

# Editing Your Work

17

1 assays vs TST in the new healthcare workers. They also found that the QFT-G and QFT-GIT  
2 strategies are more effective and less costly than the TST, whether or not the healthcare  
3 workers have been previously BCG-vaccinated. Their result <sup>was</sup> showed that the QFT-GIT is  
4 the least costly and most effective strategy if the QFT-GIT sensitivity is greater than that of  
5 the QFT (17). Femia G-Rizkhalil et al. reported the results of the first routine hospital  
6 application of IGRA for the diagnosis of TB infection and they suggested that the QFT-Gold  
7 test was feasible in routine hospital use for the diagnosis of tuberculosis infection (18). It  
8 may be worthwhile that the tuberculosis screening for the newly hired healthcare workers  
9 based on the QFT. Our study is based on their evidence and all the strategies started from  
10 the QFT not the TST on the newly-hired year. We calculated the cost effectiveness of the  
11 subsequent annual testing compared the QFT with the chest X-ray examination.  
12 This study has several limitations. Firstly, the models are not universal in the world at the  
13 point that annual chest X-ray examination is continued as the serial tuberculosis screening  
14 for the healthcare workers who had already treated LTBI or TB once. But we cannot neglect  
15 the point that all have been done with no reasonable scientific evidence in Japan. The  
16 proposed diagnostic X-ray examinations increase the risk of cancer (19). Japanese has  
17 many chances of exposure of X-ray with annual X-ray screening than the people in the  
18 other countries of the world. The potential psychological stress and anxiety with needless  
19 X-ray exposure, as well as the increasing risk of the cancer, also should be reconsidered.  
20 Secondly, our results could not be adapted to immunocompetent adults and children under  
21 5 year-old. Finally, this model wasn't considered about the revision of the QFT results.  
22 One of the impossible reasons is that there is not enough data of natural history of  
23 tuberculosis about the revisions of the QFT (11, 20). The other reason is that we have no  
24 method to take into the revision in Markov model that avoids the reported  
25 chemoprophylaxis. The further improved studies will be needed in order to refine this  
26 In conclusion, annual QFT alone strategy is more effective and less costly than the QFT

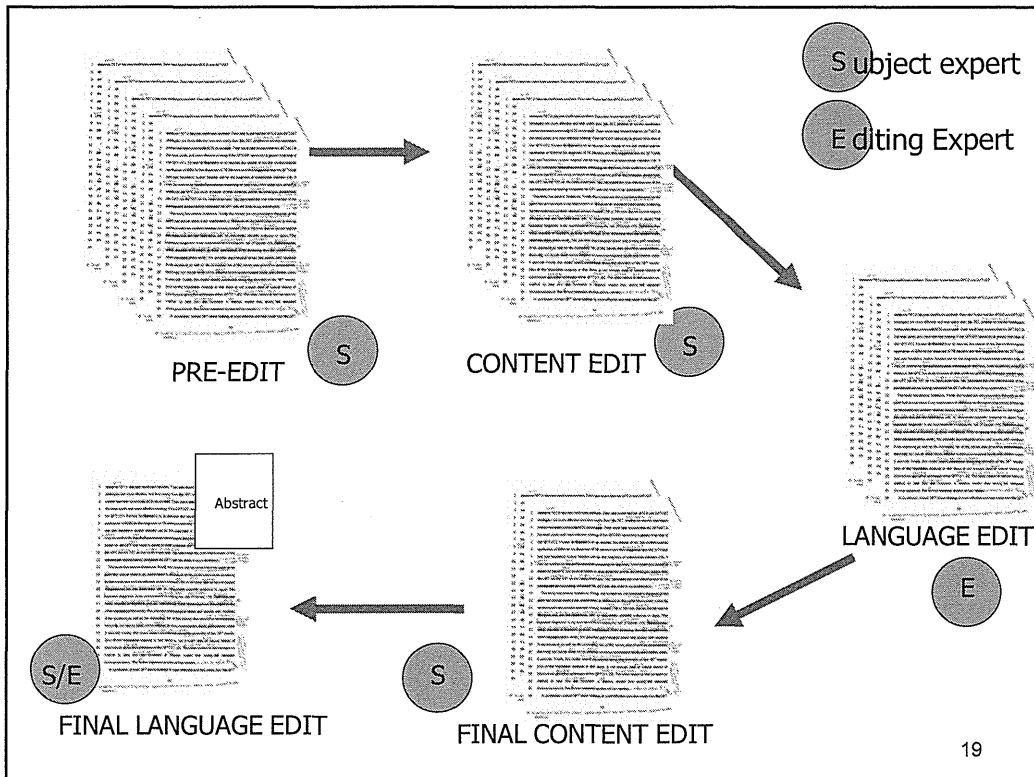
10  
\* In addition, your journal may require you to have more publications acquired in other  
publications.

- Cross-reference your work with the literature

- Do *pre-edit* prior to submitting your work to a professional editing company

- Be clear with your expectations for editors: do you want *comments or corrections*? ("Both" is only feasible in the final drafts!)

18



PRESENTATIONS

## Presentation Preparation: Your Slides

- Words Bad. Pictures Good.
  - *Maximize Figures, Tables, Graphs, and Illustrations*
  - *Minimize words and distracting animations*
  - *Words should be in bullet point format and simple to follow.*
- Complex Pictures Bad. Simple Pictures Good.
  - *Avoid large, scanned graphs/tables/figures*
  - *If you must use them, use highlights or circles to guide viewer*
- Summarize often (visual and oral):
  - *“So far we have seen how previous research has suggested the link between X and Y. We set out to explore this association further in the current study..”*

21

## Presentation Preparation: You!

- Look at the audience when you present.
  - *Don't just read the slide text! Scan the audience.*
- Use body language to accentuate and refer to “action” on your slides.
  - *When you keep moving, your audience are more likely to focus on your goals (and stay awake).*
  - *Use your pointer calmly and sparingly.*
- Stand straight, speak up. Enunciate!!!
  - *Your audience will expect you to be confident.*
  - *Modify your voice (speed, sound, pacing) to control the flow of the presentation.*

22

## Presentation Preparation: Understand Your Audience

- What is the venue?
  - *International forum or symposium?*
  - *Panel of experts?*
  - *Specialists or Generalists?*
  - *Clinicians?*
  - *Researchers?*

23

## Presentation Preparation: Understand Your Audience

- What is the venue?
  - *International forum or symposium?*
  - *Panel of experts?*
  - *Specialists or Generalists?*
  - *Clinicians?*
  - *Researchers?*
- What is the level of statistical expertise?

24

## Presentation Preparation: Understand Your Audience

- What is the venue?
  - *International forum or symposium?*
  - *Panel of experts?*
  - *Specialists or Generalists?*
  - *Clinicians?*
  - *Researchers?*
- What is the level of statistical expertise?
- What are the goals of the audience?
  - *To learn about a new technique or test?*
  - *To learn about a new topic?*
  - *To learn about a new research methodology?*

25

## Presentation Preparation: Rule of Three

- Practice, practice, practice!
  - At least 3 times FORMALLY
    - Use audio recordings if necessary
  - With 3 appropriate test audience members
  - Ask your test audience for 3 difficult questions
    - Practice your answers!

26

## Presentation Preparation: Ask for Help

- Experienced Colleagues
- Books
- Editing companies
- International colleagues and collaborators  
– *drdeshpande@gmail.com*

27



**MAHALO!**

28

# 「研究マインドを持つ臨床医を育てるプロジェクト」報告会と臨床研究ワークショップ 開催のご案内

## 2015/2/22

本会では、eCRNetでe-learningを修了した方による、研究の進捗報告をします。さらに臨床研究ワークショップとして、Rによるデータ分析法と英語のプレゼンテーションを学びます。

### 概要

開催日：平成27（2015）年2月22日（日）13：00～18：00

会場：聖路加国際大学 3階 301教室（ワークショップでは、310教室も使用）

定員：20名

対象：本プロジェクトにご参加いただいている医師

参加費：無料

※本会は、平成26年度厚生労働省科学研究補助金によって運営しています

### アクセス



学校法人 聖路加国際大学  
3階 301教室

〒104-0044  
東京都中央区明石町10-1

◆地下鉄◆

東京メトロ日比谷線 築地駅

3・4番出口 徒歩3分

東京メトロ有楽町線 新富町駅

6番出口 徒歩3分

### お知らせ

- 平成26年度厚生科研費補助金「研究マインドを持つ臨床医に対する疫学教育プログラムの開発と基盤整備」の一環として実施しています。
- データ分析では、統計分析フリーソフトウェア「R」を使用します。

### お申し込み

- 申込用紙を聖ルカ・ライフサイエンス研究所Eメール (sllsi@luke.ac.jp) へお送り下さい。
- 会議終了後には軽食（無料）もご用意しています。こちら是非ご参加ください。

# P R O G R A M

## 第1部

|               |       |  |
|---------------|-------|--|
| 13:00 - 13:05 | 開会の挨拶 | 高橋 理 (「研究マインド」を持つ臨床医に対する疫学教育プログラムの開発と基盤整備」研究代表者)<br>(聖ルカ・ライフサイエンス研究所臨床疫学センター長) |
| 13:05 - 13:20 | 研究報告① | 「6カ月未満の乳児におけるアトピー性皮膚炎の病勢評価マーカーとしての血清TARC/CCL17の臨床的有用性」<br>小泉 宗光 (愛媛県立中央病院小児科)  |
| 13:25 - 13:40 | 研究報告② | 「仕事・家庭の優先度に対する希望と現実のギャップとパートナーとの関連」<br>茶谷 有紀 (独立行政法人国立病院機構埼玉病院麻酔科)             |
| 13:45 - 14:00 | 研究報告③ | 「初診時めまい問診票の自由記載内容による患者の特性の検討」<br>五島 史行 (独立行政法人国立病院機構東京医療センター耳鼻咽喉科)             |
| 14:00 - 14:10 | 休憩    |  |

## 第2部

|               |            |  |
|---------------|------------|--|
| 14:10 - 15:40 | ワークショップ A1 | Rを利用した医療系データ分析入門<br>林 邦好 (国立大学法人岡山大学大学院環境生命科学研究科助教)<br>高橋 理 (聖ルカ・ライフサイエンス研究所臨床疫学センター長)                                   |
|               | ワークショップ B1 | The ABC's of presentation in English ※レクチャーは英語<br>デシパソングーラム (聖ルカ・ライフサイエンス研究所臨床疫学センター上級研究員)<br>浦山 健一 (東京医科歯科大学医歯学総合研究科講師) |
| 15:40 - 15:50 | 休憩         |  |
| 15:50 - 17:20 | ワークショップ A2 | The ABC's of presentation in English ※レクチャーは英語<br>デシパソングーラム (聖ルカ・ライフサイエンス研究所臨床疫学センター上級研究員)<br>浦山 健一 (東京医科歯科大学医歯学総合研究科講師) |
|               | ワークショップ B2 | Rを利用した医療系データ分析入門<br>林 邦好 (国立大学法人岡山大学大学院環境生命科学研究科助教)<br>高橋 理 (聖ルカ・ライフサイエンス研究所臨床疫学センター長)                                   |
| 17:20 - 18:00 | まとめ (軽食無料) |  |

※プログラムは変更する場合がございます。予めご了承ください。  
A1とA2、B1とB2は同じ内容です。

## お問い合わせ

公益財団法人 聖ルカ・ライフサイエンス研究所 事務局  
〒104-0044 東京都中央区明石町10-1  
電話：03-5550-4101 FAX：03-5550-4114  
URL：https://sllsi.or.jp E-mail：sllsi@luke.ac.jp



聖路加国際大学3階301教室

日時:平成27年2月22日(日)

## Rを利用した医療系データ分析入門

林 邦好

(岡山大学大学院環境生命科学研究科 助教)

高橋 理

(聖ルカ・ライフサイエンス研究所臨床疫学センター センター長)

平成26年度厚生労働省科学研究費補助金事業「研究マインドを持つ臨床医を育てるプロジェクト報告会と臨床研究ワークショップ」

## 本日の内容

1 フリーソフトウェアRについて

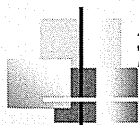
2 Rを利用した実際の統計解析

3 本日のまとめについて

1 フリーソフトウェアRについて

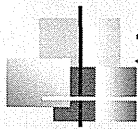
2 Rを利用した実際の統計解析

3 本日のまとめについて



## 統計解析環境「R」について

データ解析あるいは統計解析を行うための商用のソフトウェアには、S-PLUS、STATA、SPSS、SASという代表的なソフトウェアがありますが、これらのソフトウェアはフリーで提供されていないため、誰もが自由にデータ解析や統計解析に取り組むことはできません。Ross Ihaka氏、Robert Gentleman氏により開発され、1996年頃から登場したR言語はオープンソースかつフリーソフトウェアの統計解析向けのプログラミング言語およびその開発実行環境であり、これにより、誰もが自由にデータ解析や統計解析に取り組むことが可能です。R言語のその使用許諾、ライセンスはGeneral Public Licenseのもとで配布されています。このライセンスに従い、Rでは、ソースコードも自由に入手し、閲覧可能であるため、Rの処理結果はその算出過程を誰でも確認することができます。



## オープンソースの特徴

誰もがフリーで自由に同じ環境で統計解析を行うことができるため、同じデータを用いて同じ解析を行っているにも関わらず、(商用の)異なるソフトウェアの違いにより別々の出力結果を得るという事態を回避することができます。現在では、R Development Core Teamによりメンテナンスと拡張が行われています。RCRAN(<http://cran.r-project.org/>)から最新のリリース版(2015年1月15日現在では、2014年10月31日にリリースされたR-3.1.2が最新バージョンです)を入手することができます。さらに、商用のパッケージとは異なり、最近の論文で提唱された分析に関するパッケージも次々にRCRANにおいてアップロードされています。これは、商用のソフトウェアには無い特徴と言えます。

1 フリーソフトウェアRについて

2 Rを利用した実際の統計解析

3 本日のまとめについて

## データの読み込みについて

### 直接格納する方法

例えば、変数「data1」に1、2、3、4、5の5つのデータを格納するためには、以下のようなコマンドを書きます。

```
> data1<-c(1,2,3,4,5) ↩ Enterキーを押してください。  
> |
```

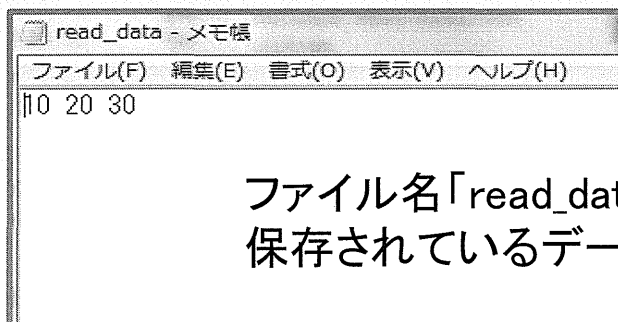
「data1」の中身を確認するために、「data1」と入力し、Enterキーを押してみてください。次のような出力を得ます。

```
> data1  
[1] 1 2 3 4 5
```

## データの読み込みについて

### ファイルからデータを格納する方法

デスクトップ上のファイルからもデータを読み込むことができます。



ファイル名「read\_data.txt」に  
保存されているデータ



## データ「birthwt」について

例えば、「birthwt」というデータフレームがあります。これは、Springfield の Baystate 医療センターの 189 の出生について低体重出生とそのリスク因子の関連を調べるためのデータセットです。

```
> birthwt
```

|    | low                                      | age  | lwt            | race | smoke         | ptl | ht  | ui  | ftv | bwt  |
|----|--|------|----------------|------|---------------|-----|-----|-----|-----|------|
| 85 | ↑ 0                                      | ↑ 19 | ↑ 132          | ↑ 2  | ↑ 0           | ↑ 0 | ↑ 0 | ↑ 1 | ↑ 0 | 2523 |
|    | ↑ 低体重出生の有無を示す 2 値変数(児の出生時体重 2.5kg 未満が 1) |      |                |      |               |     |     |     |     |      |
| 87 | 0  | ↑ 年齢 | 105            | 1    | 1             | 0   | 0   | 0   | 1   | 2557 |
| 88 | 0  | 21   | 108            | 1    | ↑ 非熟練労働経験数    | 0   | 0   | 1   | 2   | 2594 |
| 89 | 0  | 1    | ↑ 最終月経時体重(ポンド) | ↑    | 1             | ↑   | ↑   | ↑   | 0   | 2600 |
|    | ↑ 高血圧の既往(1=あり)                           |      |                |      |               |     |     |     |     |      |
| 91 | 0  | 2    | ↑              | ↑    | ↑             | ↑   | ↑   | ↑   | 0   | 2622 |
|    | ↑ 人種(1=白人, 2=黒人, 3=その他有色人種)              |      |                |      |               |     |     |     |     |      |
| 92 | 0  | 22   | 118            | 1    | 0             | ↑   | ↑   | ↑   | ↑   | 2637 |
|    | ↑ 子宮神経過敏の有無(1=あり)                        |      |                |      |               |     |     |     |     |      |
| 93 | 0  | 17   | 103            | ↑    | ↑ 喫煙の有無(1=あり) | 0   | 0   | 0   | 1   | 2637 |
| 94 | 0  | 29   | 123            | 1    | 1             | 0   | ↑   | ↑   | ↑   | ↑    |
|    | ↑ 妊娠の最初の 3 ヶ月の受診回数                       |      |                |      |               |     |     |     |     |      |
| 95 | 0  | 26   | 113            | 1    | 1             | 0   | 0   | 0   | ↑   | ↑    |
|    | ↑ 児の出生時体重 (g)                            |      |                |      |               |     |     |     |     |      |

## Rの関数を使ってみよう

### 最も基本的な関数について(その①)

| 記号       | 意味  |
|----------|-----|
| sum()    | 総和  |
| mean()   | 平均  |
| median() | 中央値 |
| max()    | 最大値 |
| min()    | 最小値 |
| range()  | 範囲  |

「birthwt」の「年齢」にこれらの関数を適用してみましょう。

```
> sum(birthwt[,2])
[1] 4392
> mean(birthwt[,2])
[1] 23.2381
> median(birthwt[,2])
[1] 23
> max(birthwt[,2])
[1] 45
> min(birthwt[,2])
[1] 14
> range(birthwt[,2])
[1] 14 45
```

Enterキーを押してください。

Enterキーを押してください。

Enterキーを押してください。

Enterキーを押してください。

Enterキーを押してください。

Enterキーを押してください。

## Rの関数を使ってみよう

### 最も基本的な関数について(その②)

| 記号         | 意味      |
|------------|---------|
| var()      | 不偏分散    |
| sd()       | 不偏標準偏差  |
| quantile() | クォンタイル点 |
| summary()  | 要約統計量   |
| scale()    | 標準化     |

「birthwt」の「最終月経時体重」にこれらの関数を適用してみましょう。

```
> var(birthwt[,3])
[1] 935.0985
> sd(birthwt[,3])
[1] 30.57938
> quantile(birthwt[,3])
 0% 25% 50% 75% 100%
 80 110 121 140 250
> summary(birthwt[,3])
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  80.0  110.0  121.0  129.8  140.0  250.0
> scale(birthwt[,3])
      [,1]
[1,] 1.706548157
[2,] 0.823600246
[3,] -0.811488478
[4,] -0.713383154
[5,] -0.746084929
[6,] -0.190154763
```



Enterキーを押してください。

Enterキーを押してください。

Enterキーを押してください。

Enterキーを押してください。

Enterキーを押してください。

## 母平均の差の検定・推定 (対応のない場合)



## Rを用いた2群の母平均の差の検定


対応のない場合：2群が互いに独立である場合

① どちらの群も正規分布に従うときは2標本のt検定を行います。

② 正規分布が仮定できない場合はWilcoxonの順位和検定を行います。

(データ「birthwt」において)2.5kg以上の児を出産した母親の「最終月経時体重」を「group1」に格納し、2.5kg未満の児を出産した母親の「最終月経時体重」を「group2」に格納しましょう。

```
> group1<-birthwt[,3][birthwt[,1]==0]
```

 Enterキーを押してください。

```
> group2<-birthwt[,3][birthwt[,1]==1]
```

 Enterキーを押してください。

① どちらの群も正規分布に従うときの2標本のt検定

<両群の分散が未知であるが、等しい場合>

```
t.test(x, y = NULL, alternative = c("two.sided",  
"less", "greater"), mu = 0, paired = FALSE,  
var.equal = TRUE, conf.level = 0.95, ...)
```

片側検定を行う場合は、こちらを指定します。

両側検定を行う場合は、こちらを指定します。

関数「t.test」を用いることで実行することができます。  
ただし、引数を「var.equal = TRUE」と指定します。

```
> t.test(group1, group2, alternative=c("two.sided"), mu=0, paired=FALSE, var.equal=TRUE, conf.level=0.95)
```



Enterキーを押してください。

```
Two Sample t-test
t値      自由度      p値
data: group1 and group2
t = 2.3537, df = 187, p-value = 0.01962
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.807157 20.521656 ← ----- 上側信頼限界
sample estimates:
mean of x mean of y
133.3000 122.1356
----- 下側信頼限界
group1の平均値      group2の平均値
```

① どちらの群も正規分布に従うときの2標本のt検定

<両群の分散が未知であり、等しいかどうか分からない場合>

```
t.test(x, y = NULL, alternative = c("two.sided",  
"less", "greater"), mu = 0, paired = FALSE,  
var.equal = FALSE, conf.level = 0.95, ...)
```

関数「t.test」を用いることで実行することができます。  
ただし、引数を「var.equal = FALSE」と指定します。

```
> t.test(group1, group2, alternative=c("two.sided"), mu=0, paired=FALSE, var.equal=FALSE, conf.level=0.95)
```



Enterキーを押してください。

Welch Two Sample t-test

| data | 値       | group1 and group2 | 自由度    | p値                |
|------|---------|-------------------|--------|-------------------|
| t =  | 2.5155, | df =              | 132.46 | p-value = 0.01308 |

alternative hypothesis: true difference in means is not equal to 0  
95 percent confidence interval:  
2.38552 19.94329 ← ----- 上側信頼限界

sample estimates:  
mean of x mean of y  
133.3000 122.1356

group1の平均値      group2の平均値

----- 下側信頼限界

(データ「birthwt」において)2.5kg以上の児を出産した母親の「年齢」を「group1.wil」に格納し、2.5kg未満の児を出産した母親の「年齢」を「group2.wil」に格納しましょう。

```
> group1.wil<-birthwt[,2][birthwt[,1]==0]
```



Enterキーを押してください。

```
> group2.wil<-birthwt[,2][birthwt[,1]==1]
```



Enterキーを押してください。

## ② 正規分布が仮定できない場合のWilcoxonの順位和検定

関数「wilcox.test」を用いることで実行することができます。

```
wilcox.test(x, y = NULL, alternative =  
c("two.sided", "less", "greater"), mu = 0,  
paired = FALSE, exact = NULL, correct = TRUE  
conf.int = FALSE conf.level = 0.95, ...)
```

信頼区間を計算するか否かを  
指定します。

連続修正を行うか否かを指定  
します。



引数についてさらに詳しく学びたい方は、`help(wilcox.test)`を実行すると参照いただけます。