

## ケースコントロールデザインを想定したシミュレーションデータの発生方法

### 1 はじめに

現在、日本人に対する前立腺がん発症と血中エクオール濃度の関係を調べるケースコントロール研究が計画されている。資源が限られているなどの理由から、この研究のデザインには予め決めておいたサンプルサイズの情報一度にすべて入手し1回のみ解析を行う固定デザインよりも、少数ずつの検体を測定していきその都度解析を繰り返して行う逐次モニタリングデザインを用いた方が研究終了までに必要なサンプルサイズの減少を期待でき、より好ましいといえる。ところが1例ずつではなく複数例ずつ解析する場合、研究が終了するまでに必要な期待サンプルサイズ (Average Sample Number; ASN) がどの程度増大するのかが分かっていない。よって本研究では、先述したケースコントロール研究で得られるデータを想定したシミュレーションデータを発生させ、解析ごとの例数を増やしたときの期待サンプルサイズの増大を確認する。

### 2 シミュレーションデータの発生方法

ケースコントロールデザインを想定したシミュレーションデータの発生方法を述べる。

観察研究では、既知・未知のリスク要因のアンバランスにより、曝露と疾患の関連が交絡されている状況が常である。そこで本研究では、(1) 年齢のような既知のリスク要因を直接測定、解析で調整する状況と、(2) 未知、あるいは測定困難なリスク要因を代替すると期待される、居住地域のような変数をデザインに用いる状況を考える。(1) のリスク要因には、サンプリング集団において各層内に十分なケース数・コントロール数を確保できるだけの限られた水準を想定し、(a) 独立サンプリング、(b) 水準内マッチングの2通りのコントロールのサンプリング法を考え、これらが層別解析の結果に与える影響を効率、検出力、あるいは期待サンプルサイズから検討する。一方、(2) のような非常に水準の多い変数の調整を前提としたコントロールサンプリングでは、マッチングとそれを考慮した解析が必要となり、今回は Mantel-Haentzel の方法で併合した条件付きオッズ比を求めることを想定している。(1) と (2) を組み合わせた状況においても、(2) と同様のサンプリングおよび解析方法を用いることになる。

### 3 境界アプローチを用いた逐次モニタリング法

境界アプローチは今回のケースコントロール研究のように中間解析の回数や時点を試験開始後に変更する必要がなく、さらに1例ごとに検定を行える状況では、連続モニタリングの下で ASN を最小にできる。三角検定は境界アプローチの代表的な手法であり、PEST でも容易に実行可能である。

PEST ではデザイン、モニタリングの他にシミュレーション実験を行うことができる。しかしシミュレーションデータの発生方法が不明であるため、例えば PEST で三角検定を行ったときの期待サンプルサイズと、同じデータを用いて SAS で群逐次検定を行ったときの期待サンプルサイズを比較することはできない。さらに PEST で発生させているシミュレーションデータや、これまでに行われているマッチングを行った場合のケースコントロール研究における逐次検定<sup>1,5</sup> に用いられているシミュレーションデータは、各層内での曝露割合が一定であることを仮定している。本研究では各層で曝露割合が異なる

ような状況を考えるため、PEST のシミュレーションは使用できない。また PEST では自分で発生させたデータに対するシミュレーション実験は困難である。そのため PEST を用いず SAS で三角検定を行いたいと考えている。

#### 4 今後の予定

まず SAS で三角検定を行うプログラムを書き、その後マッチングを行わない場合のシミュレーションデータに対する結果を SAS と PEST で比較し、同じ結果であるかどうかを調べる。その後、マッチング因子の水準間で曝露割合が異なるシミュレーションデータに対して三角検定を行い、複数例ずつ解析する場合に期待サンプルサイズがどの程度増大するのかを確認する。

さらに可能であれば、マッチングを行わない場合のデータとどのような結果の違いが表れるかを確認したい。また SAS でマッチングを考慮した群逐次検定のデザインと検定を行う方法を考え、三角検定や群逐次検定の手法を期待サンプルサイズなどにより性能比較したいと考えている。

#### 参考文献

1. Baksh MF, Todd S, Whitehead J, Lucini MM. Design considerations in the sequential analysis of matched case-control data. *Stat Med.* 2005; **24**: 853-67.
2. Whitehead J, Stratton I. Group sequential trials with triangular continuous regions. *Biometrics.* 1983; **39**: 227-36.
3. Siegmund D. Corrected diffusion approximations in certain random walk problems. *Advances in Applied Probability.* 1979; **11**: 701-19.
4. Whitehead J. *The Design and Analysis of Sequential Clinical Trials, Revised Second Edition.* Chichester: Wiley; 1997.
5. Van der Tweel I, Van Noord PAH. Sequential analysis of matched dichotomous data from prospective case-control studies. *Stat Med.* 2000; **19**: 3449-64.
6. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology Third Edition.* Philadelphia: Lippincott Williams & Wilkins; 2008.