Analyze Group-randomized Trial Data with Binary Outcome
Using Weighted Paired t-test: A SAS Macro

Don(Dongguang) Li, NCIC-CTG at Queen’s University, Kingston, ON, Canada

ABSTRACT
Group (previously also called 'cluster', will use 'group' in this article) - randomized design is an important method widely used in clinical trials and trials of public health research. Design with binary outcome is a specific type of group-randomized trials. Using the weighted paired t-test approach raised by Donner and Donald (1987), a SAS® macro-based program for a flexible data analysis is developed. The program accommodates the options such as number of groups, actual condition of intra-group correlation, data transformation method and test adjusting when number of groups is small. The algorithm of the method is introduced. An example with unpublished data is provided. This SAS macro has been validated and is very easy to use.

KEY WORDS
Group-randomized trial, Binary Outcome, Weighted paired t-test, SAS Macro

INTRODUCTION
There are many public health and clinical trial studies using group-randomized design. This design is featured with that the study objects in the trial are identifiable groups rather than individuals. The groups are allocated randomly to the trial treatment categories and the study outcomes are observed on the members of the groups. The groups are not formed at random, but rather based on some physical, social or geographical connections of the members (such as communities, schools, and clinics). In analyzing the data with group-randomized design, the variances within and between the groups need to be counted. There are different statistical analysis strategies for data with continuous outcomes and binary outcome. While the mixed-model variance analysis is the major solution for analyzing data with continuous outcomes, the weighted paired t-test is a convenient and valid analysis method for data with binary outcome in a pre-stratified design.

The weighted paired t-test for group-randomized trial data with binary outcome was introduced by Allan Donner in 1987. Derived from the idea of testing groups means of treatments with paired t-test, this approach uses group proportions (events rates) as basic observations, taking into account of intra-group correlation, measuring the intra- and inter-group variations, and producing weight scores accordingly. The t statistic is generated using the weight scores in each sub-group and summarizing them together. The calculation equations are as follows:

Calculate the rates of treatment and control
Let \( n_{i1} \) and \( n_{i2} \) denote the size of the groups randomly assigned to treatment (1) and control (2), and \( i \) is the \( i \)th stratum in the \( k \) strata \((i = 1, 2, \ldots, k)\); we have \( \hat{p}_{i1} = a_{i1} / n_{i1} \) and \( \hat{p}_{i2} = a_{i2} / n_{i2} \) estimating the events rates of treatment and control groups, respectively, in the \( i \)th stratum.

We want to test the hypothesis \( H_0: \text{Odds Ratio} = p_{i1}(1 - p_{i2}) / p_{i2}(1 - p_{i1}) = 1 \) for the stratum. The overall testing needs the weight scores according to the variance of each stratum.

Calculate the intra-group correlation
\[ MSC = \sum \sum n_{ij} (\hat{p}_{ij} - \bar{p}_i)^2 / k \]
and
\[ MSE = \sum \sum a_{ij} (1 - \hat{p}_{ij}) / (N - 2k) \]

and
\[ n_A = [N - \sum \sum n_{ij} / n_{iT}] / k \]
\[ N = \sum n_{iT} \]

then
\[ \hat{\rho} = \frac{MSC - MSE}{MSC + (n_A - 1)MSE} \]

if MSC < MSE then \( \rho = 0 \)

**Estimate the variance of each stratum**

The variance of \( i \)th stratum can be approximated as below, if \( n_{ij} \) and \( n_{i2} \) is reasonably large:

\[ \text{var}(d_i) = \bar{p}_i (1 - \bar{p}_i) \left[ \frac{1 + (n_{i1} - 1)\hat{\rho}}{n_{i1}} + \frac{1 + (n_{i2} - 1)\hat{\rho}}{n_{i2}} \right] \]

**Calculate the weight scores and the weighted paired \( t \) statistic**

\[ w_{ip} = \text{var}(d_i)^{-1} = \frac{n_{i1}n_{i2}}{n_{i1} + n_{i2} + \hat{\rho}[2n_{i1}n_{i2} - n_{i1} - n_{i2}]} p_i (1 - p_i) \]

Let \( \bar{d}_{wp} = \sum w_{ip} d_i / \sum w_{ip} \) and \( S_d^2 = \sum w_{ip} (d_{ip} - \bar{d}_{wp})^2 / \sum w_{ip}^2 \), the overall \( t \) is:

\[ t_{wp} = \frac{\bar{d}_{wp} \sum w_{ip} / S_d}{\sqrt{\sum w_{ip}^2}} \]. The \( t \) value follows \( t \)-distribution with \( (k-1) \) degrees of freedom.

**Correct the \( t \) statistic if \( k \) is small (defined as <10)**

When \( \bar{d}_{wp} \geq 0 \), replace it with \( \bar{d}_{wp} = 0.5 / N_1 - 0.5 / N_2 \) (\( N_1 \) and \( N_2 \) are total observations in treatment and control groups, respectively).

**Empirical logistic transformation**

If there is considerable variation in the events rates across strata, an analogues procedure based on the empirical logistic transformation is applied:

\[ l_{ij} = \log \left\{ \frac{a_{ij} + 1/2}{(n_{ij} - a_{ij} + 1/2)} \right\} \]

\[ \text{var}(d_{il}) = \frac{(n_{i1} + 1)(n_{i2} + 1)[1 + (n_{i2} - 1)\hat{\rho}]}{n_{i1}(a_i + 1)(n_i - a_i + 1)} + \frac{(n_{i2} + 1)(n_{i2} + 2)[1 + (n_{i2} - 1)\hat{\rho}]}{n_{i2}(a_{i2} + 1)(n_{i2} - a_{i2} + 1)} \]

Let \( \bar{d}_{wi} = \sum w_{il} d_{il} / \sum w_{il} \) and \( S_i^2 = \sum w_{il} (d_{il} - \bar{d}_{wi})^2 / \sum w_{il}^2 \), where \( w_{il} = 1 / \text{var}(d_{il}) \)

the overall \( t \) then becomes: \[ t_{wp} = \frac{\bar{d}_{wp} \sum w_{ip} / S_d}{\sqrt{\sum w_{ip}^2}} \].

From the equations we can see that the algorithms of the calculations are pretty complicated, and the procedure is determined with many conditions. The intra-group correlation depends on the values of MSC and MSN; small number of groups needs a correction of continuity; the empirical logistic data transforming is an option based on the judgement of the rates variance across strata. All these factors need to be accommodated in calculating the \( t \)-statistic.
There is no statistical software including this analysis method and ad-hoc calculation is time-consuming and easily getting errors. This SAS macro is purposed to provide an easy tool for analyzing the data from this specific sort of design.

PROGRAMMING STRATEGIES

**Macro-driven, simple I/O**
The process of the analysis is determined by some parameters to control the equation selection and data management. The input data is in the delimited text file format and can be read in directly if the data name and path are given as a macro parameter.

**Be able to read data in different formats**
The raw data can be in the format of original list per ID or of summarized frequencies per group. A macro parameter is used to make the data reading optional according its style.

**Determine the number of groups**
Although the number of randomized groups is available in design, I just simply let the program getting the group number automatically rather than hard-entering it.

**Calculate and compare MSC and MSE to determining $\rho$**
Because the value of $\rho$ depends on the values of MSC and MSE, the program calculates the two vectors and compares them internally to give a correct value of $\rho$.

**Automate the correction of continuity based on $k$**
The process of correction of continuity will be conditionally triggered when the randomized groups are less than 10.

**Make empirical logistic transformation optional**
The data logistic transformation is switched on/off with another macro parameter.

**Keep transitional value as macro variables**
The calculation (mostly adding up a summary) in the data step yields many transitional values and these values are stored in macro variables.

**Custom report**
A user-custom report is created using data _null_ approach.

AN EXAMPLE BASED ON UNPUBLISHED DATA
Suppose we have a data from a study comparing two methods of protecting children’s eyesight (the rates of worsened cases based on a pre-defined criteria between the two treatments are compared). Stratified with school and grade, we have selected 6 classes for each treatment group. The selected classes are our randomization units and are allocated randomly into the two treatments matching the stratification factors.

The data has already summarized the frequencies. It is a tab-delimited data. We call the macro to find the data and do the analysis with:

```sas
%donner(dat=wtfreqs.txt, dlm='09'x, logtrans=y, freqdata=y);
```

Here logistic data transformation is applied. The result is in a click away as:
PROGRAM VALIDATION
I used an Excel spreadsheet to manually confirm the basic calculations step by step. The results of the parameters as well as the t statistic are validated to be correct.

DISCUSSION
Group-randomized trial is simple and convenient. The advantages of the design include that the study cost may be reduced and that the study is easier to conduct. In some cases, it may eliminate the ethical problems which might otherwise arise. However, it is often that the probability of imbalance between treatment groups on important subject characteristics increases significantly. To offset this problem, pre-stratification of groups with respect to important background factors is commonly applied. Group summary measures, such as means or proportions as the basic observations, are usually used to test the effect of the treatment.

For the studies with outcome as continuous number, the analysis methods are mainly based on mixed modelling. For the studies with binary outcome, the proportions (event rates) of the randomized groups can be used as continuous values, given the number of the groups are reasonably large. A weight based on the proportion of each group is calculated to offset the imbalance. However, it is difficult and costly to have
large number of groups, and randomization per se sometimes cannot completely avoid the bias caused by the characteristics of the groups. Pre-stratification (randomizing the groups in matched pair) is purposed to control some of the major sources of bias. Weighted paired t-test is the method designed for analyzing the data from a stratified group-randomization trial.

Adapted from the paired t-test for testing the difference of means, Dr. Donner’s algorithm is an analogue for testing the difference of event rates between treatment groups. The weight of each stratum is the reverse of the variance of the rates difference. This design requires relatively smaller number of strata and is especially suitable for group-randomization designs, which match small number of paired groups of relatively large aggregate units.

When the number of strata is small, empirical logistic transformation is recommended even if there is no obvious variance of event rates across strata. As we see in the example, the variances of the rates of worsened eyesight in the unit groups are not remarkable, however the empirical logistic transformation of the rates changed the statistical conclusion.

Apparently other methods can be used to calculate the weighted paired t statistic, this macro program can however provide a simple, robust and flexible tool to do the analysis. If anyone may have a chance to do a pre-stratified group-randomized trial with event rate as study outcome, this program can help you to get quick and reliable result.

REFERENCES

APPENDIX: PROGRAM CODE
Available upon request.

ACKNOWLEDGMENT
The author thanks Dr. Joseph Pater for his input and his spreadsheet data for validation.

CONTACT INFORMATION
Don(Dongguang) Li, PhD, MPH
Senior Biostatistician
National Cancer Institute of Canada – Clinical Trials Group
10 Stuart Street, Queen’s University
Kingston, ON, Canada, K7L 3N6
613 533 6000 Ext. 78337 dli@ctg.queensu.ca