

科学研究費助成事業 研究成果報告書

平成 26 年 5 月 15 日現在

機関番号：14401

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23500349

研究課題名(和文) 臨床試験における探索的研究のデザインと統計的推測法の研究

研究課題名(英文) Study on reseach design and statistical inferences for exploratory clinical researches

研究代表者

上坂 浩之 (Uesaka, Hiroyuki)

大阪大学・臨床医工学融合研究教育センター・招聘教授

研究者番号：60446250

交付決定額(研究期間全体)：(直接経費) 4,000,000円、(間接経費) 1,200,000円

研究成果の概要(和文)：臨床試験による治療法の有効性評価のための仮説検定において、エフェクトサイズに関する情報がきわめて乏しく、かつ臨床的に意味のある最小の効果に関する情報もない状況で、被験者数の再推定を含む2相3段階デザインを導入し、その性能を評価した。真のエフェクトサイズのもとでの、2相3段階デザインの平均被験者数は単一段階デザインにおけるそれより小さく検出力は同デザインの85～95%を示し、本方式の有用性が明らかとなった。

また、複数の治療やエンドポイントがある場合の多重仮説の検定において、種々の仮説構造の下での調整p値の陽表現、及び複数の独立仮説系列で階層構造を有する場合の調整p値の一般公式の陽表現を与えた。

研究成果の概要(英文)：We developed two-phase and three-stage designs for a situation where information on sample size estimation was very scarce. In these designs sample size is re-estimated after the first phase. The second phase consists of two stages and one interim analysis is performed. Sample size is not re-estimated at this phase. Theses designs were proved to give about 85 to 95 percent of the power of the single stage design and much less average sample sizes than that design.

We also investigated multiple testing problems. When we wish to test a family of several elementary hypotheses which are structured according to the importance and logical relationships among them. We obtained explicit formulae of adjusted p-values of each test of hypothesis. Furthermore, we defined multiple parallel /fixed sequence hypothesis structure, which consists of several ordered treatments and equally important multiple endpoints. We also developed explicit formulae for the adjusted p-values of elementary hypotheses.

研究分野：情報学

科研費の分科・細目：情報学基礎・統計科学

キーワード：臨床試験デザイン sample size reestimation adaptive design multiple comparison gatekeeping procedure

1. 研究開始当初の背景

(1) 新治療の開発における初期段階ではサンプルサイズを設定するために必要な情報が極めて乏しいのが通常である。また、検証段階においても探索段階の試験と検証試験ではエンドポイントが異なり、サンプルサイズの決定に際して既存情報の活用が困難な場合もある。医師主導型臨床試験では、探索から検証へと段階的研究が困難なため、既存情報が乏しい状況であっても検証的立場で試験が実施される場合も多い。これらの状況下では被験者数再推定デザインや適応型デザインが有用とみられており、多くの研究がなされているが、被験者数の初期設定や検出力、平均被験者数等に関しての情報が乏しく適用にあたっての指針が無いと言ってよい状況であった。

(2) 探索試験のみならず検証試験においても複数の治療群と複数のエンドポイントが存在することは多い。複数の仮説の検定全体での第1種の過誤を維持しつつ、誤りである帰無仮説については可能な限り多数の検定が有意とできる方式は、試験から得られる情報を最大限取り出すという意味において重要である。このような状況に対処するために gatekeeping procedure をはじめとして幾つかの方法が提案されているが、その実地適用の仕方と検定の p 値の求め方については十分な研究がなされていなかった。そのために、多くの実務統計家にとっては、これらの方法を適用することが困難な状況であった。

2. 研究の目的

(1) 探索段階か検証段階かを問わず、エフェクトサイズに関する情報が極めて乏しい状況下で、エフェクトサイズの暫定的な初期値を与え、試験の中間情報に基づきエフェクトサイズを再推定し、効果の有無をできる限り効果的に評価する適応型デザインを考案する。この際被験者数を最小にすること、可能な限り高い検出力を有すること、デザインが煩雑でないこと、偏りの介入を可能な限り小さくすることを基本とする。あわせて試験デザインの性能を様々な側面から評価し、臨床試験デザインと被験者数の設定の指針を与える。

(2) 多重仮説の検定に関しては、多重仮説の発生する個々の状況における仮説の構造化の具体的な事例を列挙し、それらにおける多重推測の手順と調整 P 値の公式を陽の形で与える。これにより、多重仮説検定の詳細な理論を知らない実務統計家が正しく多重仮説を構造化し、検定を実施し、p 値を求められるようにする。

3. 研究の方法

研究代表者と連携研究者1名、研究協力者1名で、2つの研究目的を分担して担当して、3名で定期的に勉強会を開催し、議論しながら研究を進めた。

(1) に関しては、研究開始にあたって、被験者数再推定に関するほとんど全ての文献、適応型計画の基本文献および最近の文献について方法論の背景となる考え方や数理を調査した。そのうえで、新たな視点から試験デザインを構築し、その性能を種々の指標を用いて評価した。第1種の過誤確率、検出力、平均被験者数などは、全て数値積分によって正確に計算した。数値計算には S-PLUS7.0J および 8.2J を用いた。

(2) に関しても同様に既存文献を調査した。これらの文献では多重推測の構造化に基づく推測の一般論の記述は多いものの、現実に遭遇する場面の記述は乏しく、企業の統計家が個々の場面で仮説を構造化し、その構造の下での検定結果を具体的に p 値で表現できるほどの具体性が無いことが明らかとなった。そこで、現実に遭遇する多重推測の問題を列挙しそれに対する p 値公式を導くことが有意義だと考え、個々の構造の下での各基本仮説の p 値を陽の形式で与えることとした。さらに、投与量などで定義された複数の治療法と複数のエンドポイントがあり、各エンドポイントに関する検定において投与量間では予め検定の順序が定められており、全てのエンドポイントは同等の重要性を有する、固定順序仮説系列の平行な仮説群からなる仮説構造に対する検定の一般理論の構築が可能であることが、この研究の過程において明らかとなった。そこで、この仮説構造における基本仮説の p 値公式も陽の形式で与えることとした。

以上の(1)および(2)に関する成果は、途中経過も含め国際学会および国内の学会で報告し、そこでの議論を理論の精緻化と評価に活かした。

4. 研究成果

(1) 新治療法の既存治療あるいは擬治療に対する優越性試験における2相3段階デザインを考案した。このデザインは、試験を大きく2相に分け、第1相では比較的楽観的なエフェクトサイズの下で、十分高い検出力を確保できるように被験者数を設定する。第1相終了後、検定を実施し有意であれば試験は有効中止、p 値が予め定めた値より大きければ無効中止とする。いずれでもなければ観測されたエフェクトサイズを用いて被験者数を再推定する。第2相は2段階デザインとするが、中間解析では無効中止、

有効中止、または継続のいずれであるかを判断するが被験者数の再推定は行わない。第2相を2段階としたのは、第1相でのエフェクトサイズの推定値が過大評価である場合や過小評価である場合には中間評価で結論を出せるようにすることで、無駄な試験の続行を回避できるからであり、被験者数の再推定を実施しないのは、再推定を繰り返すことによる偏りの介入を避けるとともに、探索的性格をできる限り抑制するためである。このようにして、「(2)研究の目的」の項に述べた、できるだけ煩雑さを避け、偏りの介入を可能な限り防止し被験者数を最小限にするために2回の中間解析を許容し、エフェクトサイズに関する情報の欠如を補うための被験者数の再推定を1回のみ許容するデザインを構成した。なお試験全体の第1種の過誤確率は予め設定するが、その大きさは試験が探索的試験か検証試験とみなすかに応じて設定することが可能である。新たに構成したデザインはBauer-Köhneの2段階適応型計画を第1相と第2相に適用し第2相では2段階群逐次計画を適用する方式(hybrid design)と、有意水準を第1相と第2相に分割して割り当てる方式(-split design)である。これらの方式の他に既存のBauer-Köhne-Wassmerの3段階デザイン、Lehmacher-Wassmerの適応型群逐次デザインを2相3段階デザインとして用いた場合も考察した。既存のデザインでは第1段階を第1相と位置付け、第2、第3段階を第2相に位置づけることで、2相3段階デザインとした。群逐次デザインではWang-Tsiatisの冪型境界を用いて、最適境界、 Pocock 境界、 $0'$ Brien-Fleming 境界の比較を行った。各中間解析での無効中止基準としてはDeMets-Wareの非対称境界を用いた。被験者数の再推定において、観測したエフェクトサイズを用いると検出力の損失が生じるため、保守的な推定値として上坂(計量生物学 2003)の方法を用いた場合も検討した。

デザインの性能は、試験全体の検出力と平均被験者数、検出力と平均被験者数の比、第2相における条件付き検出力や条件付き平均被験者数などを検討したが、全体の検出力と平均被験者数に基づき、デザイン間の比較の結果を要約すると以下のごとくである。(ア)群逐次デザインとして $0'$ Brien-Fleming 境界を用いる場合は平均被験者数が他より大きく、また検出力はエフェクトサイズに依存して高い場合と低い場合が認められる。(イ) -split design は、状況によっては他のデザインに比し検出力が低下する場合がある。(ウ)

Hybrid-design と最適境界の適応型群逐次デザイン、Bauer-Köhne-Wassmer デザインは類似した性能を示した。(エ)検討したデザインパラメータの下では、真のエフェクトサイズの下でのサンプルサイズを最大サンプルサイズとした2相3段階デザインは、単一段階デザインの85ないし95パーセント前後の検出力を確保し、平均被験者数は大きく低下した。(オ)保守的なエフェクトサイズ推定値による検出力の向上は数パーセント以内であった。

(2)基本仮説数が2個の場合、3個の場合、4個の場合に関して生じうる様々な仮説構造を列挙し、それらに対して、多重仮説の検定における gatekeeping procedure を基本にした推測方式を定義した。積仮説の検定には、それを構成する基本仮説に対する多重性調整法として Bonferroni 調整、Holm 調整、仮説の重要性に関する階層的順序関係に基づく上位仮説の検定などの適用規則を定め、個々の基本仮説の調整P値の式を求めた。全ての積仮説の検定でそれを構成する基本仮説に Bonferroni 法を適用した場合、Holm 法を適用した場合の、各基本仮説の調整P値の陽表現による公式を導いた。さらに複数の治療の間に順序関係が存在し、個々のエンドポイントについては治療間の順序による階層を設定し、エンドポイントは同等に重要と位置付けた場合あるいはその逆の場合の固定順序仮説系列の平行仮説群構造を考えた。この構造を multiple parallel/fixed sequence hypothesis structure と呼びこの構造のもとでの各仮説の調整P値の陽表現による公式を導いた。この仮説構造の下では、評価すべき仮説検定の個数は、本構造を仮定しない一般的な構造で必要な検定回数に比し、大きく削減されることが導かれた。同様に Bonferroni 法を適用した場合には仮説系列ごとに Bonferroni 基準に基づき有意水準を分割して割り当て、仮説系列ごとに固定順序の検定を行う場合に同等なことが導かれた。さらにこの構造の一般化として、エンドポイント系列毎に固定順序仮説の個数が異なる場合の公式の陽表現も与えた。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔学会発表〕(計 9 件)

上坂浩之、森川敏彦、嘉田晃子. 2相3

段階適応型試験デザインのデザインパラ

メータの選択. 2013年度 統計関連学会連

合大会. 2013年9 月9 日~9 月11 日 . 大
阪大学

Uesaka, H., Morikawa, T., and kada, A
Some considerations on use of
two-phase three-stage adaptive
designs. 34th Annual Conference of the
International Society for Clinical
Biostatistics. 25-29 August 2013
Ludwig-Maximilians- University Munich,
Germany

Morikawa, T. and Uesaka, H. Multiple
Testing on Multiple/Parallel Fixed
Sequence Hypothesis Structures. 34th
Annual Conference of the International
Society for Clinical Biostatistics.
25-29 August 2013 / Ludwig-
Maximilians-University Munich,
Germany

Uesaka, H., Morikawa, T., and kada, A.
Comparisons of three stage adaptive
designs for a clinical trial. 26th
International Biometrics Conference.
26-31, August, 2012. Kobe
Morikawa, T. and Uesaka, H.
Standardization of a gate keeping
procedures. 26th International
Biometrics Conference. 26-31, August
2013. Kobe

研究者番号 : 70399608

(4)研究協力者

森川敏彦 (MORIKAWA, Toshihiko)
元久留米大学バイオ統計センター

6 . 研究組織

(1)研究代表者

上坂浩之 (UESAKA, Hiroyuki)

研究者番号 : 60446250

大阪大学臨床医工学融合研究教育センター
招聘教授

(2)研究分担者

()

研究者番号 :

(3)連携研究者

嘉田晃子 (KADA, Akiko)

独立行政法人国立病院機構 (名古屋医療セ
ンター臨床研究センター) 室長