

F. 参考文献

- 1) Anders Ahlbom and Staffan Norell: Introduction to Modern Epidemiology. 1. What is Epidemiology? , 1990; Epidemiology Resources Inc., Newton Lower Falls, MA, USA.
- 2) 大橋靖雄、荒川義弘編集：臨床試験の進め方. I. 1. a. 臨床医学と EBM、臨床試験 2006 ; 南江堂, 東京.
- 3) Stephen B. Hulley, Steven R. Cummings, Warren S. Browner, Deborah G. Grady and Thomas B. Newman: Designing Clinical Research. 3rd edition 2. Conceiving the Research Question, 2007; Lippincott Williams & Wilkins., Philadelphia, PA, USA.
- 4) 週間医学界新聞 第 2747 号. (2007 年 9 月 10 日)
- 5) Shirkey HC. Editorial Comment. Therapeutic Orphans. J Pediatr 1968; 2: 119-120.
- 6) 厚生科学研究 厚生省医薬安全総合研究事業「小児薬物療法における医薬品の適正使用の問題点の把握及び対策に関する研究」 - 「小児薬物療法における処方実態調査と医薬品添付文書解析」平成 12 年度研究報告書 (主任研究者: 大西鐘壽、分担研究者: 森田修之)
- 7) 厚生科学研究 厚生省医薬安全総合研究事業「小児薬物療法における医薬品の適正使用の問題点の把握及び対策に関する研究」 - 「医薬品の添付文書における子どもに関する記載内容についてのアンケート調査」平成 11 年度研究報告書 (主任研究者: 大西鐘壽、分担研究者: 森田修之)
- 8) UCSF-Department of Epidemiology and Biostatistics.
<http://www.epibiostat.ucsf.edu/courses/post.htm>
↓
- 9) UCSF-School of Medicine
http://medschool.ucsf.edu/news/features/research/20060911_CTSI_overview.aspx
- 10) UCSF-School of Medicine
<http://medschool.ucsf.edu/mstp/>
- 11) 座談会 社会全体が支える「医師」の養成 小中高等学校、患者が良医を育てる ; 文部科学時報 2008: 18-29.
- 12) Comments-BPCA from the FDA Pediatric Drug Development.
<http://www.fda.gov/cder/pediatric/PL107-109.pdf>
- 13) Comments-PREA from the FDA Pediatric Drug Development.
<http://www.fda.gov/cder/pediatric/S-650-PR-EA.pdf>
- 14) Comments-FDA Amendments Act of 2007 (FDAAA) from the FDA Pediatric Drug Development.
<http://www.fda.gov/oc/initiatives/HR3580.pdf>
- 15) Murphy D. Pediatric trials. The Impact of U.S. Legislative and Regulatory Effects. Applied Clinical Trials 2005; 14: 38-42.
- 16) 土田 尚. 小児領域の臨床試験と医薬品開発を促進するための海外の取り組み. 日本小児アレルギー学会雑誌. in press
- 17) ICH Topic E11 医薬品医療機器総合機構 HP より.
http://www.pmda.go.jp/ich/e/e11_00_12_15e.pdf
- 18) 厚生省医薬安全局審査管理課長通知. 小児集団における医薬品の臨床試験に関するガイダンスについて. (医薬審 第 1334 号. 平成 12 年 12 月 15 日付)
- 19) Regulation (EC) No 1901/2006 from the Official Journal of the European Union.

<http://www.emea.europa.eu/htms/human/paediatrics/regulation.htm>

- 20) Regulation (EC) No 1902/2006 from the Official Journal of the European Union.
http://ec.europa.eu/enterprise/pharmaceuticals/eudralex/vol-1/reg_2006_1902/reg_2006_1902_en.pdf
- 21) Paediatric Committee from the EMEA Medicines for children.
<http://www.emea.europa.eu/htms/human/paediatrics/pdco.htm>
- 22) Guidance for applicants-PIPs, waivers and modifications from the EMEA Medicines for children.
<http://www.emea.europa.eu/htms/human/paediatrics/pips.htm>
- 23) Guidance for applicants-Paediatric-use marketing authorisations (PUMAs) from the EMEA Medicines for children.
<http://www.emea.europa.eu/htms/human/paediatrics/pumas.htm>
- 24) 「医学教育の改善・充実に関する調査研究協力者会議」最終報告、平成19年3月28日付。

- 25) 報告 医学教育の改善・充実に関する調査研究協力者会議 高等教育局医学教育課；文部科学時報 2008: 40-43.

G. 研究発表

- 1) 第111回日本小児科学会学術集会医学教育3で「小児領域の臨床研究推進のための医学教育のあり方」について発表した（平成20年4月26日（土））
- 2) 第45回日本小児アレルギー学会シンポジウム6「アレルギー治療薬の小児適正使用に向けて」で「海外における取り組み」について発表した（平成20年12月13日（土））
- 3) 土田 尚. 小児領域の臨床試験と医薬品開発を促進するための海外の取り組み. 日本小児アレルギー学会雑誌. in press

H. 知的財産権の出願・登録状況

該当する事実・予定はない。

題名:臨床研究/試験の科学的分析が可能な人材育成に向けた
生物統計学の教育プログラム作成

分担研究者 大森 崇 京都大学大学院医学研究科医療統計学分野

研究要旨

医学部で学ぶ学生が、臨床研究に関して学ぶ時間は決して多くない。その中で臨床研究に必要な統計学の知識を学ぶ機会はさらに限られている。このような制約の中で、将来の臨床研究のための医療統計学の教育を効率よく行うことは課題である。

本研究では教材を用いた実習を通してランダム化臨床試験の基本的な考え方と、その統計解析の知識を身につけるための教育を考えることにした。データ解析に用いるデータを受講者自らが作成するには、時間的な制約からの困難が伴う。そこで、名刺サイズのカード1枚1枚に異なる対象者のデータを記載し、カードを選ぶという形式を考えることにした。

初年度は、カードを利用した教材を開発した。その内容は、連続値の2群比較を想定したもので、はじめにパイロット研究を行い、そこで得られたデータに基づいてサンプルサイズを見積もって、次の研究を行うというものである。統計学の教育としては、(1)パイロット研究では有意な結果は出ずに、(2)そこで得られた情報にもとづきサンプルサイズを設計した場合に有意な結果が得られる、ことを期待している。そこで、最終年度は、開発した教材が上記2点についてどの程度の性能を有しているかを検討し、それに基づき新たなカードを作成し直した。また、これを用いて実際にセミナーを行い、その評価を行った。

検討では、まず乱数に用いる平均と標準偏差に関していくつかの状況を設定し、検出力曲線を作成した。この検出力曲線から、カードに用いるモデルを決定し、実際に乱数を発生させた値を記入して教材のカードを完成した。次に、ヒストグラムを描き、要約統計量を計算することで、作成したカードに記載される値が想定したとおりのものになっているかどうかを確かめた。さらに、この値を用いて、実際の使用上を想定した100,000回のシミュレーションを行い、パイロット研究で想定している人数等の検出結果を調べた。その結果、1群5名のパイロット研究で有意となるのは、有意水準5%の両側検定で約20%であり、10名で約50%、20名で約90%であった。教材を使用し教育する者がこの結果を知っておくことは重要である。

また、午前中は講義、午後は教材を用いた実習という1日セミナーを実施した際のアンケートの結果は、ほとんどの受講者が内容に関して「満足」、時間に関して「ちょうどよい」というものであり、自由記載のほとんどが内容に前向きなものであった。これらの結果から、作成した教材は統計学の教育に有用なものとなっていると考える。

A. 研究目的

臨床研究/試験に携わる研究者は、事前に定めた研究仮説に答えることができるように研究を計画し、計画に基づき遂行した結果についてその解釈を行い、適切な結論を述べる必要がある。このような研究者を育成するには、医療統計学の知識が不可欠であり、そのための教育には、データ解析の方法のみならず、広い範囲での研究の計画の理解が必要である。

しかしながら、現在医学部の中で、臨床研究/試験のための統計学を教育する時間はかなり限定されていると思われる。その時間を増やすことを前提とし研究を行うことはややもすると現実的ではないかもしれず、また、研究期間内でそれを行うことはきわめて難しかった。また、仮に講義の時間を増やすことができたとしても、それによって得られる学習効果が高くなければ意味はない。このため、限られた時間の中で、効率的に医療統計学の知識を伝えることが課題である。

この課題に答えるために、本研究では、短時間で臨床研究/試験に携わる研究者を育成するための教育教材を開発することを考え、それを使用した教育を提案することを目的とした。教育教材の開発を考えた理由は、講義で聞いた内容を実習を通して考えることで理解するということが効果的だと思われるからである。現に、筆者が工学部の学生のとときには、紙ヘリコプターを用いた統計学の実験実習があった(Box & Liu (1998))。この実習は印象に強く残っており、このような教材が社会人を対象としてセミナーでも用いられている(高橋(1998))。これらの教材は、品質管理の教育のために開発されている。本研究では、このような教育教材を参考に、将来臨床研究

を行う者や現在臨床研究を行っている者を対象とした教材を開発して、教育に組み込むことを考えることにした。提案する教材を用いた教育を具体的に実施するために、本研究ではセミナーを開催することにした。その際、複数回のセミナーを行うと、すべてのセミナーに参加できる受講者は限られてしまうであろうことから教育の時間を1日と限定し、1日のセミナーでどこまで効率的に医療統計学の知識を伝えることができるかということに焦点をあてることにした。

教材を作成する場合、実際にその教材が多くの教育者に使われていくためには、長所、短所を含めてその教材がどのような特徴を持っているのか、実際に使用した場合にどのような反応を得たのかを明らかにする必要があるであろう。そこで、この研究では、教材そのものの性能評価を行うこと、また、実際にセミナーを開いて使用した際の受講者からの声を聞き評価を行うことも目的とした。

なお、筆者は分担研究者として平成19年度からこの研究に参加したため、ここでは、19年度に行った教材の開発と20年度に行ったその評価を報告することにする。

B. 研究方法

カードを用いた教材の開発

先に示した紙ヘリコプター実験は、学習者がいろいろな型の紙ヘリコプターを作成し、ストップウォッチを使い測定した滞空時間が解析するデータとなる。学習者はいろいろな形の紙ヘリコプターを作成し、その型の特徴と滞空時間の関係をデータ解析によって明らかにしていく。紙ヘリコプターを作成し、自分たちのデータを測定するという作業は、学習者にとって楽しさを感じる部分ではある。しかし、本研究で想定

している時間の制約を考慮すると学習者自らがデータを取るということは難しい。このため、事前に記載されているデータを利用し、かつ学習者が自らがデータを得れるようなものが望ましいと考えた。そこで、名刺サイズのカードにデータを記載したものを何枚か用意し、学習者がそのカードを選ぶことにし、カードに記載されたデータのデータ解析を行うということにした。

ここでは、ある降圧薬を開発するランダム化臨床試験を想定した。カードには、試験に登録された協力者のベースラインを示すものと、投与後を示すカードを用意した(図 1、図 2)。はじめにベースラインのカードをランダムにいくつかの群にわけ、それぞれの群で異なる介入結果のカードを選ぶことで、学習者は投与後の値を知ることができるようにした。

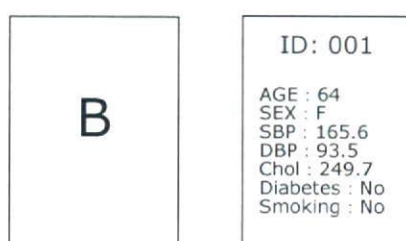


図 1 ベースラインカードの例
左：表面 右：裏面

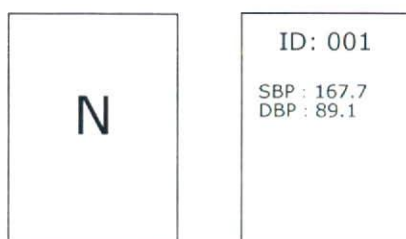


図 2 新薬カードの例
左：表面 右：裏面

当初作成したカード教材は、プラセボ、既存薬、新薬を比べるという臨床試験を想定したものである。実習には指定したソフトウェアで 3 群比較(分散分析や多重比較)を行う手順を示したテキストを作成した。作成した教材は研究期間中、講義とは別に 180 分の実習を行う機会があり、その際に試作した教材を利用することができた。その結果は、学習者は手順に従い、自らのデータを指定したソフトウェアで解析し、結果を得ることができたが、そのことで学習者が何を学習できたのを知ることは困難であった。

そこで、プラセボと新薬の 2 群比較のストーリーとし、グループワークによるサンプルサイズ的设计に焦点をあてた内容に変えることにした。そのコンセプトは、まず、1 群 5 人分のカードでパイロット研究を行うことにし、このデータに基づいてサンプルサイズ的设计を行う。そして、設計したサンプルサイズの人數分のカードを用いてデータ解析を行う、というものである。サンプルサイズ的设计は臨床研究の計画を行う上で重要な項目である。いくつかの市販のソフトウェアでは、平均値の差、標準偏差、検出力、第一種の過誤確率を入力するとサンプルサイズの値を出力する機能があり、簡単に計算を行うことができる。カード教材でこの項目を取り組むことで、学習者は平均値の差や標準偏差をどのような値で見積もるか、検出力を多少変えたときにどの程度サンプルサイズが変わるかなどを検討することが可能となる。

カードに記載するデータのモデル

カードに記載するデータは、降圧薬の特徴を考慮して収縮期血圧のモデルを作成し、乱数を用いて発生させた(藤井、芦田(2002))。

ベースラインである投与前の収縮期血圧の

値は、年齢が 40 歳のときに 160mmHg になることを基礎として、年齢、性別、総コレステロール、糖尿病の有無、喫煙状況によって値が異なるようにし、さらに正規乱数で作成される誤差を加えるモデルとした。投与後の測定値は、同じ ID の投与前の測定値に薬剤もしくはプラセボの効果の固定効果と正規乱数により作成した誤差を加えたモデルとした。

薬剤群とプラセボ群の降圧効果の差を D とし、投与前と投与後のモデルに加える誤差は、それぞれ平均を 0、標準偏差を σ_b 、 σ_a の正規分布に従う乱数とした。投与後から投与前の変化量をエンドポイントとしたとき、その標準偏差は $\sigma = \sqrt{\sigma_a^2 + \sigma_b^2}$ となることが期待される。

検討 1 カードに記載するデータの作成

教材として期待することは、パイロット研究として 1 群 5 例で検討した際には有意な差がつきにくく、データをもとに検討してサンプルサイズを見積もった際には有意な差がつくことである。このようになるような D と σ を決める必要がある。そこで、いくつかの D と σ についてサンプルサイズと検出力の関係を調べることを検討することにした。

具体的には D は 8、9、10 の場合を、 σ は $7.8 \left(= \sqrt{5.5^2 + 5.5^2} \right)$ 、 $9.9 \left(= \sqrt{7^2 + 7^2} \right)$ 、 $12.0 \left(= \sqrt{8.5^2 + 8.5^2} \right)$ とした。

検討 2 作成したカードに記載される値の分布

作成したカードは検討 1 の結果得られる検出力が期待される。しかし、それはあくまでも平均的に期待できる検出力であり、ひとたびカードを作るためのデータを決めてしまえば、カ

ードに記載された値にすべての結果が依存することになる。そこで、投与前、プラセボ投与後、新薬投与後に相当する各 100 枚のカードに用いる値から計算される、プラセボ投与後または新薬投与後の投与前からの変化量の分布、2 群間の変化量の差の分布を調べた。

この検討での、変化量の差は、同じ ID のものが新薬を投与されたときとプラセボを投与されたときの差として算出した。つまり、モデルは個々の ID は年齢や性別などが血圧値に影響するように作成したが、この定義での差はこれらの共変量の影響は除外したときの変化量の差をみていることになる。さらに、この差は個人の薬剤の効果としてみることができるが、教材の内容では現実の状況を意識して、同一 ID が新薬かプラセボのどちらかのみが投与されることが想定されている。つまり、この状況の 2 群の差は理想的な状況であり、現実の状況とは異なっている。

検討 3 作成したカードの性能評価

先に記載したように、教材の観点からは、作成したカードを各群 5 枚使い行うパイロット研究では、検定を行ったときに有意にならないことが望ましい。そこで、作成したカードのパイロット研究での検出力がどの程度かを調べることにした。投与前の 100 枚分のカードから 10 枚ずつランダムに選び、はじめの 5 枚に対応する ID はプラセボ投与の値を、残りの 5 枚の ID は新薬を割り付けることを想定し、有意水準 5% の両側 t 検定を行うという試行を 100,000 回繰り返した。

さらに、上記と同様の検討を、1 群 10 例とした場合、20 例とした場合、30 例とした場合を想定して行った。

検討4 カード教材を用いた実習の評価

上記のような検討を行い、カード教材を作成したが、実際にこの教材を用いた際に学習者がどのように感じたかを知ることも重要である。そこで、2008年8月に滋賀医科大学、2009年1月に国立成育医療センターで、午前中はパワーポイントを用いた講義、午後は教材を用いた実習という形式の1日のセミナーを行った。どちらのセミナーも午前中は2時間半から3時間かけて講義を行った。講義では「ランダム化臨床試験」、「2群間の比較」、「信頼区間」、「サンプルサイズ的设计」に関する4種類のパワーポイントスライドの用意した。「ランダム化臨床試験」では、ランダム化比較試験とは何か、なぜランダム化をして比較する必要があるのかについて、「2群間の比較」では検定の考え方とt検定とはどのような検定かについて、「信頼区間」では検定結果だけではなく区間推定による報告を行うことについて、「サンプルサイズ的设计」ではサンプルサイズ的设计に必要なパラメータと意味を説明した。

午後の実習は、グループ分けを行い、3時間から4時間使い、グループごとにパソコンを1台準備して、教材を用いた実習を行った。

セミナーの後に、受講者にアンケート調査票を配布し、

- 1) 受講して感じた問題点(複数回答可)
- 2) 受講して感じたメリット(複数回答可)
- 3) 内容について
- 4) 時間について
- 5) 医療統計学について感じる事、セミナーについて思ったこと(自由記載)

について尋ねた。

C. 結果

作成した教材の内容

作成した教材は、カルシウムチャンネル拮抗薬を開発するという事を想定し、プラセボ新薬との比較を考えるというストーリーである。カードはベースラインを示すカードと2種類(プラセボ群、新薬群に相当)の介入後を示すカードに分かれており、それぞれ100枚ずつある。

ベースラインのカードには、一枚一枚が個々の臨床試験対象者の薬剤割付前のベースライン時点の測定値に相当する値を記入した。ID番号と年齢、性別、収縮期血圧、拡張期血圧、総コレステロール、糖尿病の有無、喫煙状況である。残りの介入後のカードには、ID、群、介入後の収縮期血圧、拡張期血圧の測定値に相当する値を記載した。学習者は4から6名程度のグループを単位とすることを前提とした。

パイロット研究として、まず、シャッフルしたカードの束から10枚選ぶ。そしてそれらを再びシャッフルし、2つの束に分け、分けた束ごとに、異なる投与群に相当する介入後のカードのIDの割り当てることにより、群分けを行う。学習者はこれらの値をソフトウェアに入力し、各グループごとの平均値や標準偏差を算出したり、2群間の比較を行う。

次に、パイロット研究のデータをもとに、次の研究で有意な差が得られるようにサンプルサイズ的设计を行う。その際、ソフトウェアの操作だけでなく、グループ内で十分にデータを吟味する時間をとり、いくつかのシナリオの下で検出力やサンプルサイズの検討を行い、最終的なサンプルサイズを決定する。

最後に、決定したサンプルサイズの枚数だけカードを選び、パイロット研究と同じようにデータを入力して、解析を行う。

検討1 カードに記載するデータの作成

図3~5は、Dと σ についてそれぞれ3つの条件を指定したときの、横軸に1群あたりのサンプルサイズ、縦軸に検出力をとったグラフである。

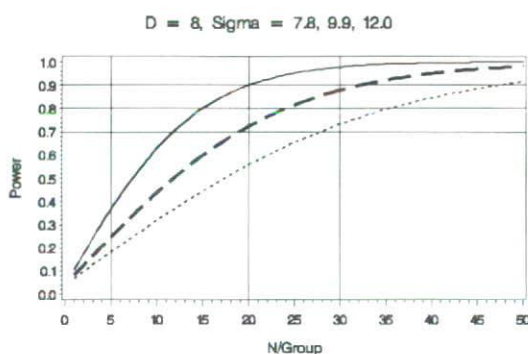


図3 Dを8とした場合の検出力
実線: $\sigma=7.8$, 破線: $\sigma=9.9$, 点線: $\sigma=12.9$

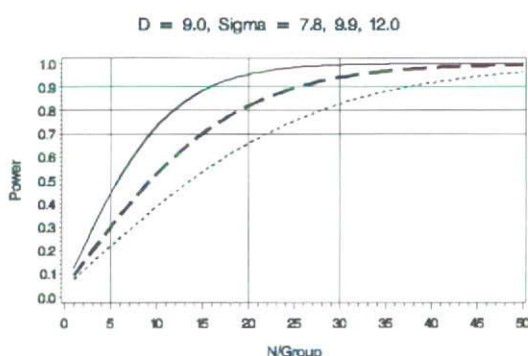


図4 Dを9とした場合の検出力
実線: $\sigma=7.8$, 破線: $\sigma=9.9$, 点線: $\sigma=12.9$

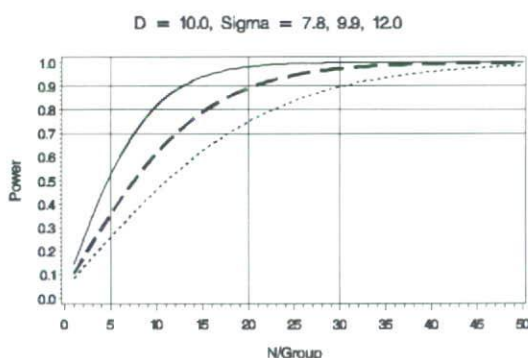


図5 Dを10とした場合の検出力

実線: $\sigma=7.8$, 破線: $\sigma=9.9$, 点線: $\sigma=12.9$

上記の3つグラフからD=9で $\sigma=9.9$ の時に1群あたり5人で30%、20例で80%、30例で90%程度の検出力であることがわかった。

そこで、新薬投与後の血圧降下を -8.5mmHg 、プラセボの効果は 0.5mmHg とし、投与前、投与後のそれぞれの標準変偏差を 9mmHg と設定して、値を算出させ、カードを完成させた。

検討2 作成したカードに記載される値の分布

図6は検討1に基づき作成されたカードに記載されたプラセボ投与後、新薬投与後に相当する値から計算されるプラセボ投与後または新薬投与後の投与前からの変化量の、2群間の差の分布である。表1はこの基本統計量を示している。

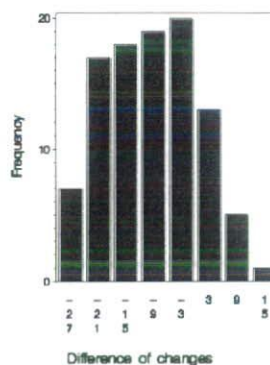


図6 変化量の差の分布

表1 各分布の基本統計量

	N	平均	標準偏差	最小値	最大値
変化量の差	100	-9.4	10.1	-29.0	14.7

表1に示すように、変化量の差は -9.4 とな

っており、検討 1 で検討した設定よりもやや大きな差となっていた。

検討 3 作成したカードの性能評価

作成したカードを重複を許さずにランダムに 10 枚選び、半数にプラセボ、残りの半数に新薬を投与することを想定した施行を 100,000 回繰り返した。図 7 は、各回の変化量の群間差(縦軸)および併合分散により計算された標準偏差(横軸)の散布図である。

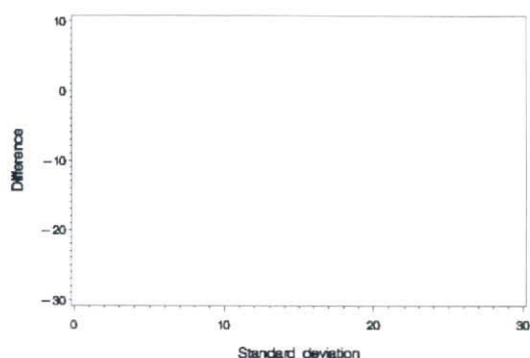


図 7 変化量の群間差と標準偏差

このデータを用いて、5%有意水準で両側 t 検定を行った結果、有意になった割合は 20.9%であった。表 2 は、このデータから 70、80、90、98%の検出力で、1 群あたりのサンプルサイズを計算した際に、どのくらいのサンプルサイズが見積られるかを割合で示したものである。

表 2 パイロット研究の値によって見積られる 1 群あたりのサンプルサイズの割合(%)

人数	検出力(%)			
	70	80	90	98
>45	9.7	11.8	15.2	22.3
≤45、>25	6.3	8.0	10.6	15.3
≤25、>5	44.3	49.2	52.9	51.9
≤5	39.7	30.9	21.2	10.5

また、作成したカードを重複を許さずにランダムに 20、40、60 枚選び、半数にプラセボ、残りの半数に新薬を投与することを想定した施行をそれぞれ 100,000 回繰り返し、有意水準 5%での t 検定で有意になる割合を計算した。その結果は、それぞれ 51.2、92.0、99.7%であった。

検討 4 カード教材を用いた実習の評価

1) 滋賀医科大学での結果

受講者 23 人中、調査票の回答者 21 名の内訳は、4 年生 1 人、5 年生 11 人、その他 8 人、大学院生 1 人であった。夏休みの特別セミナーとしてセミナー行ったこともあり、対象者全員が医療統計学への関心は「あり」と回答していた。

表 3 から 7 にセミナーに関する回答を示した。

表 3 受講して感じた問題点(複数回答)

特になし	基礎知識がない	内容が難しい
6	13	2

表 4 受講して感じたメリット(複数回答)

特になし	研究に役立ちそう	医学論文を読む上で有用	統計学を身近に感じた	その他
0	9	10	14	5

表 5 内容について

満足	やや満足	不満	計
17	4	0	21

表 6 時間について

長すぎる	短すぎる	ちょうどよい	計
1	2	18	21

表 7 自由記載

自由記載
ぜひ今後第2回、第3回と行ってほしいです。とても勉強になりました。レクチャーとディスカッションのバランスもよかったです。
実際に作業をやることで理解が深まったように感じます。ありがとうございました。
1コマ目よりも他の統計の考え方のところにも少し時間をかけていただきたかったです。値が大きいとどうなのか、その意味は、などです。
難しすぎるとまったくわからないし、あまり省略されると理解できないし、どうやって勉強したらいいかわからない！
サンプルサイズ、標準偏差を自分達で決定してから解析するという流れが初めての経験で、今までの統計学の勉強とは異なる視点で統計学を勉強できてよかったです。ありがとうございました。
今日をふまえてまた実習を通して実践的に活用していきたいです。お弁当がとてもおいしかったです。
論文を例にして読み方を教えてほしいです。次回セミナーを開くなら早めにしてほしいです。
統計の基礎的な知識が欠けている私ですが、前提となる考え方がきっちり説明して下さってなんとかついていくことができました。論文を読むときに、書いてあることをうのみにするのではなく、信頼できるデータかどうかを見極めることも大切だと感じました。
サンプル数の見積もり方法は以前から気になっていたトピックでした。実習でもとり入れていただき勉強になりました。
論文を使った統計の読み方を次回してほしい。
具体的な医学論文の読み方なども教えていただき良かったです。
はじめてt検定の意味が分かりました。とても理解しやすく教えていただいて、ありがとうございました。勉強になりました。
演習はサンプルサイズ計算の原理を考える上でとてもよかったです。
統計学は奥が深いと思いました。身近にあるデータやサンプル数には大きな意味があることを知り、実習で設計したことで、考えそのものが理解できました。
是非アドバンスコース等を企画していただきたい。
食わず嫌いになりがちですが、本日の講義のこの余韻がずっと残ることを願いつつ、日々応用につなげたいです。又是非次回受講したいです。
さならる知識向上に努めたい
t検定。とてもわかりやすく説明があり楽しかったです。なぜ？というところの理解ができました。サンプル設計については難しかったです。(楽しくグループワークはできました。)

2) 国立成育医療センターでの結果

受講者 37 人中、調査票の回答者 36 名の内訳は、臨床医 22 名、研究者 7 名、パラメディカル 1 名、その他 5 名であった。休日に行っ

たセミナーだったこともあり、35 名が医療統計学への関心は「あり」、1 名が「その他」と回答していた。

表 8 から 12 にセミナーに関する回答を示した。

表 8 受講して感じた問題点(複数回答)

特になし	基礎知識がない	内容が難しい
13	13	2

表 9 受講して感じたメリット(複数回答)

特になし	研究に役立ちそう	医学論文を読む上で有用	統計学を身近に感じた	その他
0	22	14	19	1

表 10 内容について

満足	やや満足	不満	計
30	4	1	35

表 11 時間について

長すぎる	短すぎる	ちょうどよい	計
1	2	32	35

表 12 自由記載

自由記載
大変わかりやすい内容でとけ込みやすかったです。続編を期待しています。次は多変量解析について知りたいです。
セミナーの前に勉強すべき内容を教えてもらっておくともっと多くの内容が理解できると思います
昼休みはもっとみじかくてもよい。
pre learningの教材があるとよいのでは？
自分は医学部1, 3年で統計の講義があったが臨床に出た後とか、卒後教育としてカリキュラムが組まれたらより効率的だし役に立つと感じました
どの対象にどの検定を使っていいかが臨床で一番悩んでいるところなので、"どの対象にどの検定を使うか"を教えてください
継続してレベル別の講習をしてほしい
時々思いついたように勉強するだけなので身につかないだと思います。今回の講義でモチベーションが上がりました。
グループで実習ができたことで、医療統計がかなり身近になりました。どうもありがとうございました。
統計は恐ろしい。生兵法にならないためにはよく勉強する必要があります
医学生の頃に言葉を聞いたことがある程度で、実地臨床で統計学を学ぶ機会がとても少ないように思います。今後とも続けて頂ければ幸いです。
大変有意義なセミナーでした。ありがとうございました。
第2回、3回目を期待します。
数字のマジック…というか、基本がわかっていないと勝手な解釈をしてもそれなりに感じてしまうような危なさを感じた。
PCが一人1台あたるとような人数にしてほしい
基礎知識が全くなかったので講義のみだと自分の理解度に不安だったのですが実際に実習することでイメージが掴めてよかったです。ありがとうございました

D. 考案

本教材で期待している学習内容

本研究では、現状の教育の実情を考慮して、臨床研究/試験を行う研究者の統計学の教育を行うために、教育のための教材を開発して、その教材を用いたセミナーを実践した。

この中で、統計学の知識として含めているものは、極めて限定されている。また、講義でも数理的な事項はほとんど触れてはいない。ここで想定したように教育時間を1日に限定したとしても、講義の時間を増やすなどして、より広い知識を伝えたり、数理的な事項に触れること

も可能である。しかし、そのようなことをせずに実習を中心にしたのは、いくつかの理由がある。その一つは、1日で広い知識を伝えることでどれほど学習効果があるのかが明確ではないことである。教科書の目次を紹介して終わるような内容にすれば、非常に多くのことを教えたことにはなるかもしれないが、学ぶ側からすれば、それが学んだことになるとは思えない。もう一つの理由は、この研究で育成対象にしている者は、臨床研究の専門家であり、統計学の専門家ではないことである。このことから数理の部分よりもソフトウェアの出力を読むということを重視することにした。

特に、教材はサンプルサイズの設定に焦点をあてており、その部分に特化していると言っても過言ではない。このようにした理由の一つは、サンプルサイズの設定が将来臨床研究に携わる研究者がプロトコルを作成する段階で必要な知識であるからであり、現に関心が高い項目であるからである。サンプルサイズの設定は、ソフトウェアに適切な値を人力しさえすれば得ることができる。しかし、現実的なサンプルサイズの設定を行うためには、適切な見積もりを行う必要がある。そして、そのためにはデータの解釈やデータ解析が必要である。この教材のもう一つのねらいはここにある。学習者は、パイロット研究の結果から見積もりを得るためにデータ解析を行い、結果の解釈を行わなくてはならない。この教材では SAS 社の統計解析ソフトウェア JMP を用いることを前提としているが、JMP はわずかに数回のクリックを行うだけで膨大な出力が得られる。学習者は、この出力の中から自分の必要な情報を読み取らなくてはならない。ここでは出力される統計量の意味が学習できることを期待している。

各検討の考察

昨年度の報告にも示したとおり、この教材は実際に教材を使用した教育を実践しながら発展させてきた。検討 1 から 3 は実際の教材使用の場で生じた問題を解決するために行ったものである。

検討 1 の結果から、 $D=9$ 、 $\sigma=9.9$ と設定し、乱数を発生させたカードを作成したが、後から乱数を発生させるシードを変えると値がかなり変わることから、得られた値が狙ったとおりになっているかどうかを確かめる必要があることに気づいた。これは、作成するカードがわずか 300 枚 (=100 人分×3 つの条件) しかないことに起因していると考えられる。このため、検討 2 で示した最終的な結果を確定するまでに、表 1 に記載した値 (特に変化量の差の平均値) を見ながら何度か乱数を発生させる必要があった。

さらに確認のために、実際の実習で使われる状況を想定して検討したものが検討 3 である。この検討から、パイロット研究の段階で有意差が得られるのはわずかであることが確認できた。しかしながら、両側検定の結果から、新薬群に効果がある方向に有意になるのは約 20% あることもわかった。また、図 9 からパイロット研究でプラセボ群の方が降圧降下を示すような結果がでることもあることや、表 2 からサンプルサイズを見積もっても、用意してあるカードでは足りなくなってしまうたり、5 例以下になってしまうことがあることもわかる。実習は複数のグループで行うことを前提としているので、これらの結果はいくつかのグループで生じる可能性があることを示しているともいえる。カードを用いて実習を行うものは、この点をよく理解して、実習をすすめる必要がある。筆者の場合には、わずか 1 群 5 名の検討であるから不安定だと

いうことを十分吟味してサンプルサイズを行うように言うようにしている。また、パイロット研究でプラセボ群の方が効果があるような結果が得られたグループには、この段階で新薬に見込みがないと判断するのも一つの考えであるからグループ内で十分に話し合って先に進むことをやめるとすることも選択肢の一つであることを言うようにしている。

実際、カードを用いたセミナーは、このようなグループがあった。滋賀医科大学のセミナーでは、本報告で示したカードの前のバージョンであったため、その性能が十分にはわからないが、成育医療センターのセミナーでは、本報告でその性能を示したカードを用いることができた。その結果は 10 グループ中、3 グループがパイロット研究で有意な差が生じた。ほぼ想定したとおりの結果が得られたことになる。

この研究では、教育が行える時間は 1 日という限定を課して教材の作成に取り組んだ。検討 4 はカードを使った実習を含むセミナーの評価である。いずれも医療統計学に興味があるという受講者であったという制約ではあるが、表 5 から 12 からは、教材を用いたセミナーの評価は悪くはないといえるであろう。詳細については、1 日のセミナーであるため、統計学特有の用語が一度に多くでてくることでのとまどいが生じやすいことや、学生は臨床研究をイメージすることが難しいので論文の読み方につながることを期待していることがわかった。これらは、今後セミナーを行う際の準備で考慮する必要があるといえる。

小児の臨床研究推進のために

この研究は小児の臨床研究推進に必要な人材育成を目的にしているが、ここで作成した教材そのものは、小児の臨床試験で生じる困

難さ(Benjaminら(2006)、Bush(2006)、Caldwellら(2004)、Jacobson(2007)、Word & Kauffman(2007))が反映できていない。それは、本邦では小児領域で医薬品開発や治験を含めた臨床研究に精通した医師が不足している状況を考慮すると、現時点でできることとして、臨床研究を行っていく人材の育成が必要であろうということを優先した結果である。したがって、行ったセミナーでの講義や実習の内容は、小児の臨床試験の特殊性はほとんど含まれていない。しかし、現実の小児の臨床研究の困難さを無視すべきではないであろう。そこで、実習テキストの最後に課題として「この実習で行うほど単純ではありませんが、新薬が認められるまでには臨床試験が実施されます。しかし、多くの場合対象者は成人であり、小児が対象となる臨床試験はほとんど実施されていません。この理由を考えてみましょう。それらをもとに、小児の臨床試験を実施するにはどのようにすればよいかを考えてみましょう。」という課題を含めた。さらに、参考となる文献を示し、関心があるものはこの課題を考えることができるようにした。

E. 結論

限られた時間で医学部生を中心に臨床研究を行なう上で必要な統計学の教育を行うための教材を作成した。開発した教材はこれらの対象者の医療統計学の教育に有用である。

F. 参考論文

Altman DG, Machin D, Nryant TN, Gardner MJ. Statistics with confidence, 2nd ed. BMJ books, 2000.
Armitage P, Berry G, Matthews JNS.

Statistical methods in medical research, 4th ed. Blackwell Publishing, 2002.

Benjamin DK, Smith PB, Murphy MD, Roberts R, Mathis L, Avant D, Califf RM, Li JS. Peer-reviewed publication of clinical trials completed for pediatric exclusivity. Journal of the American Medical Association 2006; 296: 1266-1273.

Box GEP, Liu PYT. Statistics as a catalyst to learning by scientific method part I – An example. Journal of Quality Technology 2999; 31: 1-15.

Bush A. Clinical trials research in pediatrics. Pediatric Drugs 2006; 8: 271-277.

Caldwell PHY, Murphy SB, Butow PN, Craig JC. Clinical trials in children. Lancet 2004; 364: 803-811.

Matthews JNS. Introduction to randomized controlled clinical trials, 2nd ed. Chapman & Hall/CRC, 2006.

Jacobson RM. Pediatrics and evidence-based medicine revisited. The Journal of Pediatrics 2007; 150: 325-326.

Ward RM, Kauffman R. Future of pediatric therapeutics: Reauthorization of BPCA and PREA. Clinical Pharmacology and Therapeutics 2007; 81: 477-479.

藤井潤, 芦田映直. 降圧薬の選び方と使い方 改訂第3版, 南港堂, 2002.

高橋武則. 模擬生産・模擬実験と統計的品質管理, 日本規格協会, 1998.

G. 研究発表

・大森崇, 安藤友紀. カードを用いた医療統計学の学習教材, 2008 年度統計関連学会連合大会.

Data
Card

カードを用いた 医療統計学の学習教材

大森崇¹、安藤友紀²

¹京都大学大学院医学研究科社会健康医学系専攻

²医薬品医療機器総合機構

おことわり

(個人の見解であり、
所属する団体の見解ではありません)

統計関連学会大会 2008/9/8

2

この研究は

- 「小児の臨床研究推進に必要な
人材育成と環境整備のための
教育プログラム作成」
(主任研究者 中川雅生)
の支援を受けています。

統計関連学会大会 2008/9/8

3

はじめに (大森の環境)

- 医学研究科社会健康医学系専攻
- 社会健康医学を学ぶ各分野の学生に
授業を提供
 - 医療統計を学ぶ学生に限らない
- 中川班の分担研究者

統計関連学会大会 2008/9/8

4

はじめに (安藤の環境)

- 医薬品医療機器総合機構
- 行政官、医者や薬学専門家らとともに
新薬の審査、開発の相談
- 最近では、~~中川班で~~統計セミナーを主催

詳しくは書籍コーナーでの
パンフレットで!

統計関連学会大会 2008/9/8

5

はじめに

- 統計学を専門としない人の需要
- 数理的な事項の詳細よりも、
臨床研究の報告内容の理解
 - デザインや解析結果
- 臨床研究を体験することが有益、
しかしそのような時間は現実にはない

統計関連学会大会 2008/9/8

6

はじめに

- パワーポイント、統計ソフトの利用にはやや限界を感じる
- 答えが合う、合わないというより、考える時間を確保した演習のようなものが望ましいのでは？

統計関連学会大会 2008/9/8

7

この研究の目的

- 対象者
 - 医学部、薬学部等の学生の臨床研究に関する生物統計の知識を必要とするもの
- 短期間でその要素を学ぶための教育教材を作成し、評価する

統計関連学会大会 2008/9/8

8

カードを用いた教材のコンセプト

- 市販の名詞作成カードに擬似的な臨床研究データ(乱数)を記載
- カードをランダムに選ぶことでランダム化
- 汎用統計ソフトを利用しデータ解析

統計関連学会大会 2008/9/8

9

カードを用いた教材

- 身近な例
 - 降圧薬を開発する臨床研究を想定
 - エンドポイントは収縮期血圧
 - 2群比較
- 実習手引きに沿って作業を進める
- グループワーク

統計関連学会大会 2008/9/8

10

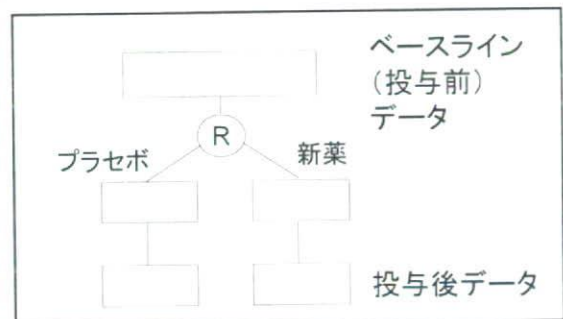
カードの記載例

ベースライン	プラセボ 投与後	新薬投与後
ID: 001 AGE: 64 SEX: F SBP: 165.6 DBP: 93.5 Chol: 249.7 Diabetes: No Smoking: No	ID: 001 Group: P SBP: 166.9 DBP: 100.0	ID: 001 Group: N SBP: 167.7 DBP: 89.9

統計関連学会大会 2008/9/8

11

カードを用いた教材



カードを用いた教材(試作段階)

- 試作版
- 選んだカードに記載されたデータをソフトに入力して解析
- 考える要素が少なくソフトの使用法の説明と変わりなかった
→改善が必要

統計関連学会大会 2008/9/8

13

教材の内容 (Version. 1.0)

- はじめは少ないカードでパイロット研究
➢1群あたり5例
- パイロット研究のデータをもとに学習者はサンプルサイズを設計
- 設計したサンプルサイズをもとに再度ランダム化し、データ入力後解析

統計関連学会大会 2008/9/8

14

Version 1.0の適用(2007秋)

- 授業(180分)で使用
➢時間的に問題はなし、良好な反応
➢方法については事前に講義が必要

統計関連学会大会 2008/9/8

15

Version 1.0の問題点

- ① 4グループのうち2グループがパイロット研究の結果で有意差
- ② とにかく多くのサンプルサイズを見積もりたがる
→改善が必要

統計関連学会大会 2008/9/8

16

問題①の考察

モデルの設定下での検出力

5例/群: 76.2% ←高すぎる

8例/群: 95.0%

12例/群: 99.5%

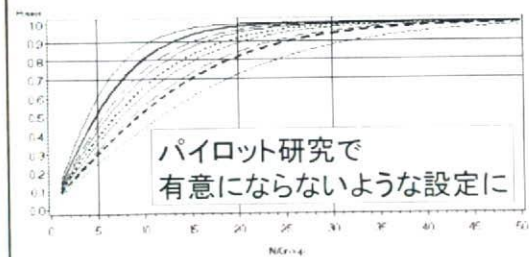
こういう検討をきちんとすべきだった。。。。

統計関連学会大会 2008/9/8

17

問題①改善のための検討

条件を変えて、検出力の検討



統計関連学会大会 2008/9/8

18

問題②の改善

- 少しだけゲームの要素を入れる
「統計的に有意な差が出る
かつ
一番小さなサンプルサイズ」
のグループを目指して
実習に取り組むように示唆

統計関連学会大会 2008/9/8

19

その他の改善

- 実習の結果は
CONSORT (Consolidated Standards of Reporting Trials) 声明に沿った報告の作成を課題に
➢ 臨床研究の報告を学ぶ

<http://www.consort-statement.org/index.aspx?o=1011>

統計関連学会大会 2008/9/8

20

この教材 (Version 2.0) で学べること

- ✓ 臨床試験における統計的データ解析の基礎
- ✓ サンプルサイズ的设计
- ✓ 研究報告の内容を理解する

統計関連学会大会 2008/9/8

21

Version 2.0の適用

- 8月20日に滋賀医大でセミナーを開催
➢ 医学生13人、その他10人
➢ 午前中はパワーポイントによる講義
RCT、t検定、95%CI、サンプルサイズ
➢ 午後が6グループで実習

統計関連学会大会 2008/9/8

22

Version 2.0の適用

- なかなか良好な手ごたえ
➢ 他の場所でも適用してみることに
- しかし、その評価はどうする？

統計関連学会大会 2008/9/8

23

考察

- 対象者の興味は統計学ばかりではない
➢ 「時間が限られている」ということは重要な要素
- 教材を用いた実習は考える時間を重視しているが、これは合理的か？
➢ 多くの知識を伝える方が望ましいか？

統計関連学会大会 2008/9/8

24

考察

- どのトピックをどのように教えるか
 - サンプルサイズ的设计は需要がある
 - 差やばらつきを見積もるグループ作業はおそらく印象に残る
- その次の機会はどこで？だれが？
 - アンケートでは、論文の読み方、応用編などの希望

統計関連学会大会 2008/6/8

25

考察

- 作成した教材はまだ改善の余地あり
- 対象者によって変えていくことが望ましい
- 教材の評価をどう行うかは課題

統計関連学会大会 2008/6/8

26

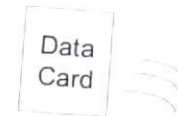
まとめ

- テキスト作成や講義のみで、人材育成は困難
- 学習者が(擬似的な)体験を通じて、臨床研究の基礎を学ぶ教育
- カードを用いた教材の導入とその検討

統計関連学会大会 2008/6/8

27

ご清聴ありがとうございました



統計関連学会大会 2008/6/8

28

参考にした文献

- Aliage M, Cobb G, Cuff C, Garfield J, Gould R, Lock R, Moore T, Rossman A, Stephenson B, Utts, J Velleman P, Witmer J. The GAISE (Guidelines for Assessment and Instruction in Statistical Education) college report, The American Statistical Association 2004.
- Altman DG, Machin D, Bryant TN, Gardner MJ. Statistics with confidence, 2nd ed. BMJ books, 2000.
- Berenson ML, Utts J, Kinard KA, Rumsey DJ, Jones A, Gaines LM. Assessing student retention of essential statistical ideas: perspectives, priorities, and possibilities. The American Statistician 2008; 62, 54-61.
- Matthews JNS. Introduction to randomized controlled clinical trials, 2nd ed. Chapman & Hall/CRC, 2006.

参考にした文献

- 大村はま、苅谷剛彦、苅谷夏子. 考えることの復権. 筑摩書房 2003.
- 津谷喜一郎、小島千枝、中山健夫訳. CONSORT声明: ランダム化平行群間比較試験報告の質向上のための改訂勧告. JAMA日本語版 2002; 6月号, 118-124.
- 芳賀敏郎、橋本茂司. 実験データの解析(2). 日科技連 1990.
- 藤井潤、芦田映直. 降圧薬の選び方と使い方 第3版. 南江堂 2002.

2008年8月20日 厚生労働科学研究費補助金 医療技術実用化総合研究事業 臨床研究基盤整備研究「小児の臨床研究推進に必要な人材育成と環境整備のための教育プログラム作成」

医学生のための医療統計学セミナー

模擬研究をおこなってみよう

～連続値データの2群間の比較～

0. はじめに

さあ午後は実習の時間です。この実習は班ごとに行います。チームワークが重要ですから班のみんなで相談し、協力しながらすすめていきましょう。

この実習の目標は

- 統計ソフトを使った2群間の比較の操作方法を知る、
 - サンプルサイズの設計方法を知る、
 - 調べたいことと対応した解析方法を考えてみる
- です。ではさっそくはじめましょう。

1. 状況設定

みなさんは今、新しい降圧薬の効果を調べるというプロジェクトに参加しているとしましょう。このプロジェクトでは、用量設定が終わり、降圧薬の有効成分を除いた偽薬（以下：プラセボ）、新しい降圧薬の降圧効果を比較する段階にあります。主要な評価項目として収縮期血圧とすることが決まったのですが、どのくらいの効果が得られるのかがはっきりせず、このためサンプルサイズが設計できません。そこで、10例（プラセボ5例、実薬5例）によるパイロット研究を行うことによって、差と標準偏差を調べて、それにもとづいたサンプルサイズを設計することにしました。さらに、設計したサンプルサイズで行なった研究結果の解析を行ないましょう。

2. カードでの模擬研究（その1）

2.1 模擬研究の準備 カードボックスに、「B」、「N」、「P」というカードの束が入っています。「N」と「P」は順番にならんでいますから、順番を変えないでください。

「カードB」には、ID、年齢（AGE）、性別（SEX）、収縮期血圧（SBP）、拡張期血

圧 (DBP)、総コレステロール (Chol)、糖尿病の有無 (Diabetes)、喫煙状況 (Smoking) のデータが記載されています。残りのカードには、ID、群 (Group)、収縮期血圧 (SBP)、拡張期血圧 (DBP) のデータが記載されています。

まずは班で以下の役割を決めましょう。

割付担当者 (1人) : カードボックスを管理して、カードを配布します。

データ入力担当者 (1人) : データを入力します。

補助作業担当者 (数人) : 作業の補助を行います。

役割分担が終わったら、割付担当者はカードボックスから、「カード N」、「カード P」をとりだし、2人の補助作業担当者が束のいずれかを持つように渡してください。

2.2 模擬研究 1 割付とベースライン値の入力

課題 パイロット研究に同意した 10 人の対象者を、ランダムにプラセボグループ 5 人、実薬グループ 5 人に 2 つの群に分けましょう。

- 1) 割付担当者は「カード B」を取り出し、よくシャッフルしてください。そして、上から 10 枚選び、残りのカードをカードボックスに戻しましょう。
- 2) データ入力担当者は、JMP を立ちあげ、データテーブルに ID、投与前の収縮期血圧を入力するための変数「ID」、「SBP_B」を作成しましょう。そして 1) で選んだ 10 枚の「カード B」に記載してあるデータを誰かに読み上げてもらい、10 人分のデータを入力しましょう。(ID の大きさの順に入力した方があとで楽かもしれません。)
- 3) 入力が終わったら、割付担当者はデータを入力した 10 人分の「カード B」を再びよくシャッフルして、「カード N」と「カード P」を持っている補助作業担当者に 5 枚ずつ配りましょう。これで「カード N」を持っている補助作業担当者に配られた ID の人が N 群に、「カード P」の補助作業担当者に配られた ID の人が P 群に割り付けられたことになります。

課題 さて、10 人の人たちは、P 群にはプラセボ、N 群には新しい降圧薬が処方されたとしましょう。みなきちんと処方された薬を飲んで 2 ヶ月経ち、新たに収縮期血圧の値が得られたとします。この 2 ヶ月後の測定値を入力しましょう。

- 6) 補助作業担当者は、自分の持っているカードの束から配られた「カード B」に記載してある ID と同じ ID のカード 5 枚を選びましょう。(残りのカードは輪ゴムでとめておいてください。)
- 7) データ入力担当者は、群と 2 ヶ月後の収縮期血圧を入力するための 2 つの変数、