

クロスオーバー計画を用いた超音波治療器・光線治療器の施術効果の測定
The measurement of effect of therapies - ultra sound treatment and
phototherapy using the crossover plan (RCT).

四国医療専門学校 四宮 英雄・山本 幸男

要旨

目的：

現在、エビデンスのある研究が求められている。ランダム化比較試験(RCT)はエビデンスの高い研究に属する。従来、多くの RCT が実施されてきたが、研究者の所属する専門学校・大学の学生を用いたものが多く、実質的にランダム化されていないものが多い。さらに、比較するための必要サンプル数の検討は、ほとんど実施されていない。そこで本研究では、以上のことを考慮して研究を実施した。

仮説：

超音波治療は光線治療に比べて温熱効果が深部まで達するので、前者は後者に比べて、筋硬度をより低める。ただし、温熱効果により、筋の膨張が生じることもあるので、その差は、有意になることはない。

方法：

被験者 5 人の肩井部の筋硬度を測定、その後、10 分間超音波治療をし、2 分休憩後に筋硬度を測定した。2 週間後、同じ被験者 5 人に同様の光線治療を行った(クロスオーバー研究)。5 人の結果を基に、ブートストラップ法を用いて 1000 人分のシミュレーション・データを取得した。これらのデータから 15 人分をランダムに抽出し、実測した 5 人分のデータと合わせて 20 人分のデータ(必要サンプル数)を用いた。なお、施術者と測定者は異なり、測定者は治療方法を知ることはなかった。

結果：

超音波治療器の筋硬度平均低下 2.84kg、光線治療器のそれは 2.04kg、 $t=0.72$ ($-1.53 \leq t \leq 3.13$) ($p > 0.05$)。

結論：

超音波治療器の効果がより高いが、2 治療器の治療効果に有意差があると

はいえない。その理由は温熱効果により、筋の膨張が生ずることもであると推察されるからである。

提案：

専門学校・大学で実質的な RCT を実施することは難しい。また被験者の数も少数であることが多い。そこで今回、RCT に近似させるために、クロスオーバー研究を用い、必要サンプル数 20 を得るためにブートストラップ法を用いた。今後、各先生方が RCT を実施する際の参考になれば幸いである。

Abstract

Background: Much RCT (randomized Controlled Trial) has been carried out conventionally, but they are insufficient. So, the purpose of this study carried out RCT in earnest.

Hypothesis: Because a warm temperature effect reaches in comparison with phototherapy to the deep part, as for ultra sound treatment, an effect is higher in the former than the latter. But the significant difference does not occur.

Methods: We tested it using some methods - cross-over study, boot strap method, and number of the necessary samples (20).

Results: (1) Mean drop of muscle hardness of ultra sound treatment: 2.84 kg, (2) Mean drop of muscle hardness of phototherapy: 2.04 kg, (3) $t=0.72$ ($-1.53 \leq t \leq 3.13$) ($p > 0.05$)

Consideration: The effect of ultra sound treatment is higher than phototherapy, though the significant difference does not occur. The reason is because the expansion of the muscle may occur because of a warm temperature effect.

Proposition: We wish this study design - cross-over study, boot strap method, and number of the necessary samples - is useful for other researchers.

キーワード：

クロスオーバー計画、ブートストラップ、超音波治療器、光線治療器、施術効果

本文

はじめに：

柔道整復学の研究においてもエビデンスのある研究が求められている。この研究の上位に位置するのは、ランダム化比較試験である。従来の研究でも、ランダム化比較試験が実施されているが、研究者の所属する専門学校・大学の学生を用いたものが多く、実質的にランダム化されていない。しかも、比較するためのサンプル数の検討がされている研究は、ほとんどない。そこで本研究では、以上のことを考慮し、2 治療器の有意差、および無治療（プラセボ）との有意差を検討することにした。

現在、本校の付属接骨院では、多くの物療機器を使用している。これらの有効性については、多くの研究成果が発表されている。しかし、それらの研究の質には、さまざまな問題点がある。本研究では、そのなかで、

- 1 単なる p 値による有意差の検討ではなく、施術成果と関連したサンプル数の確定
- 2 クロスオーバー計画の導入と持ち越し効果の解消法の提示
- 3 マスキングの導入によって、従来の研究の質の改善を図ることを目的とする。

この研究によって、超音波治療器・光線治療器の施術効果を客観的で、質の高い方法で測定することが可能となる。これらの研究成果を活用して、今後、付属接骨院での治療にあたりとともに、学生に対する物理機器の指導を行いたい。さらに、物理機器の使用法の工夫を行って、より有効な施術方法を検討する。

対象および方法

1 サンプル数は各治療が、他の治療・治療なしと比較して、20%以上改善した場合を施術において有効と認め、事前に確定する。単なる p 値による有意差の検討ではなく、施術成果との関係を問う。

2 クロスオーバー計画： 各グループの治療担当者は異なる。各治療の持ち越し効果を解消するため、各治療後 2 週間休む。治療時間は各々 10 分間とする。治療なしは治療器をあてるが照射しない。

グループ 1 超音波治療 2 週間休む 治療なし 2 週間休む

光線治療

グループ 2 光線治療 2 週間休む 超音波治療 2 週間休む
治療なし

グループ 3 治療なし 2 週間休む 光線治療 2 週間休む
超音波治療

3 測定者・測定方法： 測定者は上記の治療担当者とは異なる。また、被験者が、どの治療を受けたかをマスクキングするため、治療後 10 分経過後に測定する。測定部位は、肩井部とし、筋硬度計で事前・事後の測定を行う。

なお、使用する超音波治療器はインテレクトモバイル US(PH-U2776CG)、光線治療器はスーパーライザー HA-2200 LE1 (東京医研株式会社)。また、筋硬度計は、NEUTONE TDM-NA1(DX) 有限会社トライオールを使用した。

被験者 5 人の肩井部の筋硬度を測定し、その後、10 分間超音波治療をし、2 分休憩後に筋硬度を測定した。2 週間経過後、同じ被験者 5 人の肩井部の筋硬度を測定し、その後、10 分間光線治療をし、2 分休憩後に筋硬度を測定した (クロスオーバー研究)。これら 5 人の結果を基に、ブートストラップ法を用いて 1000 人分のシミュレーション・データを取得した。これら 1000 人のデータから 15 人分をランダムに抽出し、実測した 5 人分のデータと合わせて 20 人分のデータ (必要サンプル数) で超音波治療と光線治療の有意差を解析した ($p < 0.05$)。なお、施術者と測定者は異なり、測定者は、どちらの治療を行ったかを知ることはない (マスクキング)。

結果：

治療なし・超音波治療・光線治療の 3 群による分散分析では、有意差は認められなかった。そこで、超音波治療・光線治療の 2 群による平均値の差の検定を実施した。結果は表 1 のとおりである。

表 1 超音波治療・光線治療の 2 群による平均値の差

母平均の差の検定 (対応あり)

基本統計量

| 変数 | n | 平均 | 不偏分散 | 標準偏差 | 標準誤差 |
|----|----|--------|--------|-------|-------|
| L | 20 | -2.040 | 8.512 | 2.918 | 0.652 |
| U | 20 | -2.840 | 16.003 | 4.000 | 0.894 |

対の差と効果量(対の差 = |「L」-「U」|)

| 対の差 | 標準偏差 | 効果量 | Cohen's d | Hedges' g | Cohen's d' |
|-----|--------|--------|-----------|-----------|------------|
| 平均 | 0.8000 | 4.9711 | 0.2344 | 0.2285 | 0.1609 |

差の平均の 95%信頼区間(両側検定/対立仮説:対の差 \neq 0)

| | 下限値 | 上限値 |
|------|---------|--------|
| t 推定 | -1.5265 | 3.1265 |

差の平均の検定(両側検定/対立仮説:対の差 \neq 0)

| | 統計量:t | 自由度 | P 値 | *:P<0.05 **:P<0.01 |
|------|--------|-----|--------|--------------------|
| t 検定 | 0.7197 | 19 | 0.4805 | |

検出力(両側検定/対立仮説:対の差 \neq 0)

| | $\alpha=0.05$ |
|------|---------------|
| t 検定 | 0.1051 |

考察：

今後の課題：

1 今回はパイロット研究のため、被験者 5 名で研究を行ったが、今後、適切なサンプル数である 20 名で実施したい。その際、ランダムサンプリングを行うことが非常に難しい(柔道整復学科の学生に偏ってしまう恐れがある)。しかし、本学の鍼灸マッサージ学科、鍼灸学科、理学療法学科、作業療法学科、看護学科、さらに教職員の協力も得ることによって、できるだけランダムサンプリングに近い形で研究の質を高めたい。

2 付属接骨院に設置している、他の物療機器についても、今回の方法をさらに改善して、その施術効果を客観的で、質の高い方法で測定する。

3 これらの研究成果を活用して、今後、付属接骨院での治療にあたりとともに、学生に対する物理機器の指導を行いたい。さらに、物理機器の使用方の工夫を行って、より有効な施術方法を検討したい。

結論：

超音波治療器の効果がより高いが、2 治療器の治療効果に有意差があるとはいえない。その理由は温熱効果により、筋の膨張が生ずることもあると推

察されるからである。また、「治療なし」との差も認められたが、20パーセント改善するという有意な差は生じなかった。

(補論)

1 クロスオーバー比較試験について

代表的な臨床試験には、並行群間比較試験とクロスオーバー比較試験がある。並行群間比較試験では、被験者をいくつかの試験群にランダムに割り付け、試験開始から試験終了まで試験群ごとに異なった治療を行い、試験群間の比較を行う。たとえば、試験群 A に割り付けられた被験者は試験開始から試験終了まで A 薬（治療法）を投与され（受け）、試験群 B に割り付けられた被験者は B 薬（治療法）を投与される（受ける）。クロスオーバー比較試験では、各被験者は 2 つ以上の試験治療を行う順序をランダムに割付けられる。たとえば、2 つの薬剤（治療法）を 2 期に分けて比較する場合、ある被験者は、第 1 期には A 薬（治療法）で治療を受け、第 2 期には B 薬（治療法）で治療を受ける。別の被験者は、第 1 期には B 薬（治療法）で治療を受け、第 2 期には A 薬（治療法）で治療を受ける。したがって、同一の被験者について、異なった治療法を比較することができる。

クロスオーバー比較試験の最大の利点は、同一の被験者について異なった治療法を比較することができる点である。異なった治療法の比較において被験者間のばらつきを除くことができるため、並行群間比較試験に比べて少ない症例数で試験が実施可能である。クロスオーバー比較試験の最大の問題は、持ち越し効果(carry-over effect)が生じる可能性があることである。持ち越し効果とは、先行する試験治療が次に続く試験治療期間において及ぼす残存効果である。不均等な持ち越し効果が試験治療の直接の比較を偏らせる。そのために、2 つの治療期間の間に適切な休薬（休息）期間（washout period）をおくことが必要になる。長い持ち越し期間が必要になる場合には、クロスオーバー比較試験は適切でないので、並行群間比較試験を考えることになる。期間の長い試験では、季節の変化など外部要因による時間の影響を受ける可能性もある。病気（病態）が治療期間中に進行したり軽くなったりするかもしれないので、対象とする疾患は慢性的で症状が安定しているべきである。薬剤に関連のある効果は、その試験治療期間中に完全に現れなければならない。休薬（休息）期間は薬剤（治療）効果が完全に消失するよう十分長くすべきである。

（本研究では、筋硬度の測定であり、各治療器の効果は経験上、2 週間以上は持続しないので、2 週間という休息期間は適切であると考えられる）

クロスオーバー比較試験のもうひとつの問題は、被験者が途中で試験から外れた場合に、解析とその結果の解釈が複雑になることである。

(本研究でも、2週間後の、同時帯に、学生を確保することに、非常に苦労した。今後、サンプル数 20 名に治療する場合に、どのように学生・教職員を確保するかが課題である)

また、持ち越し効果が存在する可能性があることから、後続の試験治療期間に発生した有害事象が、どちらの試験治療によって生じたものか判断することが困難になる。クロスオーバー比較試験は、一般に試験から外れる被験者が少ないと期待できる場合に限定すべきである。個々の被験者にとっては、並行群間比較試験にくらべて試験に参加する期間が長くなるために、試験から脱落する確率が高くなる。

第 1 期の測定結果と第 2 期の測定結果に違いが生じることがある。これは、異なる大きさの持ち越し効果によるものかもしれないし、真の交互作用によるものかもしれない。異なった時期における状態の違いが、何らかの理由により薬剤間の差に影響を及ぼす可能性がある。

(本研究でも、今後、この点に注意して、事前に詳細に検討された研究デザインを作成する必要がある)

一般的に、よく計画された試験では薬剤-時期交互作用は小さいが、交互作用や持ち越し効果を確認できる試験デザインが望まれる場合がある。そのような場合には、 2×2 クロスオーバー比較試験を拡張した反復クロスオーバー試験デザインなど高次の試験デザインを考えることができる。検定の感度は 2 つの治療法間の相関性に依存しているので、あらかじめ相関性についても注意しておく必要がある。

() 内は、四宮による記述である。

(光森達博. クロスオーバー比較試験で用いられる統計解析手法, PhamStage, 12(12), 17-20 (2013))

(厚生省. 「臨床試験のための統計的原則」について (医薬審 第 1047 号, 平成 10 年 11 月 30 日) (1998)

http://www.pmda.go.jp/ich/e/e9_98_11_30.pdf)

2 ブートストラップ (bootstrap) は統計学者 Efron が考案した計算方法である。コンピュータによるシミュレーションの一種であるが、データを乱数で生成するのではなく、実際のデータに基づいた推論をするのが、ブートストラップの特徴である。例えば平均値と中央値のどちらが安定か (ぶれが少ないか) を調べるために、通常のシミュレーションでは「乱数でデータを

生成して平均値と中央値を計算する」という作業を何度も繰り返し、その分布を比較する。これに対して、ブートストラップでは、「実際のデータ(n 個)からランダムに n 個を復元抽出(重複を許した抽出)し、その平均値と中央値を計算する」という作業を何度も繰り返す。例えば実際のデータが(2, 3, 5, 7)であれば、それから復元抽出した(2, 3, 3, 7), (3, 5, 7, 7), (2, 5, 5, 5)などについて統計量(平均値, 中央値など)を計算し、その分布を調べる。このような単純な方法で、従来は正規分布など数学的に理想化された分布についてしか求めることができなかつた統計量の分布(例えば中央値の分布)を、現実のデータとほぼ同じ分布について求めることができる。

(奥村 晴彦 <https://oku.edu.mie-u.ac.jp/~okumura/stat/bootstrap.html>)

3 統計的検定について

統計的検定は多くの研究において使用されているものの、適切に用いられているとは限らない。研究の信頼性を向上させるためには、統計的検定において少なくとも、次の 4 つの要素が明確にされなければならない。すなわち、(1) 有意水準(p 値)、(2) サンプル数(N): 研究前に適切なデータ数を計算する、(3) 効果量(ES): 帰無仮説と対立仮説のズレの量、(4) 検定力($1-\beta$): 対立仮説を正しく採択する確率である。根拠に基づく医療が求められている今日、研究報告の信頼性を高めることは緊急の課題となっている。

例えば「5%で有意」あるいは「1%で有意」と報告した論文があっても、その結果が適切かどうかは検定力の大きさに依存している。仮に「5%で有意」という結果を得たとしても、「検定力 0.3」だったならば、この結果は 30%の確率で得られるに過ぎないということになる。またサンプル数によって「有意」のあるなしが左右される。よって検定力分析(有意水準・サンプル数・効果量・検定力)が必要となる。検定力分析は「p 値、N、ES、 $1-\beta$ 」の 4 つの要素によって構成され、そのうちの 1 つの要素は残りの 3 つの要素の組み合わせによって導出されるという関係にある。必要なサンプル数「N」をあらかじめ割り出すためには、残りの「 α 、ES、 $1-\beta$ 」の 3 つが決まらないと計算できない。通常 p 値は 0.05 や 0.01 を用いるので問題はないが、ES と $1-\beta$ は別途定めなければならない。ES は 2 群間の平均値の差と 2 群共通の標準偏差から導出できる。また J. Cohen(1992)は「 $1-\beta$ 」を慣例的に「0.8」と設定している。

例えば、1 例としてフリーソフトの EZR(<http://www.jichi.ac.jp/saitama-sct/SaitamaHP.files/statmed.html>)を利用して事前検定と事後検定を行い、統計的検定を適切に用いることが可能である。もちろん、その他のソフトで

