

多重検定の結果 (Shirley-Williamsの方法)

各群のデータ

	第1群(対照群)	第2群(処理群)	第3群(処理群)	第4群(処理群)
平均値	4.846485	5.847370	5.999199	7.345175
分散	1.550319	1.260604	1.106185	0.660575

検定の結果

群の数	検定統計量	ウィリアムズ基準点(自由度 ∞)			判定
		5%点	2.5%点	0.5%点	
4	4.152207	1.739000	2.032000	2.615000	有意水準0.5%で有意
3	2.053485	1.716000	2.015000	2.607000	有意水準2.5%で有意
2	1.954751	1.645000	1.960000	2.576000	帰無仮説を保留する。(P > 0.025)

第3群以上で有意差が認められます。(有意水準2.5%)

結果のみかた

帰無仮説棄却された時、判定には各 5.0%、2.5%、0.5% のどの水準で有意であるかが表示されます。
(ただし、表示される水準は、最初に指定している水準以下となります。)

帰無仮説が最初に指定した有意水準で保留された場合、その時点で検定を終了します。

図 9. MEPHAS 新規サイトにおける Shirley-Williams の方法の計算結果の表示ページ



ERROR!!

残念ながらデータの受理に失敗しました。
以下の点を確認してください。

各群のデータ数は整数で入力されていますか？

以上の点を確認したうえで、検定がうまくいかない場合、プログラムの不備の可能性が
あります。

お手数ですが、上のメニューからメールにて、どの手法でどのようなデータを用いた場合
にエラーが発生したのかご連絡いただけましたら、こちらでも調査いたします。

[戻る](#)

図 10. MEPHAS 新規サイトにおけるエラーメッセージ (データ数が整数で入力されていない場合) のページ



ERROR!!

残念ながらデータの受理に失敗しました。
以下の点を確認してください。

「各群のデータ数」とデータ入力数があるか確認してください。

以上の点を確認したうえで、検定がうまくいかない場合、プログラムの不備の可能性があります。

お手数ですが、上のメニューからメールにて、どの手法でどのようなデータを用いた場合にエラーが発生したのかご連絡いただけましたら、こちらでも調査いたします。

[戻る](#)

図 11. MEPHAS 新規サイトにおけるエラーメッセージ (データ数があていない場合) のページ

分担研究報告書

統計解析の理解促進に向けた Web サイトコンテンツの開発

研究分担者 川下 理日人 国立大学法人大阪大学大学院薬学研究科 助教

研究分担者 岡本 晃典 国立大学法人大阪大学大学院薬学研究科 助教

研究要旨

本事業において修正の基としている MEPHAS は 2002 年から公開されている Web サイトであり、本 Web サイト上において Web ブラウザーを通じて統計解析手法を利用可能である。Web サイトのコンテンツとしては統計解析手法の利用の他に、それら解析手法の簡単な説明と適切な手法の選択を補助するフローチャートなどがある。しかしそれらのコンテンツは、試験公開時点である 10 年以上前のデザインと技術で作成された Web ページを現在も用いており、現在の環境下で利用し易いとは言い難い。そこで本分担研究では、MEPHAS に設けられているコンテンツの内、特に統計解析手法の説明と選択補助についてより見易く、より使い易くなるよう修正することにより、それら説明と選択補助のコンテンツがより活用されるようになれば、MEPHAS 利用者の統計解析に対する理解の促進に寄与すると考え、修正、作成を行うことにした。以下、平成 22 年度に取り組んだ内容について報告する。

A. 研究目的

本研究において改訂を進めている Web サイト、MEPHAS[1] では統計解析プログラムの提供として、Common Gate Interface (CGI) を通じて使用者が解析したいデータをサーバー側へ送信し、サーバー側では統計計算を実施し、その結果を使用者の Web ブラウザーへと返す機能を Web 上にて提供している。また MEPHAS ではそれら統計解析プログラムの簡単な説明を提供し、かつ手法選択を補助するためのフローチャートを提供している。

本分担研究では特に MEPHAS 上において統計解析プログラムとして提供されている各手法の解説のための Web ページと手法選択補助のためのフローチャートの改訂を中心に修正、改訂を行うことにより MEPHAS 利用者のユーザビリティの向上を目的とした。また、特に手法の解説などは読み易く、かつ分かり易く情報を提供するこ

とも重要であり、それらは文字の大きさや 1 ページに表示される文字の量なども含めた Web サイトとしてのデザインに影響を受けるため、MEPHAS 全体としての Web サイトのデザインを修正することも重要な実施内容の一つと考え、Web サイト全体のデザインやメニュー表示なども検討、修正を行い、ユーザビリティの向上を図ることとした。

B. 研究方法

まずは現在の Web ページのデザイン、手法の説明文章、手法選択のフローチャート等の確認から始め、修正点を検討し、実際の修正を進めた。なお、MEPHAS の Web ページの中で、実際の統計解析手法を実施するためのデータ送信のページと結果表示を行うためのページに関しては、サーバー側のプログラム群との関連もあるため、CGI を介したデータ授受や統計解析の計算プログラ

ム等の修正の検討を行う別の分担研究において実施することにした。

(倫理面への配慮)

本研究の実施において取り扱うものは数値情報のみであり、それら数値情報も個人の特定につながり得る情報は含まれないため、倫理面への配慮等は必要ない。

C. 研究結果



Figure 1. 現在の MEPHAS トップページ

Figure 1 は現在の MEPHAS トップページを示している。図中に記載があるように、MEPHAS は 2002 年から公開しており、開発は 2000 年頃から 2002 年にかけて主に行われた。つまり Web サイトの構築方針や実際の作成方法なども 10 年以上前の考え方や技術などが使われている。2000 年当時のネットワーク回線では、ブロードバンドとして ADSL による 1Mbps (Mega bit per second; 1 秒間に 1M bit を送信できる回線が 1Mbps = 125 kbytes/s に該当する。ちなみに現在の光回線は Gbps, 1Gbps = 125Mbytes/s が主流である) の回線が一般家庭用として普及し始めた年代である。つまり Web サイト利用者の通信回線は現在よりも遥かに通信回線の容量が少なく、従って Web ページを作成する際に、例えば多数の画像や動画、Flash や Javascript などのプログラムをページに配

置すると表示に時間がかかるため控えるなど、可能な限り通信の容量を多くしない工夫が尊ばれる時代であった。MEPHAS の作成においても同様に、画像などは少なく、また動的なデザインなどは取り入れず、できる限り利用者のストレスとならないことを、Web サイト構築時の主方針として作成した経緯がある。

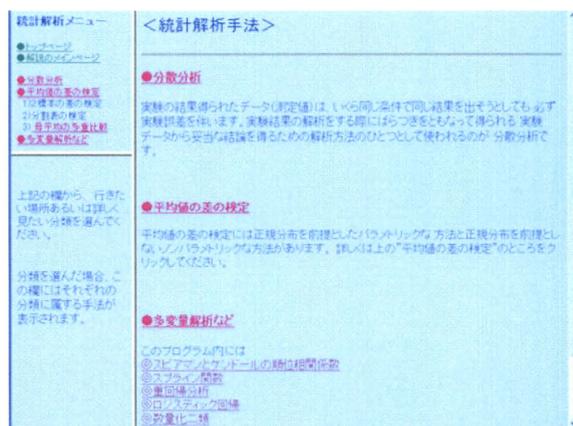


Figure 2. 現在の MEPHAS における『手法の解説』ページ

Figure 2 は Figure 1 のトップページの右側に表示されているメニューのボタンから『手法の解説』を選んだ際の移動先ページである。このページに関しても先のトップページと同様、可能な限り文字のみで情報を表現しようとしており、そのことは内容の分かり難さを助長していると思われる。また、現在の Web ページ作成においては検索エンジン最適化 (Search Engine Optimization: SEO) のためなどにより避けられる傾向のあるフレームによるページレイアウトが使用されている。勿論 MEPHAS は広告などのための Web サイトではないため、SEO などは関係のないものではあるが、World Wide Web Consortium (W3C) の HTML 4.01 や XHTML 1.0 の勧告[2, 3] においてもフレームは厳密 (strict) な文書定義にも、移行上の (transitional) 文書定義にも属しておらず、フレームセット (Frameset) として別に文書定義が設けられており、実質的に推奨されていないものである。

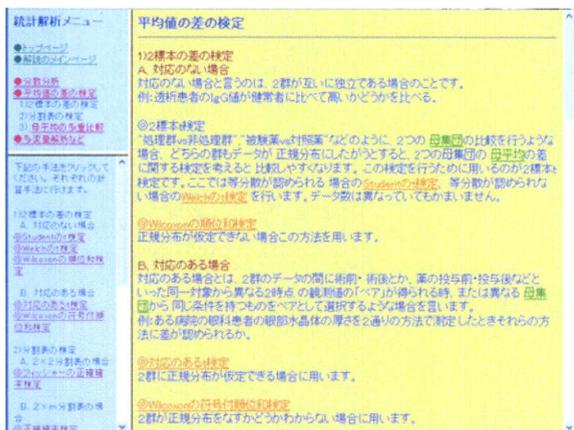


Figure 3. 『平均値の差の検定』の解説ページ

Figure 3 は Figure 2 の左フレームのメニューから『平均値の差の検定』を選択した際に表示されるページである。中央部フレームと左下部メニューのフレームが変化している。図に示されるように容量に負担をかけない方針もあり、図表も使用していない説明であり、また文章にしても決して読み易さや理解し易さに配慮したものとはなっていないように思われる。ほぼ Figure 3 と同様のため本報告書には記載しないが、他の解説ページも同様に図表などはほとんど使われていない。

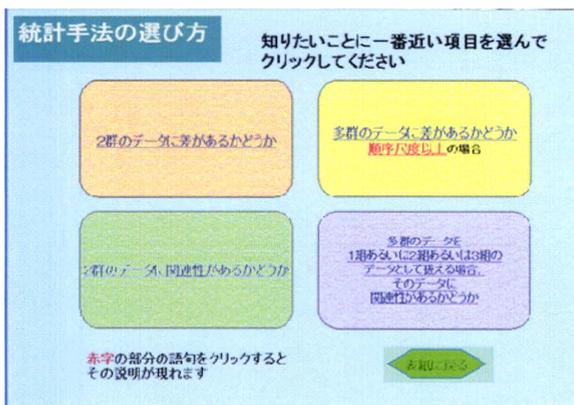


Figure 4. 『統計手法の選び方』のページ

Figure 4 は Figure 1 のトップページの右側に見えるメニューから『手法の選択』を選んだ際の移動先ページである。このページにおいても先の Figure 2 に示した『手法の解説』ページと同様、『表紙に戻る』として Web サイトのトップページへのリンクは準備されているものの、他ペー

ジへと移動するためのメニューはページ内に設けられていない。また先の『手法の解説』ページとは異なりフレームは使われていないが、その結果としてページごとにページデザインが変化し一貫性がないため、その意味でも Web サイトの利用者を迷わせる一因になり得ると考えられた。Figure 4 に示されたページの内容としては、ページ全体を図として表そうとしているものの、それにより内容を分かり易く伝えられているかには疑問が残る。Figure 4 に示した現在のページでは箇条書きと同様であり、改善の余地があるものと考えられた。

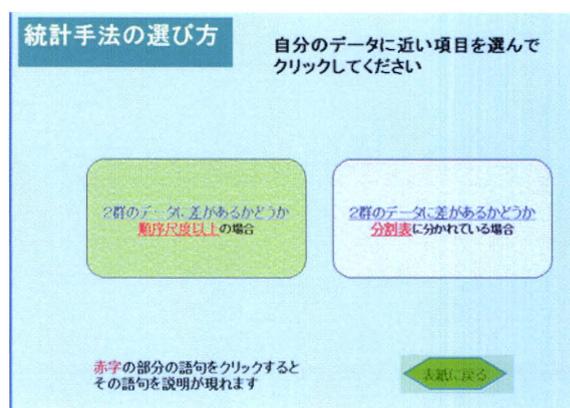


Figure 5. 『2群のデータに差があるかどうか』のページ

Figure 5 は Figure 4 に示したページの左上、『2群のデータに差があるかどうか』を選んだ際に移動するページである。Figure 4 から Figure 5 への移動を考えた場合、ページのデザインとしては一貫したものであるが、一方で、Figure 4 の『2群のデータに差があるかどうか』以外を選択した場合、例えば『多群のデータに差があるかどうか 順序尺度以上の場合』を選択した場合には Figure 6 に示すフローチャートが表示されるページへ移動する。

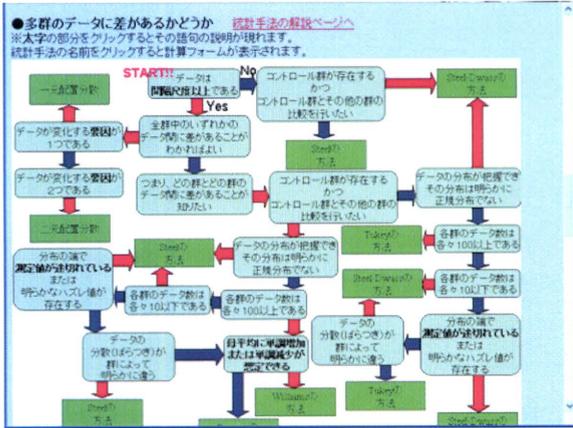


Figure 6. 手法選択のフローチャート掲載ページ中、『多群のデータに差があるかどうか』の箇所

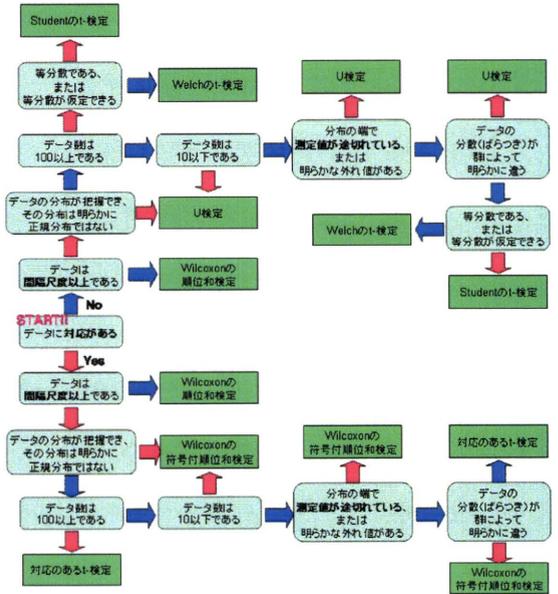


Figure 7. 『2群のデータに差があるかどうか 順序尺度以上の場合』のフローチャート

Figure 4 に示される『統計手法の選び方』のトップページでは、ユーザーが現在解析を考えているデータセットの特徴に応じて4つのカテゴリーが示されている。上記のように Figure 4 からの移動先は Figure 5 に示されるページを挟む場合と挟まない場合があった上で、フローチャートへとリンクしている。それらフローチャートへ導くための設問は必ずしもその意図が分かり易い文章とはなっていない。また、単純な質問形式とはいえ、後に使用されるフローチャートとは異なる形式での選択肢の提示は一貫性が無く、やはり使用者を戸惑わせる要因の一つとなり得る点であまり好ましくないと判断し、修正を行うことにした。

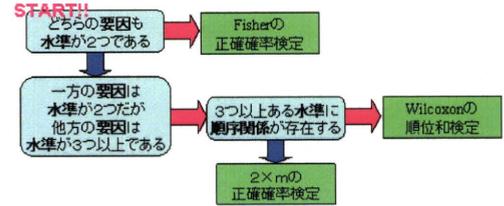


Figure 8. 『2群のデータに差があるかどうか 分割表の場合』のフローチャート

Figure 7 から Figure 11 までは各フローチャートを示している。Figure 7 から順に、『2群のデータに差があるかどうか 順序尺度以上の場合』のフローチャート、『2群のデータに差があるかどうか 分割表の場合』のフローチャート、『多群のデータに差があるかどうか 順序尺度以上の場合』のフローチャート、『2群のデータに関連性があるかどうか』のフローチャート、そして『多群のデータを1組あるいは2組あるいは3組のデータとして扱える場合、そのデータに関連性があるかどうか』のフローチャート、計5つのフローチャートである。

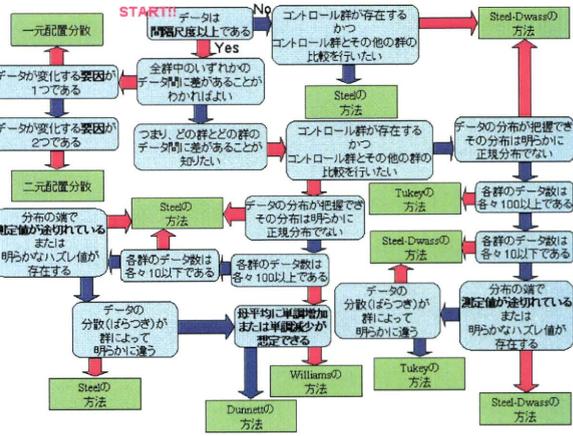


Figure 9. 『多群のデータに差があるかどうか 順序尺度以上の場合』のフローチャート

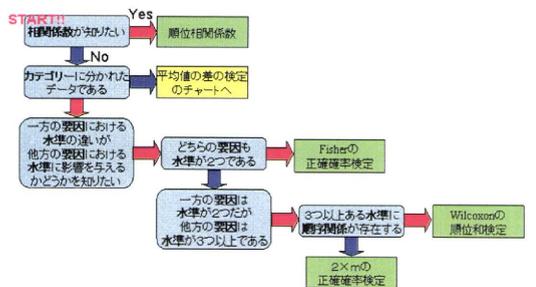


Figure 10. 『2群のデータに関連性があるかどうか』のフローチャート

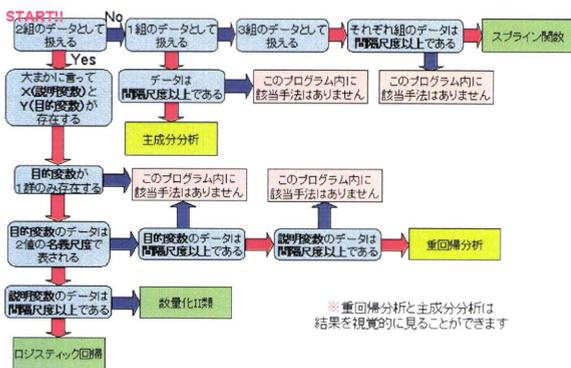


Figure 11. 『多群のデータを1組あるいは2組あるいは3組のデータとして扱える場合、そのデータに関連性があるかどうか』のフローチャート

Figure 10 のフローチャートは Figure 8 のフローチャートの中に含んでおり、さらにその直前には Figure 7 のフローチャートへ分岐する設問が設けられている。それらフローチャートの重複の解消も含め、手法選択のための手順を整理する必要があったと考えられた。

また Figure 7 及び 9 のフローチャートは情報が詰め込まれ過ぎており、見づらいこともユーザビリティを悪くする要因の一つと考えられた。さらに Figure 11 のフローチャートでは『このプログラム内に該当手法はありません』との結果に行き着くことが多く、該当手法を MEPHAS に追加するか、あるいはフローチャート自体の構成を見直すことが必要であると考えられた。

以上の点を考慮し、また MEPHAS 作成時期が 10 年以上前であることも踏まえ、Web サイトのデザインも含め、以下の点を平成 22 年度にまず修正を行う項目として考えた。

- 1). ナビゲーション機能（メニューの表示など）の充実
- 2). フレームの廃止
- 3). 統計手法の解説ページの読み易さの向上
- 4). 手法の選択に至るフローチャートの整理と修正



Figure 12. 新規サイトのトップページ

新規サイトに関する画面表示などは全て開発中のものであり、最終公開時とは異なる可能性がある。以降、明記しないが新規サイトに関する情報は全て同様に、最終公開時とは異なる可能性がある。

Figure 12 は新規サイトのトップページを示している（以降、特に明記しないが、本報告書中の新規サイトに関する情報は全て開発中のものであり、最終公開時とは異なる可能性がある）。フレームを用いずに各コンテンツへの移動メニューを上部にまとめたテンプレートを使用している。例えば『Explanation』からは現在の MEPHAS における『手法の説明』のコンテンツへと移動できる。Figure 13 はその移動先である『Explanation』のトップページである。



Figure 13. 新規サイトの『Explanation』のトップページ

先の Figure 12 のトップページにおいて記載したように、各コンテンツへの移動メニューが上部にまとめられ、トップページと同様に構成されており、利用者にとっては現サイトよりも分かり易く、使い易いことが期待できる。また、この『Explanation』以下のコンテンツに関するメニューは左側にまとめられている。



Figure 14. 新規サイトの『2標本の差の検定』の説明ページ

Figure 14 は Figure 13 の『Explanation』のトップページ左側の『menu』から『2標本の差の検定』をクリックし移動したページである。外見上は、フレームを使っている現在の MEPHAS と同様に中央の本文の箇所のみが変化しているが、フレームを使わずに上部と左側のメニューの箇所を同一のまま中央の本文の箇所を変えた別 HTML ファイルへと移動している。また本文についても先の Figure 3 に示した現サイトの説明とほぼ同内容の説明箇所を Figure 14 に示しているが、背景色の変更によってより見易くなっている他に、文

章の題字や箇条書きなどを使用し、また行の間隔などもより広く取っており、評価に個人差はあるにしても、読み易くなっているものと期待できる。

本文の内容についても解析手法を使用した例も含めて記載するよう、現在も修正を進めている。



Figure 15. 新規サイトでの手法の説明のページノンパラメトリックな多重比較法の解説のページを取り上げている。

Figure 15 にも示すように、手法の説明のページは見易くなったものの、現在のところ図表や事例が充実しているわけではなく、文章表現のみの説明も多いため、今後更なる修正が必要と考えている。



Figure 16. 新規サイトでの手法の説明のページ多変量解析のロジスティック回帰の解説のページを取り上げている。

数式についても全てを数式で説明することは利用者によっては読むことを避ける要因の一つになり得ると思われるが、Figure 16 に示すように説明上必要な箇所については使用して、適切であり、かつ分かり易い説明になるように文章を作成した。



Figure 17. 新規サイトでの『Selection』のトップページ

Figure 17 は新規サイトでの『Selection』(現サイトでの『手法の選択』に該当する)のトップページである。このページにおいても先の『Explanation』のページと同様、上部に Web サイト全体のメニュー、左側に『Selection』以下のコンテンツのメニューが設けられている。また中央部の本文箇所には、現在の MEPHAS におけるカテゴリーによる分類ではなく、分類用のフローチャートを表示させている。Figure 18 はそのフローチャートを示している。

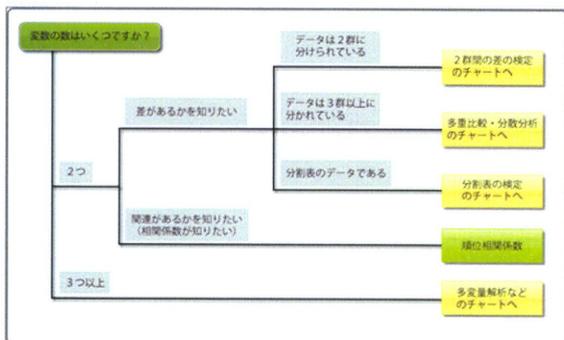


Figure 18. データの状況ごとのフローチャートを選択するためのフローチャート

Figure 18 のフローチャートは先の Figure 4, 5 に示された、使用者が解析したいデータの状況に応じた、各手法のフローチャートへの導入部分に該当する。右側の黄色のボタンから各手法のフローチャートへ、また黄緑色のボタンから解析手法のページ(順位相関係数)へと移動する。Figure 10 に示した『2群のデータに関連性があるかどうか』

のフローチャートの最初の設定問を Figure 18 に示されるようにフローチャート中に組み込むことにより、Figure 10 の『2群のデータに関連性があるかどうか』のフローチャートを削除し、Figure 8 の『2群のデータに差があるかどうか 分割表の場合』のフローチャート (Figure 18 では『分割表の検定のチャートへ』の移動先のフローチャートが該当する)へとまとめた。

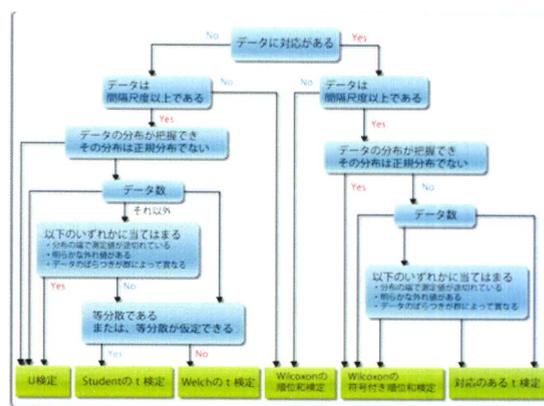


Figure 19. 新規サイトでの『2標本の差の検定』のフローチャート

Figure 19 は新規サイトでの『2標本の差の検定』のフローチャートを示している。このフローチャートは Figure 7 に示される『2群のデータに差があるかどうか 順序尺度以上の場合』のフローチャートに該当している。Figure 7 のフローチャートはできる限りブラウザ上でスクロールをせずに済むようまとめることも考えられており、その結果としてかなり入り組んでしまっている。従って、一見しただけでは開始点すら分かり難いほどであったが、Figure 19 のフローチャートでは上から下へとフローを流しており、一見して分かり易いフローチャートへと修正した。但し選択のための設問については、フローチャートのデザインを変えることはないにしてもより分かり易く、選択を進め易くなるよう、試験公開までにもう一度詳細に見直しを行う予定である。

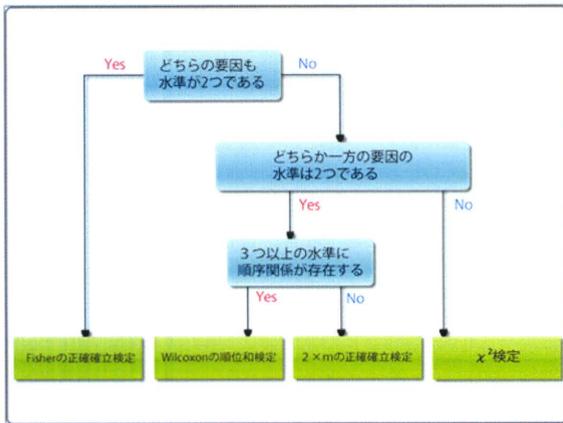


Figure 20. 新規サイトでの『分割表の検定』のフローチャート

Figure 20 は新規サイトでの『分割表の検定』のフローチャートを示している。これは先の Figure 8 に示される『2群のデータに差があるかどうか 分割表の場合』のフローチャートを修正したものである。フローチャートのデザインは先の Figure 19 に示したフローチャートと同様である。設問に関しては、Figure 8 の二つ目の設問に No の選択肢が無かったため、その選択肢を追加し、該当する手法として現状は χ^2 検定の追加を考えている。

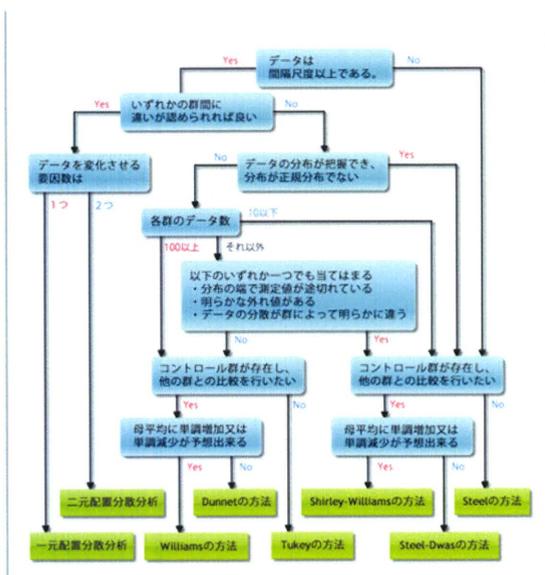


Figure 21. 新規サイトでの『多重比較・分散分析』のフローチャート

Figure 21 は新規サイトでの『多重比較・分散分析』のフローチャートを示している。こちらも現在の MEPHAS において該当する、Figure 9 に示した『多群のデータに差があるかどうか 順序尺度以上の場合』のフローチャートに比べて、上から下へとフローが流れるようにし、使用者が見易くなるよう設問等の配置も整理した。また、Williams の方法（序列を設けることができる多群の場合に適用される、パラメトリックな多重比較法）に対するノンパラメトリックな手法である、Shirley-Williams の方法を新たに MEPHAS に追加することで、多重比較法のパラメトリックとノンパラメトリックな手法の対が構成され、フローチャートもよりまとまりが改善された。なお、本フローチャートに新たに追加した Shirley-Williams の方法に必要なプログラム群は別分担研究にて作成しており、実装可能である。

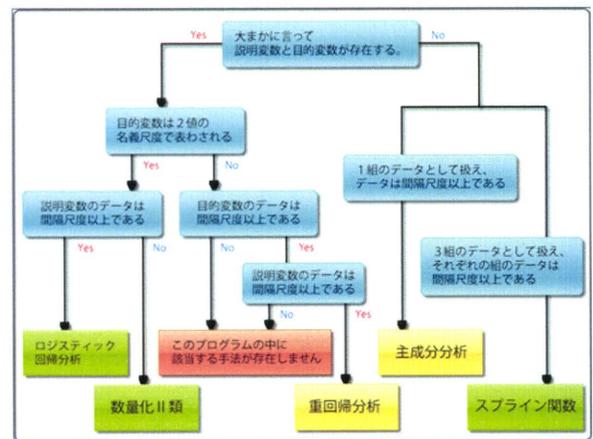


Figure 22. 新規サイトでの『多変量解析など』のフローチャート

Figure 22 は新規サイトでの『多変量解析など』のフローチャートを示している。現サイトで該当する Figure 11 の『多群のデータを1組あるいは2組あるいは3組のデータとして扱える場合、そのデータに関連性があるかどうか』のフローチャートに比べ、見易さの面では向上したと考えているが、該当する手法が無いとの結果に行き着く点

は必ずしも解決されてはいない。この点については、可能であれば新規サイトでは手法を追加することにより対応し、もし手法の追加が困難である場合には適切な手法を提示することまでは行えるよう、修正を検討したい。

D. 考察

まず、現在の Web サイトは 2000 年代前半、特に 2000 年から 2001 年にかけて作成されたものであり、Web サイトの作成方針や使用されている技術などに関して、現在の WWW や通信環境の状況、Web サイト構築の方向性などに合致していない点が多々見受けられたため、ユーザビリティ向上の一環としてデザインの再検討も含めて、Web サイトの見直しを行った。現在ではあまり推奨されず実際にほとんど使われなくなったフレームによる Web サイトのデザインを、現在の MEPHAS では使用しているため新規サイトでは修正し、上部に Web サイト全体のメニュー、左側にコンテンツ内のメニューを表示する形式で統一した。また現在の一般的なネットワーク環境において数枚の画像程度ではさほど通信容量へのストレスにはならないと考え、メニュー表示等に関してもより画像を利用し、現在のサイトにおける容量の軽減重視の方針からユーザーの見易さをより重視した Web サイトへの修正を実施した。これらにより Web サイトを使用する際に、ユーザーの行いたいことや知りたいことに、より行き着き易くなったことが期待できる。

次に『手法の説明』というコンテンツに関しては現在の Web サイトでは文字情報がほとんどであり、また背景色と相まって非常に読みづらいものとなっている。そこで、先のデザインの統一と同様、ユーザーの見易さにも配慮した本文へと修正を行った。今後は図表等もより追加し、かつ実例等も追加してコンテンツの充実を図ることも必要であると思われる。

最後に、『手法の選択』というコンテンツに関して、現在の Web サイトではまずユーザーの解析したいデータを 4 つのカテゴリーへと分類し、該当するフローチャートへ誘導していた点を修正し、全てフローチャートの形式へと統一した。また現在の Web サイトで用いられているフローチャートでは、先述の容量の軽減重視の方針からユーザーの見易さ使い易さを重視せず、一般的な Web ブラウザーの画面中に収まるように作成されていたため、複数のフローチャートにおいて開始点すら揃っていないこと、YES の回答しか行き先が指定されていない設問があること、現在の Web サイトでは計算できないとの結果に行き着く箇所が多いことなどがユーザビリティを下げるとして挙げられたため、それらに関して修正したフローチャートを作成した。それらフローチャートの修正に伴い、追加が必要になった手法に関しては、別分担研究にて構築を進めている。これらのフローチャートの修正はユーザーの使い易さの向上につながることを期待できる。今後も引き続き、フローチャートに設定する設問とその回答の見直しを図ることでよりユーザビリティの向上が期待できると考えている。

追加実施事項として、ユーザーの範囲をより広げるために英語版 Web サイトの構築も進めている（プログラムの計算結果も含め、文章の英訳中であり、まだ成果としては出ていないため結果からは割愛した）。

E. 結論

平成 22 年度に実施し、得られた成果としては以下の通りとなる。

- a. Web サイト修正の方向性の決定とデザインの作成
- b. 手法の説明文章の修正

- c. 手法の選択のためのフローチャートの整理と改訂版の作成

引用文献

- [1]. MEPHAS,
<http://www.gen-info.osaka-u.ac.jp/MEPHAS/>
- [2]. W3C XHTML™ 1.0 The Extensible HyperText Markup Language (Second Edition) A Reformulation of HTML 4 in XML 1.0,
<http://www.w3.org/TR/xhtml1/>
- [3]. W3C HTML 4.01 Specification,
<http://www.w3.org/TR/html401/>

F. 研究発表

1. 論文発表

特になし。

2. 学会発表

福井大介¹, 田雨時², 岡本晃典², 川下理日人^{2,3}, 後藤直久³, 安永照雄³, 高木達也^{2,3} (¹阪大薬,² 阪大院薬,³ 阪大微研), 医薬学データ用統計解析プログラム、MEPHAS の更新について, 29P-0115, 日本薬学会 第131年会, 静岡, 2011年3月 (東日本大震災のため開催中止となったが、組織委員会より要旨等も配布済みのため年会は成立したものとするとのお知らせあり)

G. 知的財産権の出願・登録状況

1. 特許取得

特になし。

2. 実用新案登録

特になし。

3. その他

特になし。

厚生労働科学研究費補助金（医薬品・医療機器等レギュラトリーサイエンス総合研究事業）
分担研究報告書

統計解析の理解促進に資する Web サイト構築のためのプログラム群の開発

研究分担者 高木 達也 国立大学法人大阪大学大学院薬学研究科 教授

研究分担者 岡本 晃典 国立大学法人大阪大学大学院薬学研究科 助教

研究要旨

本研究において開発の対象としている MEPHAS は CGI によりデータの授受を行う統計解析プログラムパッケージであり、それらを可能としているプログラム群はサーバー側にて稼働している。それら現在のプログラム群は 2002 年の試験公開から稼働しており、利用者からの指摘を中心にプログラムの見直しを適宜検討してきた。しかしながらそれら見直しの多くは、計算実施上のプログラミングミスや表示におけるミスなどであった。そこで本分担研究では MEPHAS における利用者の利便性の向上による MEPHAS 利用の促進、ひいては統計解析への興味を促進するよう、プログラム群を検討し直し、修正を行うことを目的に実施し、データ入力時の間違いの内容に応じた対応ができるよう、プログラムの修正を行った。また、Web サイトのコンテンツを中心に見直しを図る分担研究と連動し、解析手法の追加のためのプログラム作成等も実施した。以下、実施内容を報告する。

A. 研究目的

現在の MEPHAS[1] は Unix（もしくは Linux）上において HTTP サーバープログラムである Apache [2] を用いて提供されており、Common Gate Interface (CGI) を利用した統計解析プログラムパッケージをその機能として提供している。サーバー側にて動いているプログラムは、シェルスクリプト、C 言語、Perl などの言語を用いて作成されている。それらプログラム群は Web サイトと同様、2000 年代の前半に作成されたものであるが、プログラムは Web サイトの HTML ファイルなどと異なり、外見上の見た目やデザイン等に関わるものでなく作成年代が古いことが更新の要因とはならないため、プログラムとしての間違いや統計手法の計算上の間違いなどが無いのであれば、特に更新の必要性は無いものと思われる。従って、主に Web サイトの更新に伴い必要となるプログ

ラム群への機能の追加や新規プログラムの作成などを行うことを目的とした。

B. 研究方法

現在の MEPHAS において稼働中のプログラム群の構成を再検討した。2002 年の公開以来、10 年以上稼働させていたプログラムであり、作成者自身で気付いた箇所と利用者からの不具合連絡により判明した箇所の双方が修正されてきた経緯から、正常に動作する範囲はかなり広くかつ計算結果等にも誤りが含まれている可能性はあまり高くはないものと思われる。その一方で、原因はさておき、プログラムが正常に動作しなかった際の挙動については情報の蓄積がさほど多くはないため、特にその点に関し見直しを図ることにした。

また、デザイン等も含めたユーザビリティ向上を目的とした Web サイトの見直しを検討する

別分担研究における検討の結果を受けて、プログラム等に関して追加、あるいは修正が必要になった事項については適宜、対応を行うこととした。

(倫理面への配慮)

本分担研究は、Web サイトとその統計解析手法の計算機能を提供するサーバーにおいて動作しているプログラム群の修正や、新規プログラムの作成を主に実施するものであり、そこで使用されるものは数値データのみであるため、倫理面での特別な配慮を必要とする研究ではない。

C. 研究結果

まず、現在の MEPHAS における統計解析手法の計算を実施する際に動作しているプログラムについて整理したものが Figure 1 である。

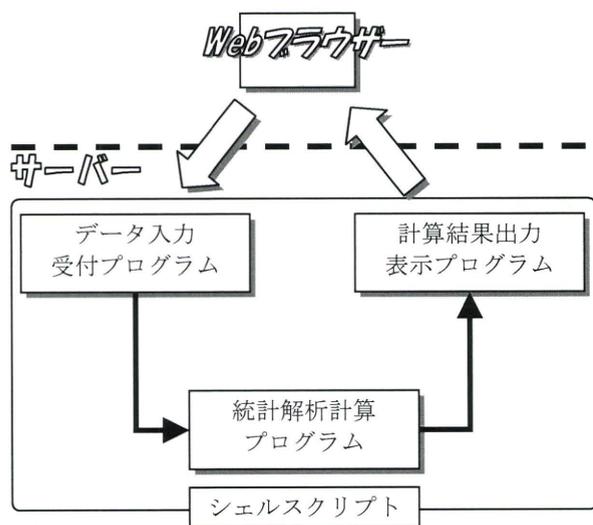


Figure 1. 現在の MEPHAS において動作しているプログラム群の概要図

Figure 1 に示されるようにユーザーは Web ブラウザーにより MEPHAS へアクセスし、CGI を通じてサーバー側へデータを送信する。サーバー側では3つのプログラムがシェルスクリプトにより1つにまとめられ、使用者からのデータを処理し、計算を実施し、計算結果を出力する。

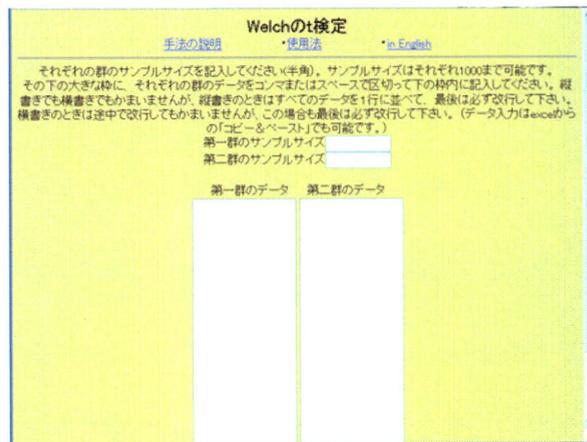


Figure 2. 現在の MEPHAS における Welch の t 検定のデータ入力画面

Figure 2 は現在の MEPHAS における Welch の t 検定のデータ入力画面を示している。上部に手法の説明や簡単な使用法のページ等へのリンクが準備されており、本文の箇所に入力項目の概説とサンプルサイズや実際の数値データを入力するためのボックス (CGI によるデータ送信のためのテキストボックスやテキストエリアなど) が用意されている。『手法の説明』は MEPHAS のコンテンツの一つである『手法の説明』の該当するページへのリンクとなっており、『使用法』のページは Figure 3 に示すように、先の Figure 2 に示された入力用のページにサンプルデータが入力され、少し説明が追加されている。

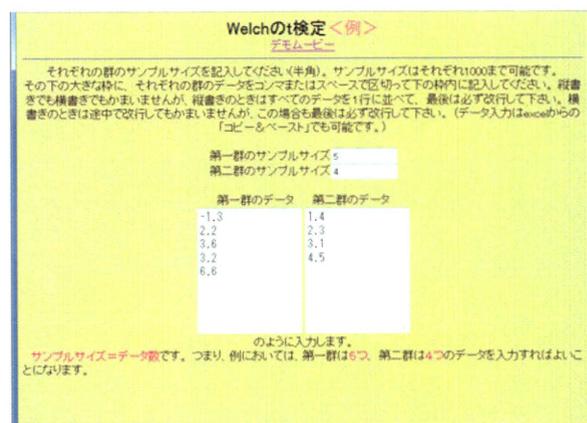


Figure 3. 現在の MEPHAS における Welch の t 検定の使用法のページ

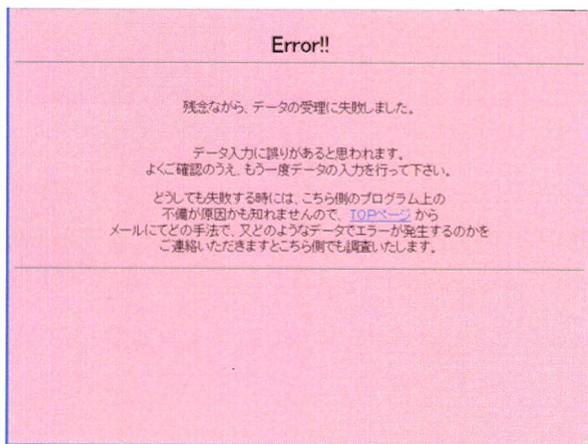


Figure 4. 現在の MEPHAS におけるデータ入力時の間違いに対するエラーメッセージの画面

Figure 4 は現在の MEPHAS においてデータを入力し間違いがあった際に、そのまま計算はせず使用者（入力者）の側へ間違いがあったことを知らせるためのエラーメッセージのページを示している。MEPHAS に限らず、計算結果を返すアプリケーションにおいて注意すべきことの一つとして、使用者が意図した入力と違うデータセットを基にした計算結果を返してしまうことが挙げられる。何らかの計算結果が返ってきた際に使用者がその結果を疑うことはまずあり得ないため、システム側で使用者の意図したデータ入力を受け取れているのかどうかをチェックすることが必要になる。従って MEPHAS においても、入力される数値データのみあればサンプルサイズや群数はその数値データから求めることが可能であり、Web サイトでの入力は必要ないが、使用者の意図しているデータの形式となっているかどうかの確認のためにも群数やサンプルサイズなどの入力を要する形式を採用している。しかしながら現在の MEPHAS ではエラーメッセージの種類に乏しく、使用者から入力されたデータを受け付け、統計解析計算プログラムの入力データへと変換する部分である、Figure 1 における『データ入力受付プログラム』におけるエラーメッセージ

は実質的には Figure 4 に示した一種類のみであった。

そこでまず、使用者の側がエラーメッセージを受け取った際に改善を図り易くするために、入力されたデータのチェックにおいて間違いを見つけた際のエラーメッセージのレポーターを増やすようプログラムの修正を行うことにした。

また、Web サイトの HTML ファイルやコンテンツを中心に改訂を行っている分担研究と連携し、多重比較法の一つである、Shirley-Williams の方法を MEPHAS で取り扱う解析手法の一つとして新たに追加できるよう、プログラムの作成を行った。



Figure 5. 新規サイトでの『多変量解析など』のフローチャート

Figure 5 は改訂したデータ入力を受け付ける Web ページである。現在の MEPHAS と同様、使用者からのデータ送信には CGI は使用しているため、デザインのみ他のページに合わせたものとしている。



Figure 6. 新規サイトでの『多変量解析など』のフローチャート

Figure 6 はデータを入力した際に空欄があった場合に表示されるエラーメッセージである。中央部分のフォントサイズが大きくなっている箇所に何がエラーの原因となったのかを表示している。



Figure 7. 新規サイトでの『多変量解析など』のフローチャート

同じく Figure 7 は入力されたデータ中に半角数字が予定される箇所に別の文字情報や記号などが入力された場合のエラーメッセージを示している。先の Figure 6 と比較すると、中央部のフォントが大きくなった箇所の文章が異なっていることが分かる。

上記も含めて以下のエラーメッセージのレパートリーを作成した。

- データの入力量が多すぎる場合
- 入力に空欄がある場合
- 入力が半角数字で行われていない場合
- 信頼区間など入力範囲があるもので範囲外の入力の場合
- 群数など整数値が期待されるもので整数値でない場合
- サンプルサイズと入力データが合致しない場合

それぞれにおいて、『データ入力受付プログラム』内でのデータチェック結果に応じて、エラーに該当するときには、Figure 6, 7 に示したようなエラーメッセージが表示される。

MEPHAS の Web サイトとしてのデザインやコンテンツの見直しを進めている別分担研究での検討を受けて、多重比較法の一つである Shirley-Williams の方法を追加するために必要となるプログラム作成等を行った。Shirley-Williams の方法について、以下概説する。

Shirley-Williams の方法は、例えば薬剤投与量が 0mg/kg (コントロール群), 10mg/kg, 50mg/kg, ... などのように、多群間に序列があるような数値データについてコントロール群との対比較を複数回行う際に、その序列も考慮に入れた多重比較を行うための手法である。同様の解析手法として Williams の方法(こちらは MEPHAS にて使用できる)があり、Shirley-Williams の方法との違いはパラメトリックな手法かどうか、つまり各群の数値データに正規分布を仮定するかしないかの違い(ちなみに Williams の方法はパラメトリックな手法)である。Shirley-Williams の方法では、他のノンパラメトリックな手法と同様、間隔尺度以上の

数値データを順序尺度へと変換する（数値データに順位をつける）ことにより、数値データの分布形の条件から解放されている。この Shirley-Williams の方法も含めて、ノンパラメトリックな多重比較法を収載している既存の統計解析ソフトウェアはあまり多くは存在しておらず、MEPHAS の利用の主目的の一つとなっている。

上記の Shirley-Williams の方法を計算できるようなプログラムを作成した。新たに作成したプログラムであるため、複数のサンプルセットを乱数を用いて作成し、Shirley-Williams の方法を収録している既存の統計解析ソフトウェアの一つである Excel 統計[3] と計算結果の比較を行い、違いの無いことを確認することで、プログラムが適切に作成されていることの確認とした。



Figure 8. Shirley-Williams の方法による計算結果

Figure 8 は Shirley-Williams の方法による計算結果を示している。

D. 考察

本分担研究では主に、MEPHAS の統計解析プログラムを実行するためにサーバー側で稼働しているプログラム群に関し、ユーザビリティの向上につながる修正を中心にプログラム群の見直しと実際の修正を実施した。

現在の MEPHAS のプログラム構成を見直し、使用者の使い易さの向上につながるよう、修正点

を検討した。その結果、CGI を介して使用者からデータを受け取るためのプログラム内でのデータチェックにおいて間違いを見つけた際の対応（エラーメッセージの内容も含む）がほぼ同一であり、使用者に間違いのあることを伝えるのみとなっており、使用者がその間違いの修正しづらいことが、MEPHAS 上の統計解析手法を利用することの妨げとなり得ると考えられた。そこで、間違いの内容に応じた対応ができるようプログラムを修正することで、使用者の使い易さの向上が図れると考え、プログラムの修正を行い、間違いの内容に応じたエラーメッセージが表示されるようにした。今後はデータ入力時のみならず、統計解析手法の実行時に生じるエラーについてもエラーメッセージの表示など使用者側へエラー内容について情報提供できるように修正することも必要と考えている。また統計に関する学習の面では、使用者が選択した統計解析手法が適切でない場合に、その理由を明示すると共に適切な手法へと誘導できるように検討する予定である。

また、別分担研究の実施項目である手法選択のフローチャートの修正に伴い、追加が必要となった多重比較法を実施するためのプログラムの作成を行った。追加された多重比較法は、群間に序列がある場合のコントロールとの対比較を行うパラメトリックな手法の Williams の方法と対となる、ノンパラメトリックな手法の Shirley-Williams の方法である。作成したプログラムは既存の統計解析ソフトウェアにて行った Shirley-Williams の方法の解析結果と比較し、違いが無いことを確認した。今後も手法を選択する上で必要となった統計解析手法を中心に、手法の追加は引き続き検討する予定である。

E. 結論

平成 22 年度に実施し、得られた成果としては以下の通りとなる。

- a. 入力データの誤りへの対応に向けたプログラムの修正
- b. 統計解析手法の追加のためのプログラム等の作成

引用文献

- [1]. MEPHAS,
<http://www.gen-info.osaka-u.ac.jp/MEPHAS/>
- [2]. Apache, <http://www.apache.org/>
- [3]. Excel 統計, <http://software.ssri.co.jp/ex2010/>

F. 研究発表

1. 論文発表

特になし。

2. 学会発表

福井大介¹, 田雨時², 岡本晃典², 川下理日人^{2,3}, 後藤直久³, 安永照雄³, 高木達也^{2,3} (¹阪大薬, ²阪大院薬, ³阪大微研), 医薬学データ用統計解析プログラム、MEPHAS の更新について, 29P-0115, 日本薬学会 第131年会, 静岡, 2011年3月 (東日本大震災のため開催中止となったが、組織委員会より要旨等も配布済みのため年会は成立したものとすとの通知あり)

G. 知的財産権の出願・登録状況

1. 特許取得

特になし。

2. 実用新案登録

特になし。

3. その他

特になし。

研究成果の刊行に関する一覧表

学会発表

発表者氏名	発表タイトル名	発表学会名	開催場所	発表年月
福井大介, 田雨時, 岡本晃典, 川下理日人, 後藤直久, 安永照雄, 高木達也	医薬学データ用統計解析プログラム、MEPHASの更新について, 29P-0115	日本薬学会第131年会	静岡	2011年3月