_2 x_2, \log \left(\frac{q_1}{1-q_1} \right) = \theta_1 + \gamma_1 x_1 + \gamma_2 x_2$$

治療効果を表す γ_1 、共変量の効果を表す γ_2 は q_0 , q_1 によらず一定である。治療効果は「 $H_0: \gamma_1 = 0$ 」, 「 $H_1: H_0$ ではない」という仮説で検証する。さらに共変量の効果についての仮説, 「 $H'_0: \gamma_2 = 0$ 」, 「 $H'_1: H'_0$ ではない」についても検討する。パラメータ推定は尤度 L' に基づく最尤法で推定する.

$$L' = \prod_{d=1}^2 \prod_{j=1}^{N_d} h(Y_{dj}|\lambda')$$

III. 結果

治療群と性別を共変量とした場合の、パラメータ推定結果は以下ようになった.

$$\begin{aligned} \hat{\beta}_{11} &= -0.944(0.099), \hat{\beta}_{12} = 0.140(0.024) \\ \hat{\beta}_{21} &= 0.872(0.203), \hat{\beta}_{22} = -0.183(0.050) \\ \hat{\theta}_0 &= -0.550(0.281), \hat{\theta}_1 = 5.157(0.916) \\ \hat{\gamma}_1 &= -1.814(0.590), \hat{\gamma}_2 = -1.692(0.799) \\ \hat{\sigma}^2 &= 0.208(0.013) \end{aligned}$$

この結果、グリチロン錠群のプラセボ群に対する治療効果は有意に優れ ($p = 0.0021$, Wald 検定), 改善オッズ比は 6.13 (95% 信頼区間: 1.93-19.50) と推定された。また、性別に有意な違いがあることが認められ ($p = 0.034$, Wald 検定), 男性の女性に対する改善オッズ比が 5.43 倍 (95%

信頼区間: 1.13-26.00) と推定された.

IV. 考察

本研究では、治療を含めた共変量の効果の推定に比例オッズモデルを用いた。これは治療と共変量に関するオッズ比が順序カテゴリーの区切り方に無関係に一定である、という仮定をおいていることになる。この仮定は一見して、成立困難な条件とも思われるが、成立する場合も少なくなく、順序カテゴリーデータの解析に比例オッズモデルが提案され利用されている⁵⁾。本研究では、新しい方法の定式化を主な目的としたため、比例オッズ性については検討しなかったが、今後の研究課題としたい。

V. まとめ

本研究では、臨床研究における経時測定データに対し、丹後のモデルをベースにして、比例オッズモデルを用いて共変量を調整した新しい混合分布モデルの提案を行った。適用例として用いたグリチロン錠二号のプラセボ対照二重盲検無作為化比較試験のデータに適用した結果、1) グリチロン錠二号の改善オッズは 6.13 (95% 信頼区間: 1.93-19.50), 2) 男性の女性に対する改善オッズは 5.43 (95% 信頼区間: 1.13-26.00) と推定され、新しい知見が得られた。

参考文献

- 1) 丹後俊郎. 臨床試験における経時的測定データ解析のための混合分布モデル. 応用統計学. 1989; 18(3): 143-161.
- 2) Tango T. A mixture model to classify individual profiles of repeated measurements. In: Data science, classification and related methods. Hayashi, et al. (eds.) . Tokyo:Springer-Verlag;1998. p.247-254.
- 3) Everitt BS, Pickles A. Statistical aspects of the design and analysis of clinical trials, revised edition. London; Imperial College Press: 2004.
- 4) 丹後俊郎, 編. 統計モデル入門. 東京: 朝倉書店; 2000.
- 5) 丹後俊郎, 山岡和枝, 高木晴良, 編. ロジスティック回帰分析. 東京: 朝倉書店; 1996.
- 6) 岩崎学, 編. 統計的データ解析のための数値計算入門. 東京: 朝倉書店; 2004.

〈教育報告〉

平成19年度専門課程 II
生物統計分野

複合形質遺伝子座の統計学的検出法

西山毅

Statistical Detection Methodologies for Complex Trait Loci

Takeshi NISIYAMA

Abstract

Objective We compared the accuracy of a chi-square test to two Bayesian model, the hierarchical logistic model and the stochastic search variable selection (SSVS) model in a genetic association study to search for the disease susceptible gene. **Methods** To compare the accuracy, receiver operating characteristic (ROC) analyses were implemented using simulated data with varying levels of genotype relative risks (GRR) and minor allele frequencies (MAF), generated via vast amount of International HapMap Project genotype data. **Results** The hierarchical logistic model were superior to the other methods for a data with MAF less than five percent and GRR less than 1.6. Otherwise, all methods used were almost equal at the accuracy. **Conclusion** The hierarchical Bayesian model is suggested to have advantage of a genetic association study, in particular, fine-mapping of complex disease over the other method used.

Keywords: complex disease, single nucleotide polymorphism (SNP), Linkage disequilibrium (LD), Bayesian statistics, receiver operating characteristic curve (ROC curve)

Thesis Advisor: Kunihiko TAKAHASHI, Tetsuji YOKOYAMA, Kunihisa KISHINO, Toshiro TANGO

I. 目的

複数の遺伝要因と環境要因により発症する複合疾患の病因遺伝子探索には、主に関連研究が用いられる。そして、関連研究ではゲノム上の SNP と疾患との関連を調べる際に χ^2 乗検定が使われることが多い。これは、サンプル数よりも調べる SNP 数の方が多いため等の理由によるが、非常に多くの SNP 間に連鎖不平衡 (linkage disequilibrium: LD) が存在するため、多重比較の補正は困難である。そこで、ベイズ統計学に依拠した解析法が考案された。しかし、頻度論とベイズ統計学に基づく両者について詳しく比較検討はなされていないため、本研究では現実的なシナリオに基づくシミュレーションによって両者を比較した。

II. 方法

1. 比較する統計解析法

現在関連研究で主に使われている、遺伝子型に基づく χ^2 乗検定 (図 1) とアレルに基づく χ^2 乗検定と、Xu¹⁾

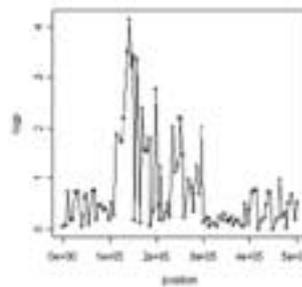


図 1. SNP のゲノム上の位置 (bp) に対する $-\log(P$ 値) プロット例

らによる階層ベイズロジスティックモデル, Xu¹⁾・Swartz²⁾らによる SSVS (Stochastic Search Variable Selection) モデルとを比較した。この SSVS モデルは、階層ベイズロジスティックモデルを改良したもので、変数選択の要素が加味されている。

2. シミュレーションの設定

まず有病率を 1% に決め、病因 SNP の minor アレル頻度 (minor allele frequency: MAF) として 35~45%・10~20%・1~5% を、病因 SNP の GRR (Genotype Risk Ratio) として 2.0・1.6・1.2 を選べば (それぞれ高 MAF・中 MAF・低 MAF, 高 GRR・中 GRR・低 GRR と呼ぶ)、症例・対

指導教官：高橋邦彦，横山徹爾，丹後俊郎（技術評価部）
岸野洋久（東京大学）

照群の各遺伝子型頻度が定まる。そこで、それに応じて国際 HapMap 計画³⁾の日本人および漢民族系中国人の併合集団のハプロタイプを resampling することによってシミュレーションデータを各シナリオごとに50個ずつ作成した。この際、症例・対照のサンプル数はそれぞれ100人とし、SNP 数は LD のない場合とある場合の2通りを考え、前者では108個、後者では100個の SNP を選び、この中から病因 SNP として前者では5つ、後者では3つを定めた。

3. シミュレーションの評価法

χ^2 乗検定については、個々の SNP の $-\log(P$ 値) を、ベイズモデルについては、回帰係数の事後平均値の絶対値を指標として ROC 解析を行った。そして、その ROC 曲線の曲線下面積の平均値と標準誤差によって解析法の評価を行った (図2)。

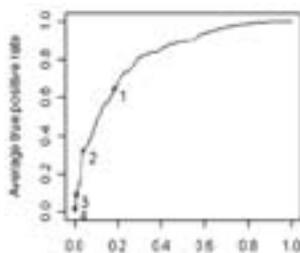


図2. χ^2 乗検定による $-\log(P$ 値) を指標にした ROC 曲線の例

III. 結果

LD がない場合、4つの解析法は概ね同等であった。LD が存在する場合も、4つの解析法は概ね同等であったが、低 MAF・中 GRR/低 GRR の場合については階層ベイズロジスティック解析が他の3つの解析法より優れており、特に低 MAF・低 GRR においては他の解析法よりも平均曲線下面積は約0.1高かった。

IV. 考察

従来のベイズ解析では、ベイズ統計学の枠内でモデルの優劣を比較しており χ^2 乗検定との比較はほとんど行われ

ていない。そこで本研究では χ^2 乗検定と2種類のベイズ解析法との比較をシミュレーションで行った。その結果、LD が存在する場合には低 MAF かつ中 GRR・低 GRR において階層ベイズロジスティック解析法の成績は他と比べ著明に優れていたが、それ以外では4つの解析法は概ね同等であった。ただし、本研究ではサンプルサイズは症例100人・対照100人に固定していたため、サンプル数が結果に及ぼす影響については不明である。実際の関連研究では数百~数千人ものサンプルを用いることを考えると、今後サンプル数を増やした場合の影響を検証する必要がある。Swartzら¹⁾は、国際 HapMap 計画より入手可能な SNP 間の LD 係数を事前情報に組み込むことによって SSVS モデルの改良を提案しているが、これも今後の検討課題である。

V. まとめ

関連研究に向け提案されたベイズ解析法と現在主に用いられる χ^2 乗検定との比較は従来行われておらず、本研究では病因 SNP のアレル頻度と GRR の様々な組み合わせに対して両者の比較を行った。LD が存在し、病因 SNP のアレル頻度が非常に低く、その GRR も比較的小さい場合は、階層ベイズロジスティック解析法が χ^2 乗検定や SSVS 法より優れていた。それ以外では、いずれの解析法の成績も概ね同等であった。

参考文献

- 1) Xu S. An empirical Bayes method for estimating epistatic effects of quantitative trait loci. *Biometrics* 2007; 63:513-21.
- 2) Swartz MD, Kimmel M, Mueller P, Amos CI, Stochastic search gene suggestion: a Bayesian hierarchical model for gene mapping. *Biometrics* 2006; 62:495-503.
- 3) International HapMap Project. Available from :<http://www.hapmap.org/index.html> ja

〈教育報告〉

平成19年度専門課程 II
生物統計分野

集団における栄養学的リスク者割合の新しい推定法

横道洋司

A New Method to Estimate the Proportion of Nutritionally-at-risk Individuals in a Population

Hiroshi YOKOMICHI

Abstract

To prevent lifestyle-related diseases or nutrient-deficiency, it is important to assess the proportion of high-risk individuals whose usual intake of certain nutrient exceeds (or is short of) a recommended reference value in a population. For that purpose, a dietary survey is sometimes conducted in the population and the individual's intakes of nutrients are measured. However, it is insufficient to simply calculate the proportion of individuals whose dietary intakes during a few days exceed (or are short of) a reference value to estimate the proportion because such intakes include a large day-to-day variation (within-person variation). Analysis of variance (ANOVA) is sometimes used to estimate the proportion considering the within-person variation, but the accuracy of estimation could be small when the analysis is done by subgroups of sex and age where the sample size in each group is small. To improve this problem, we propose a new method to estimate the proportion of nutritionally-at-risk population based on a mixed-effects model. A simulation study showed that the proposed model more accurately estimated the proportion of high-risk individuals and hence was more useful as compared to the traditional ANOVA method.

Keywords: nutrition survey, usual intake, nutritionally-at-risk population, variance components, age

Thesis Advisor: Tetsuji YOKOYAMA, Kunihiko TAKAHASHI, Toshiro TANGO

I. 背景

個人が、ある栄養素 E について、習慣的に 1 日あたり摂取している量をその個人の習慣的摂取量というが、栄養調査においては、対象集団の習慣的摂取量の分布（ヒストグラム）を推定することや¹⁾、その栄養素の摂取量の目安となる基準値 M について、“対象集団の何%の人間が、「習慣的な 1 日あたり摂取量」が M を超えているか、または M に不足しているか”を見積もることによって、その集団における栄養素 E の摂取状況を評価することが多い^{2,3)}。この、習慣的摂取量が基準値を超える者をリスク者と記すことにする。このリスク者割合 Q の推定を考える。

II. 方法

$E(Y_{jk} | Age_j) = a + b \times Age_j + c \times Age_j^2$
と仮定した次のモデルを考える：

$$Y_{jk} = a + b \times Age_j + c \times Age_j^2 + \alpha_j + \varepsilon_{jk}$$

$$\alpha_j \sim N(0, \exp(l + m \times Age_j)),$$

$$\varepsilon_{jk} \sim N(0, \exp(p + q \times Age_j))$$

すると、 Age_j 歳の個体 j がリスク者である確率

$$Q(Age_j) = \Pr\{E(Y_{jk} | Age_j) > g(M)\}$$

$$= 1 - \Phi\left(\frac{g(M) - (a + b \times Age_j + c \times Age_j^2)}{\sqrt{\exp\left(\frac{1}{2}l + \frac{1}{2}m \times Age_j\right)}}\right)$$

である。

これらパラメータ a, b, c, l, m, p, q は最尤推定量を求めることとし、次にこれらのパラメータ推定値 $\hat{a}, \hat{b}, \hat{c}, \hat{l}, \hat{m}, \hat{p}, \hat{q}$ を用いて、 Age_j 歳の個体 j がリスク者である確率 $Q(Age_j)$ を推定する。

指導教官： 横山徹爾，高橋邦彦，丹後俊郎（技術評価部）

III. シミュレーションによる比較

年齢階級別に平均値、個体間・個体内分散が一定値をとるとした正規乱数からの階段型シミュレーションデータと、平均値と個体間・個体内分散がそれぞれ連続的に変化するものとした正規乱数からの連続型シミュレーションデータについて、標本数を1500, 1000, 500としてシミュレーションにより分散分析法と新法による Q の推定を1000回行い比較した。それぞれのデータ型に用いる平均値、個体間・個体内分散は吉池・石脇らの行った12日間の繰り返し食事調査³⁾から推定されたものである。

(表1)女性 連続型データにおける Q の比較

%	全対象者	18~29歳	30~49歳	50~69歳	70~85歳
真値	85.0	62.9	84.2	91.0	89.1
分散分析法					
3日					
推定値平均	85.0	63.2	84.3	90.9	89.4
バイアス	0.06	0.24	0.13	-0.05	0.25
標準誤差	0.98	3.88	2.07	1.31	1.92
$\sqrt{\text{MSE}}$	0.98	3.89	2.08	1.31	1.93
新法3日					
推定値平均	85.1	63.0	84.2	91.0	89.3
バイアス	0.08	0.05	0.08	0.05	0.17
標準誤差	1.20	3.35	1.52	1.15	1.94
$\sqrt{\text{MSE}}$	1.20	3.35	1.52	1.15	1.94
分散分析法					
2日					
推定値平均	85.1	63.8	84.4	91.2	89.4
バイアス	0.14	0.90	0.24	0.23	0.27
標準誤差	1.33	5.93	3.16	1.88	2.61
$\sqrt{\text{MSE}}$	1.34	6.00	3.17	1.90	2.62
新法2日					
推定値平均	85.1	62.9	84.4	91.1	89.3
バイアス	0.16	-0.05	0.23	0.16	0.17
標準誤差	1.59	4.26	2.26	1.50	2.47
$\sqrt{\text{MSE}}$	1.60	4.26	2.27	1.50	2.47

IV. 考察

Q の推定精度において(表1)1500人の連続型データでは、男女を問わず分散分析法より新法の推定精度の良さが際立った。階段型データで有利となることが予想された分散分析法が新法に劣らない程度の推定精度しかもたないことから、この比較において新法が優れていることが示唆された。

実際の栄養調査データに対するモデルの適合については、栄養摂取量という世代を連続的に伝播する嗜好が根源

となるデータであるだけに、その習慣的摂取量の平均値や分散が連続的に変化しない例のほうが本当は稀であると考えられる。年齢階級はそういった個人の嗜好を世代別に分けたものに過ぎず、世代の変化と独立に年齢階級の平均値や分散が変わることが現実に起こり難いと考えられるからである。

つぎに、標本数を小さくして両者の推定の精度を検討したところ、標本数が小さいほど両者モデルの推定精度は低下していたが、ここでも新法はおおむね分散分析法に優っていると考えられた。

本稿シミュレーションで両者の方法の推定精度に差が出にくかった理由としては、標本数1500は比較的大標本であったことと、食塩の目標量が比較的习惯的摂取量の中央値に近く、各パラメータの推定精度が Q の推定精度に与える影響が少なかったことが挙げられよう。

今後の課題として、本研究の食塩だけでなく、他の栄養素および他の調査の場合についても広く検討し、新法のモデル適合が良くない栄養素についても両者の方法を比較する必要がある。

V. まとめ

栄養素の習慣的摂取量リスク者割合の推定について、年齢階級別の解析から離れ大局的に記述する方法を提案した。食塩について現実の栄養調査に近いシミュレーションデータを作成し、従来の分散分析法と新法によるリスク者割合の推定精度を比較したところ新法の十分な優越性を見た。今後ほかの栄養素についての検討や様々な性質をもつ栄養素の習慣的摂取量に対応可能なモデルへのさらなる改善が望まれる。

参考文献

- 1) National Research Council, Subcommittee on Criteria for Dietary Evaluation. Nutrient adequacy: In: Assessment using food consumption surveys. Washington, DC: National Academy Press; 1986.
- 2) Nusser SM. Semiparametric transformation approach to estimating usual daily intake distributions. Journal of American Statistical Association 1996;91:1440-9.
- 3) Ishiwaki Asako. A statistical approach for estimating the distribution of usual dietary intake to assess nutritionally at-risk populations based on the new Japanese Dietary Intakes (DRIs). J Nutr Sci Vitaminol 2007;52:337-44.