

Analysis of Longitudinal Data in Presence of Informative Observation Process with Application to Pediatric Cancer Study

Shesh N Rai
and
S Phipps

St. Jude Children's Research Hospital

Conference on QOL Research in Asia
Hong Kong Society for Quality of Life
May 20, 2006

Outline

- Motivating Example
 - Distress in Bone-Marrow Transplant Patients- background, design, measurement
 - Problems
- Usual approach
- Alternatives
 - Transformation
 - Joint Modeling
- Concluding Remarks

Background

- Bone marrow transplant (BMT) is used to treat a variety of malignancies with leukemia being the most common
- Generally requires 6-8 weeks of hospitalization, but longer stays are common due to complications
 - First 6-10 days: conditioning with high dose chemotherapy and possibly total body irradiation
 - Next 2-4 weeks: Donor marrow directly transplanted to patient
 - Next 2 weeks: Transition phase where precautions for infection are decreased.
 - 4-6 months after discharge: Medication to ward off a variety of complications and immune system “boosts”

Background

- BMT is very stressful for children and affects Quality of Life
 - Pain, Fatigue/Malaise, Mouth Sores, Nausea/Vomiting, Anxiety/Depression

- Past studies have been limited
 - Adults
 - Retrospective
 - Small samples
 - Few longitudinal assessment points
 - Lack of appropriate instruments

Quality of Life: BASES

- Behavioral, Affective and Somatic Experiences Scales (BASES)
- Developed specifically for assessing acute and short-term outcomes in children undergoing BMT (Phipps *et al*, Development of the BASES scale, *J Ped Onc Nursing*, 1994)
- Parent, Patient, and Nurse versions

Quality of Life: BASES

- Somatic Distress
 - 11 items, Range: 5 – 55

- Mood Disturbance
 - 14 items, Range: 14 – 70

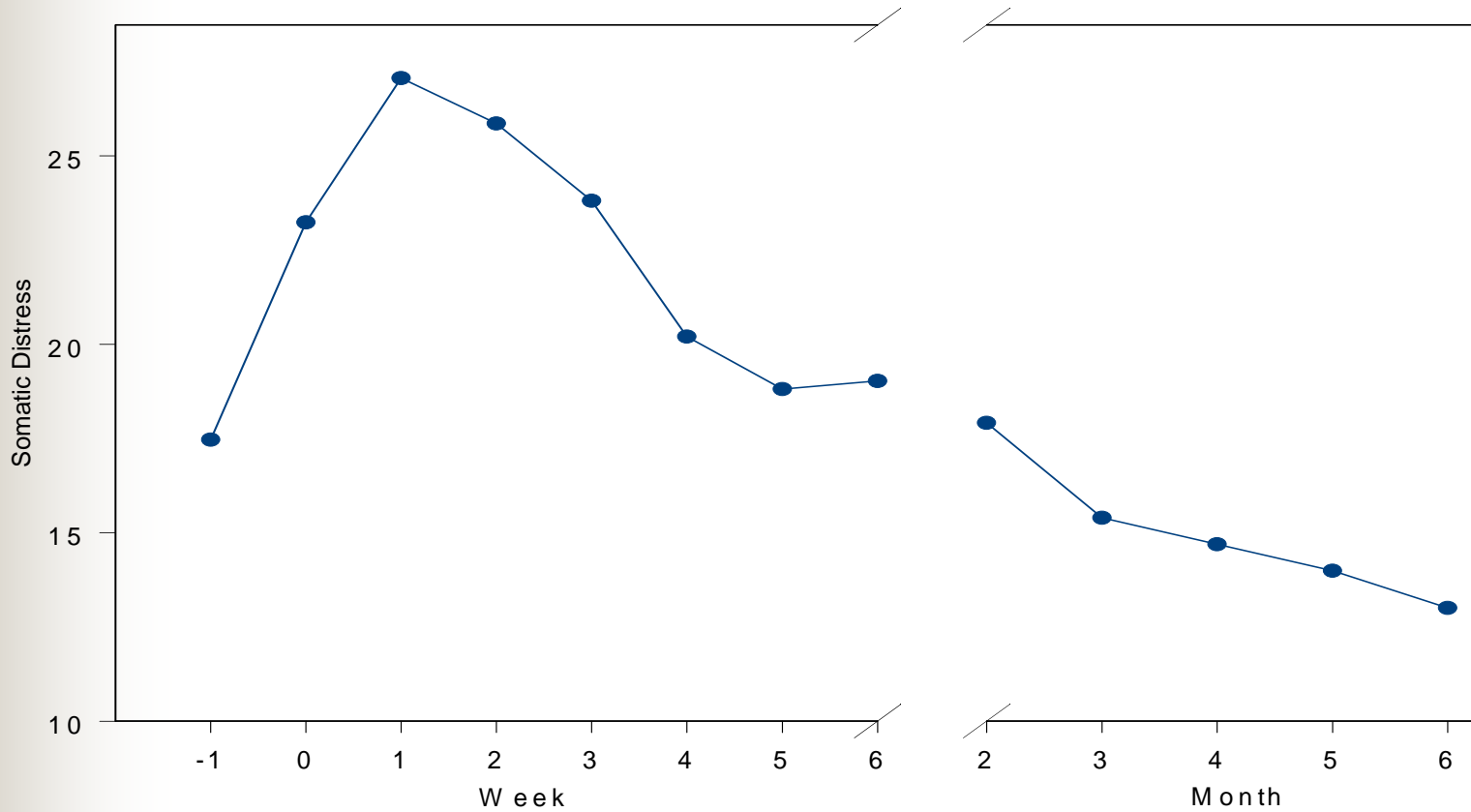
- Compliance
 - 8 items, Range: 8 – 40

- Quality of Interactions
 - 4 items, Range: 4 – 20

- Activity
 - 1 item, Range: 1 - 5

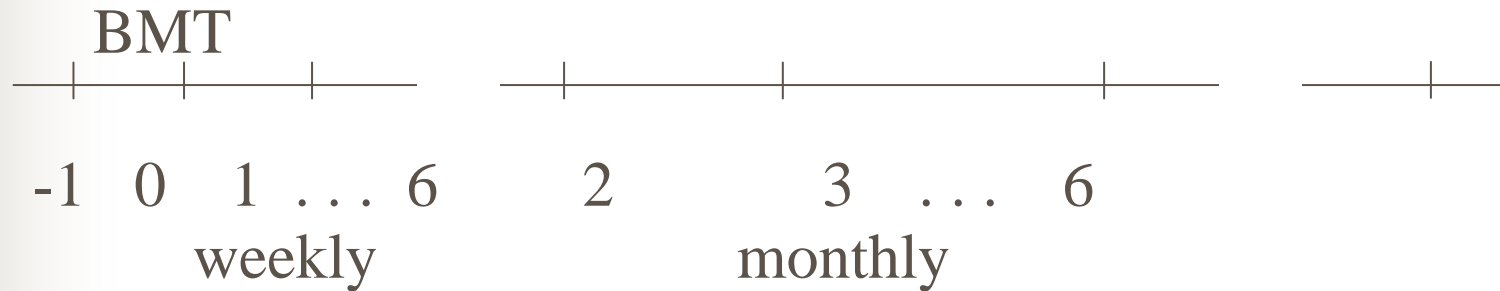
Quality of Life: BASES

Figure 1. Parent report of child's somatic distress



Study Design

- Dependent variable: 13 longitudinal BASES observations



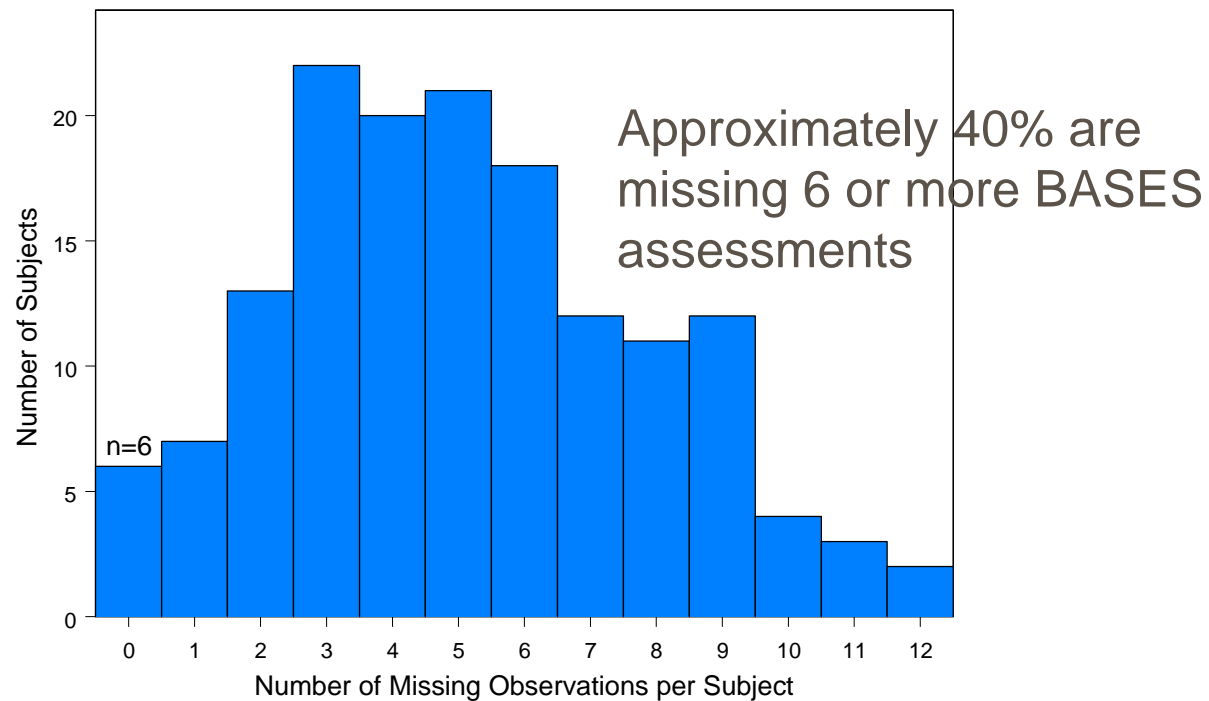
- Independent variables: One-time baseline assessments of
 - Psychosocial status and demographics of patient and parent
 - Treatment and disease related variables
 - Intervention (to improve QOL) related variables



Study Objective

- Determine the predictors of longitudinal BASES measures

Messy Data: Number of Missing BASES Responses per Subject



1194 out of a possible 1963 (60.8%) assessments obtained

Messy Data: Missing at Random?

- Death not a substantial cause for missingness. By week 6, 2 patients had died, and by 6 months 13 patients had died
- No strong evidence overall that missingness not at random:

| # Missing Responses per Subject | Median Somatic Distress Averaged Over Time |
|---------------------------------|--------------------------------------------|
| 0-5 | 19.8 |
| >5 | 20.0 |

- At individual time points (4, 8, and 16 weeks), some indication of differences (all $p=0.03$)

Messy Data: Item Missingness

- How should scales with missing items be scored?
 - Not scored – treat as missing?
 - Impute from other items for that person?
 - Impute from other subjects items?
- For BASES subscales, imputation with average from the other nonmissing items when $<30\%$ were missing

Messy Data: Missingness in the Independent Variables

- For any one independent variables, at most 20% of the responses were missing
- This problem is compounded with multivariable modeling, since data for subjects with any missing data are excluded
- Imputation *could* be used to address this

Messy Data: Fuzziness

- Although BASES observations were *a priori* specified to occur at fixed time points, some didn't
- Only 3.4% off, but for later time points, it was more of an issue
- Examples:
 - *Week 6* (n=84): 2 at wk 5, 7 at wk 8
 - *Month 5* (n=47): 1 at mo 3, 4 at mo 4
- Observations were reassigned to observed time periods. When there were 2 observations, the average was used

Messy Data: Different Scales

- Due to different age groups, different psychometric scales measured the same domain
- For example, for IQ, there were 5 different scales depending on age

When there is a standardized scale, this problem is minimized

Numerous Independent Variables

Demographic/Medical

SES

BMT Type

Diagnosis Group

Premorbid Adjustment

Behavioral Check List –Int

CBCL-External

CBCL-Total

Piers-Harris

IQ

Draw A Person

Prior Illness-Child Distress

PIES-Cooperation/Int

PIES-Parent Distress

Play Performance Scale

Child Personality & Coping

Adaptive Style

CEFT/Embedded Figures Test

Social Support Scale

Family & Parental

FES-Cohesion

FES-Expressiveness

FES-Conflict

FES-Support

FRS-Expressiveness

FRS-Conflict

Ways of coping – confrontive coping

Ways of coping – distancing

WOC – self-controlling

WOC – seeking social support

WOC – accepting responsibility

WOC – escape-avoidance

WOC – playful problem solving

WOC – positive reappraisal

WOC - Total

Inventory for Socially Supportive Behaviors

Simplify!

- Strategy: Simplify the problem
 - Reduced continuous variables to categorical variables with 2 or 3 categories
 - Symmetric – 3 groups
 - Skewed – 2 groups
- Problem: Cut points were primarily based on data. However, we are not really making a big deal about the cut-point values. We are interested in “low”, “medium”, and “high” values
- Benefits:
 - Modeling is easier – don’t have to determine the best functional relationship between all these variables and the response
 - Allows combining of different scales for different age groups (assumes the distribution is somewhat similar between age groups)
 - Results are easier to explain

Modeling Approach

- Simple longitudinal mixed model
 - No covariates
 - Time included as a factor, not as a continuous variable
 - Auto-regressive order 1, AR(1), covariance and random subject effect

- Formulation in SAS:

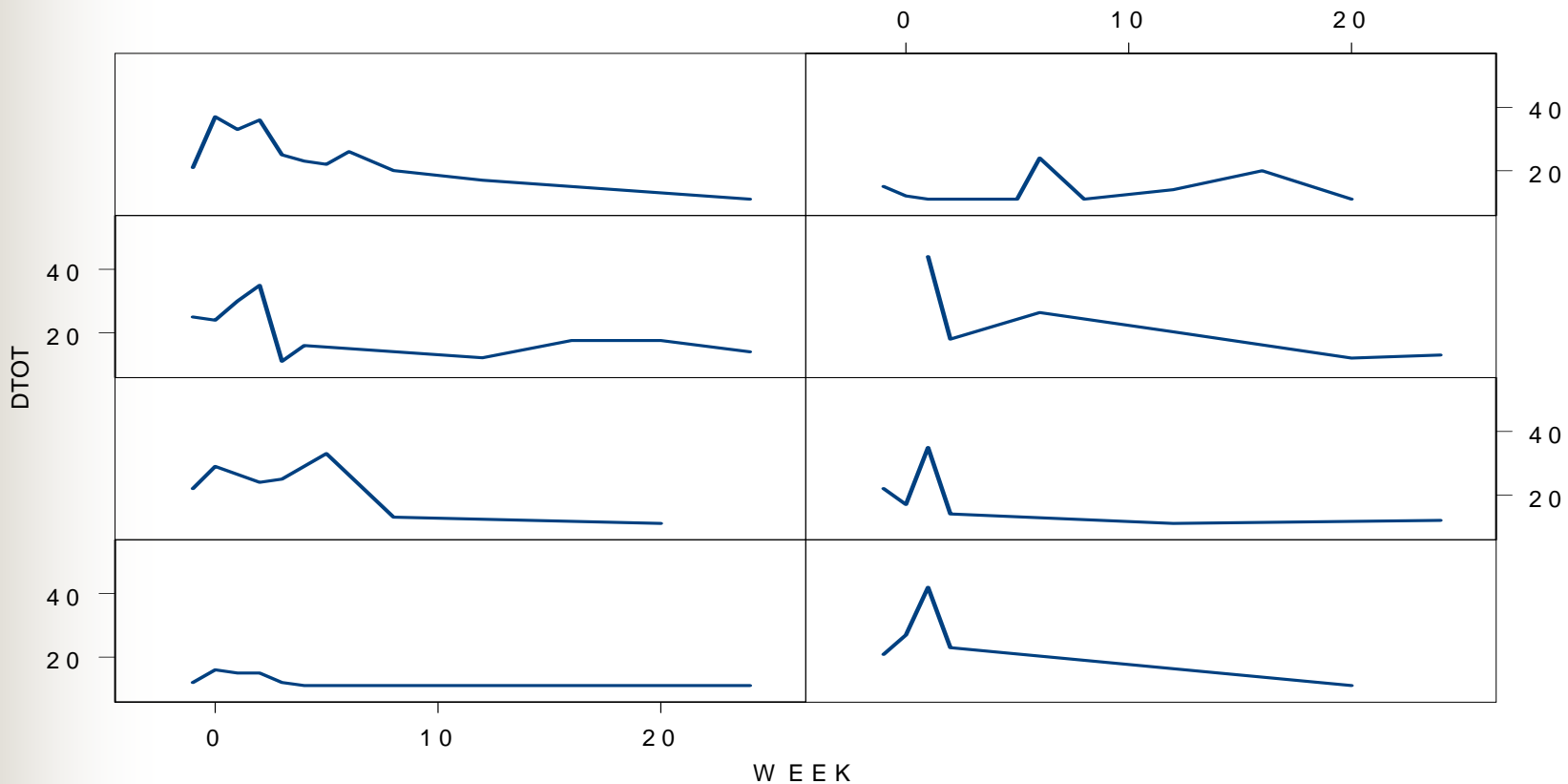
$$y = X\beta + Z\gamma + \varepsilon$$

$$\text{with } E(\gamma) = E(\varepsilon) = 0$$

$$\text{and } \text{Var}(\gamma) = G, \text{Var}(\varepsilon) = R$$

$$\rightarrow \text{Var}(y) = ZGZ' + R$$

Plots for 8 subjects: Somatic Distress vs. Week



Modeling Approach

Mathematical formulation:

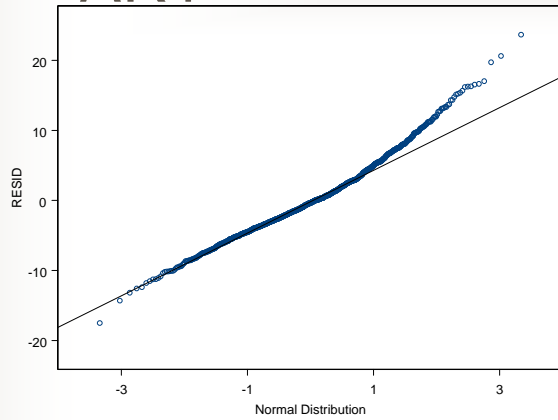
$dtot_{ij} = a + time_i + d_j + e_{ij}$, where $Var(d_j) = \sigma_s I$ and $Var(e_{ij}) = R$

$$R = \sigma \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \dots & \rho^{k-1} \\ \rho & 1 & \rho & \rho^2 & \dots & \rho^{k-2} \\ \dots & & & & & \\ \rho^{k-1} & \dots & \dots & \dots & \dots & 1 \end{pmatrix}$$

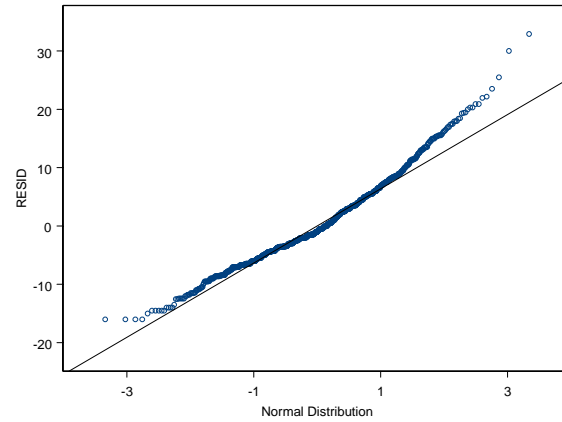
SAS code:

```
PROC MIXED DATA=one covtest noclprint;  
CLASS week ID;  
MODEL DTOT= week;  
random id;  
repeated week/type=ar(1) sub=id; run;
```

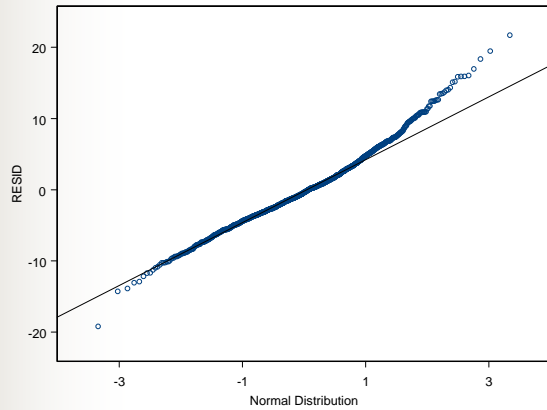
AR1



AR1 & random subject



Random subject



Modeling Approach

To reduce the number of variables under consideration, first investigated “single covariate” models adjusting for age & sex ($P < 0.15$)

- **Demographic/Medical**
 - SES *
 - BMT Type *
 - Diagnosis Group *
- **Premorbid Adjustment**
 - Prior Illness-Child Distress
 - PIES-Parent Distress*
- **Child Personality & Coping**
 - *Adaptive Style* <- *Investigated later*
- **Family & Parental**
 - Ways of coping – distancing
 - WOC – positive reappraisal*
 - WOC - Total
 - Inventory for Socially Supportive Behaviors*

Modeling Approach

- Systematically, added covariates to “simple” model ($P < 0.10$)
- Start with demographic model
- Add most significant variable one at a time
- Then, investigate interactions
- Go to next set of variables: Premorbid Adjustment, etc.
- After final model determined, allow excluded main effects one last time to enter
- Same with interactions with time

N=134

Final Model

Dimensions

Observations Not Used

132

Total Observations

1194

Iteration History

| Iteration | Evaluations | -2 Res Log Like | Criterion |
|-----------|-------------|-----------------|------------|
| 0 | 1 | 6755.57651349 | |
| 1 | 3 | 6531.23830657 | 0.00016756 |
| 2 | 1 | 6530.81776555 | 0.00000337 |
| 3 | 1 | 6530.80980335 | 0.00000000 |

Convergence criteria met.

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate | Standard Error | Z Value | Pr Z |
|----------|---------|----------|----------------|---------|--------|
| ID | | 8.0060 | 1.9604 | 4.08 | <.0001 |
| AR(1) | ID | 0.3103 | 0.04074 | 7.62 | <.0001 |
| Residual | | 29.9562 | 1.7339 | 17.28 | <.0001 |

Fit Statistics

| | |
|--------------------------|--------|
| -2 Res Log Likelihood | 6530.8 |
| AIC (smaller is better) | 6536.8 |
| AICC (smaller is better) | 6536.8 |
| BIC (smaller is better) | 6545.5 |

Final Model

| Effect | DF | DF | Num F Value | Den Pr > F |
|---------------|----|-----|----------------|---------------|
| WEEK | 12 | 706 | 36.78 | <.0001 |
| BMTTYPE1 | 2 | 129 | 14.19 | <.0001 |
| DDISTRES | 2 | 121 | 6.57 | 0.0020 |
| WEEK*BMTTYPE1 | 24 | 748 | 1.59 | 0.0371 |
| NEWSES | 2 | 119 | 2.41 | 0.0944 |

| Variable Description | Values |
|-------------------------------------------------|---------------------------------------------------------|
| WEEK = Week from BMT | -1, 0, 1, 2, 3, 4, 5, 6, 8, 12, 16, 20, 24 |
| BMT TYPE = Bone Marrow Transplant Type | allo-MUD or mismatched allo-matched sibling auto |
| CHILD DISTRESS = Child Distress from PIES scale | Low ≤ 11 Medium $> 11 - \leq 16$ High > 16 |
| SES = Socioeconomic Status | 1,2 3 4,5 |

Child Distress & SES

| Child Distress | No. Subjects | Predicted Avg. (SE) |
|-----------------|-----------------------------|---------------------|
| Low | 48 | 17.5 (0.55) |
| Medium | 51 | 19.9 (0.56) |
| High | 35 | 20.0 (0.68) |
| Comparison | Difference Pred. Avg. (SE) | P |
| Low vs. Medium | 2.5 (0.76) | 0.002 |
| Low vs. High | 2.5 (0.85) | 0.004 |
| Medium vs. High | 0.02 (0.85) | 0.98 |

| SES | No. Subjects | Predicted Avg. (SE) |
|-------------|-----------------------------|---------------------|
| 1,2 | 47 | 18.4 (0.57) |
| 3 | 49 | 18.7 (0.57) |
| 4,5 | 38 | 20.2 (0.64) |
| Comparison | Difference Pred. Avg. (SE) | P |
| 1,2 vs. 3 | 0.3 (0.77) | 0.73 |
| 1,2 vs. 4,5 | 1.7 (0.84) | 0.04 |
| 3 vs. 4,5 | 1.5 (0.83) | 0.08 |

Another St. Jude Study: SMOKN1

An Educational Intervention Targeting Tobacco Use In Pediatric Patients

Alternative Approach

- Problems with these analyses
 - Assumption of normally distributed response - wide ranging ordinal scaled data
 - Often data are skewed
 - Different scales for different age groups
- Missing due to deaths are not random
- Unfortunately, these problems are common

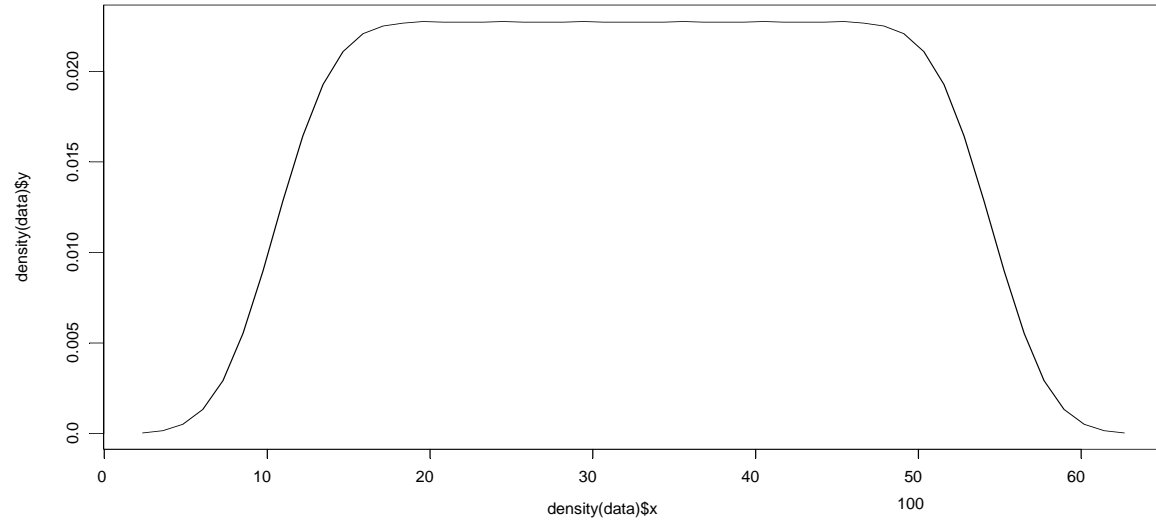
Logit Transformation

- Convert responses to proportions
 1. Subtract theoretical minimum minus 1 from scale
 2. Divide by theoretical range plus 2 (to avoid 0 and 100%) to get p
 3. Transformed response = $\ln (p/(1-p))$

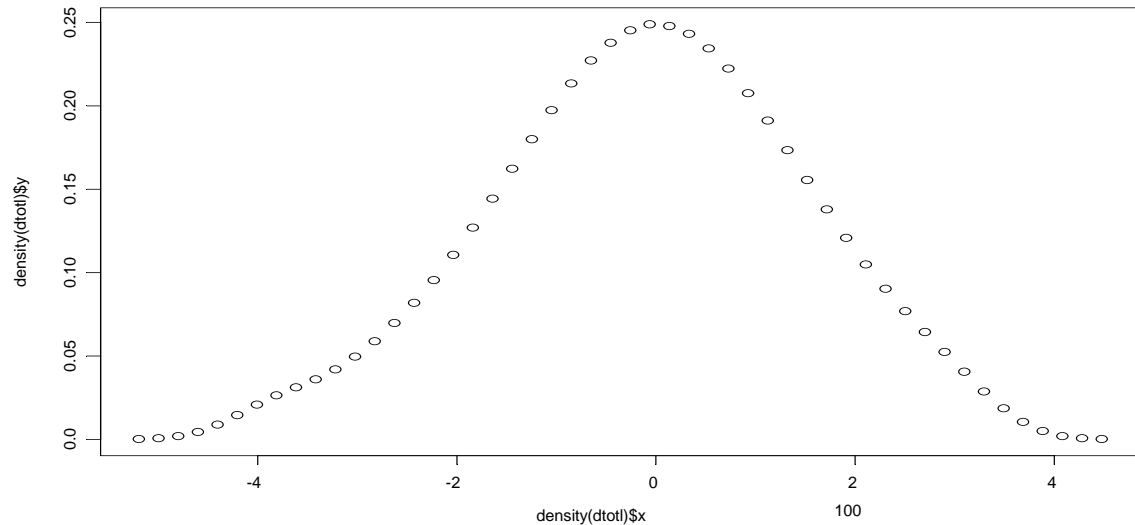
- Example: Response of 30 for scale ranging from 3 to 30
 1. $30-2=28$
 2. $28/29=0.966$
 3. $\ln (0.966/(1-0.966)) = 3.35$

Uniform Data

Density Plot for Untransformed Data - Uniform

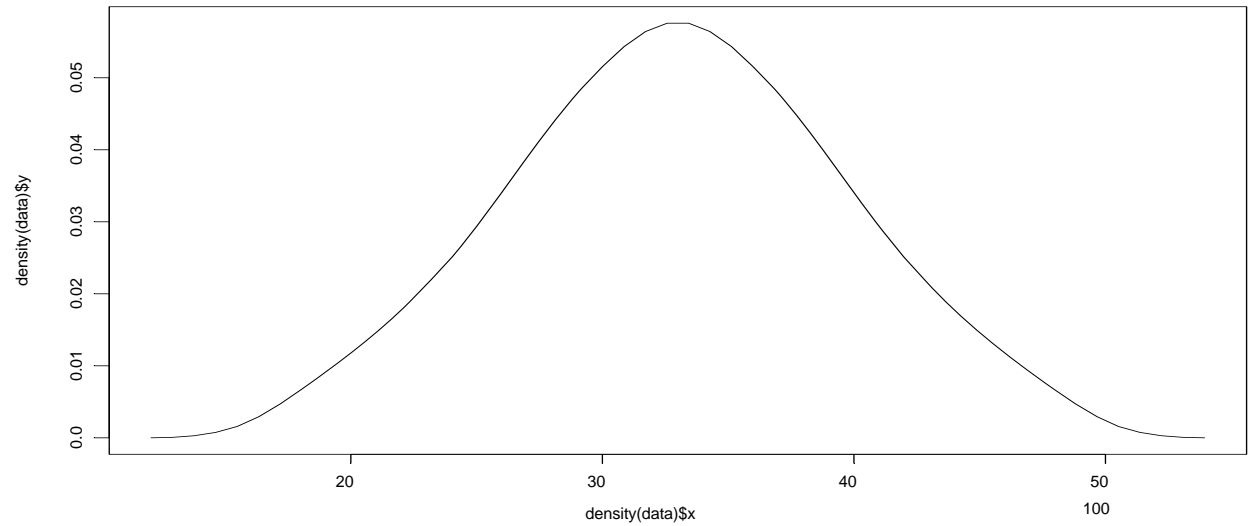


Density Plot for Transformed Data - Uniform

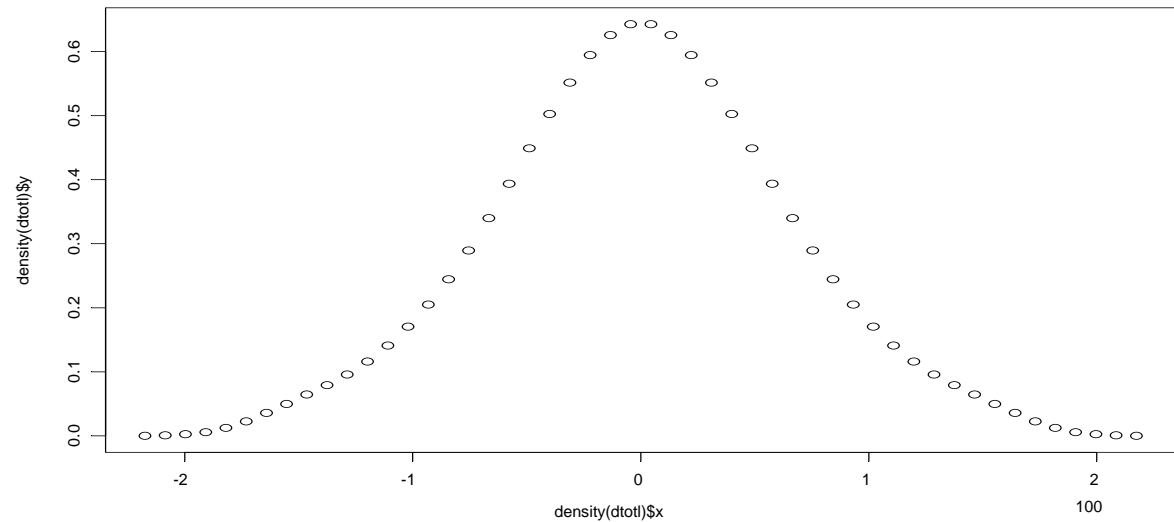


Normal Data

Density Plot for Untransformed Data



Density Plot for Transformed Data



Impact of Transformation on Normal Assumption

Uniform Data

| N | Transformed | Untransformed |
|------|-------------|---------------|
| 100 | 0.5 | 0.5 |
| 200 | 0.5 | 0.0151 |
| 1000 | 0.5 | <0.0001 |

Normal Data

| N | Transformed | Untransformed |
|------|-------------|---------------|
| 100 | 0.5 | 0.5 |
| 200 | 0.5 | 0.5 |
| 1000 | <0.0002 | 0.0498 |

*Kolmogorov-Smirnov test

1. Using mixed model, for all the three time points mean test

Rho=0.8

| | Normal N(35,10 ²)(30,10 ²) | | |
|---------------------|----------------------------------------------------|-------------------------------|-------|
| Sample Size (total) | Untransformed (#P≤0.05/1000) | Transformed (#P≤0.05/1000) | Ratio |
| 50 | 483 | 579 | 20% |
| 100 | 682 | 783 | 15% |
| 200 | 867 | 942 | 9% |
| 500 | 998 | 999 | 0% |

Rho=0.5

| | Normal N(35,10 ²)(30,10 ²) | | |
|---------------------|----------------------------------------------------|-------------------------------|-------|
| Sample Size (total) | Untransformed (#P≤0.05/1000) | Transformed (#P≤0.05/1000) | Ratio |
| 50 | 433 | 551 | 27% |
| 100 | 668 | 797 | 19% |
| 200 | 885 | 966 | 9% |
| 500 | 998 | 1000 | 0% |

Rho=0.2

| | Normal N(35,10 ²)(30,10 ²) | | |
|---------------------|----------------------------------------------------|-------------------------------|-------|
| Sample Size (total) | Untransformed (#P≤0.05/1000) | Transformed (#P≤0.05/1000) | Ratio |
| 50 | 452 | 574 | 27% |
| 100 | 704 | 847 | 20% |
| 200 | 934 | 986 | 6% |
| 500 | 1000 | 1000 | 0% |

2. Results for each time point means comparison when $\rho=0.2$

1st time point

| Sample Size (total) | Uniform (11-52)(14-55) | | | Normal $N(35,10^2)(30,10^2)$ | |
|---------------------|---------------------------------------|-------------------------------------|-------|------------------------------|-------------|
| | Untransformed (# $P \leq 0.05/1000$) | Transformed (# $P \leq 0.05/1000$) | Ratio | Untransformed | Transformed |
| 50 | 183 | 246 | 34% | 353 | 346 |
| 100 | 320 | 435 | 36% | 591 | 561 |
| 200 | 567 | 714 | 26% | 883 | 851 |
| 500 | 933 | 979 | 5% | 1000 | 999 |

2rd time point

| Sample Size (total) | Uniform (11-52)(14-55) | | | Normal $N(35,10^2)(30,10^2)$ | |
|---------------------|---------------------------------------|-------------------------------------|-------|------------------------------|-------------|
| | Untransformed (# $P \leq 0.05/1000$) | Transformed (# $P \leq 0.05/1000$) | Ratio | Untransformed | Transformed |
| 50 | 175 | 228 | 30% | 324 | 322 |
| 100 | 324 | 443 | 37% | 589 | 543 |
| 200 | 575 | 732 | 27% | 859 | 835 |
| 500 | 923 | 986 | 7% | 999 | 998 |

3rd time point

| Sample Size (total) | Uniform (11-52)(14-55) | | | Normal $N(35,10^2)(30,10^2)$ | |
|---------------------|---------------------------------------|-------------------------------------|-------|------------------------------|-------------|
| | Untransformed (# $P \leq 0.05/1000$) | Transformed (# $P \leq 0.05/1000$) | ratio | Untransformed | Transformed |
| 50 | 190 | 239 | 26% | 343 | 332 |
| 100 | 329 | 445 | 35% | 623 | 592 |
| 200 | 561 | 727 | 30% | 879 | 862 |
| 500 | 922 | 985 | 7% | 999 | 997 |



Impact of Transformation BMT study:

Normality assumptions were met for may dependent variables

Valid Results

Change in Scores from Baseline (%)



12-Month Multivariate Comparison of SSC and SEI: $p=0.002$

Without transformation

$P=0.08$

Outcome Measure

Modeling of Survival Outcome

- Longitudinal measurements Y_{ij} at t_{ij}
- Time to Event S_i (death time)
- Baseline covariates x_i
- Parameter θ .
- Model $[Y, S \mid \theta]$

Modeling of Survival Outcome

- $S \sim f(s)$
- $F(s) = P(S \leq s)$
- $h(s) = f(s) / \{1 - F(s)\}$
- Many parametric and non-parametric forms for S
 - $\text{Log } S \sim N(\mu, \sigma^2)$; accelerated life model

Joint Model

- $\{\log(Y, S)\} \sim \text{MVN}(\mu, \Sigma)$
- where $\mu = (\mu_S, \mu_Y)$ and

- $\Sigma = \begin{bmatrix} \sigma^2 & \rho' \\ \rho & V \end{bmatrix}$

Joint Model

- Known results:
- $\log(S|Y) \sim N(\mu=(\mu_{S|Y}, \sigma^2_{S|Y}))$
- where
 - $\mu_{S|Y} = \mu_S + \rho' V^{-1} (Y - \mu_Y)$
- and
 - $\sigma^2_{S|Y} = \sigma^2 - \rho' V^{-1} \rho$

Joint Model

Likelihood Function

$$L = \prod_{i=1}^n L_i$$

where

for the i th uncensored patient

$$L_i = [Y_i] \times [S_i | Y_i]$$

and for the i th censored patient

$$\begin{aligned} L_i &= [Y_i] \times [P [S_i > s_i | Y_i]] \\ &= [Y_i] \times [1 - \phi\{(S_i - \mu_{S_i | Y_i}) / \sigma_{S_i | Y_i}\}] \end{aligned}$$

Joint Model

Estimation: EM type procedure based on pseudo-likelihood

- Step1: Use working covariance matrix
- Step2:
 - E-step: If S_i is censored, replace by its expectation
 - M-step: Estimate the mean parameters
- Step 3: Iterate between step 1 and 2.

Concluding Remarks

- Messy data present many interesting challenges and are common
- Dealing with data issues before analysis requires careful consideration
- Determining the “right” model is important
- Transformation may be needed
- Joint Modeling
- Multivariate Joint Modeling

Very selected references

- Littell RC, Milliken GA, Stroup WW, Wolfinger R. *SAS[®] System for Mixed Models*, Cary, NC: SAS Institute Inc., 1996. 663pp.
- Qian W, Parmar MKB, Sambrook RJ, Fayers PM, Girling DJ, Stephen RJ. Analysis of messy longitudinal data from a randomized clinical trial. *Statist. Med.* 2000;19:2657-74.
- Fairclough DL, Peterson HF, Chang V. Why are missing quality of life data a problem in clinical trials of cancer therapy? *Statist. Med.* 1998;17:667-77.
- Rai, S.N., Kovacevic, M.S. and Srivastava, D.K. Estimation in Hierarchical Linear Models using Pseudo Maximum Likelihood Method *Far East Journal of Theoretical Statistics*, 5(2): 205-221, 2001.
- Xu, J. and Zeger, S.L. Joint Analysis of Longitudinal Data Comprising Repeated Measures and Times to Events. *Applied Statistics*, 50, 375-387.
- Phipps et al, Development of the BASES scale, *J Ped Onc Nursing*, 1994.
- Ribaud HJ et al. A Joint Analysis of Quality of Life and Survival Using a Random effect Selection Model. *Statistics in Medicine*, 19, 3237-3250, 2000
- Tsiatis A and Davidian M. Joint Modeling of Longitudinal and Time to Event Data: An overview. Personal communications.

SLEEP2: Study design and other statistical issues

Shesh N. Rai

Department of Biostatistics

St. Jude Children's Research Hospital



SLEEP2: Primary Aim

To compare sleep efficiency and sleep duration during 3 consecutive 24-hour periods of hospitalization between patients randomized to receive standard care and those randomized to participate in an enhanced physical activity (EPA) intervention.

Study Design and Analyses are driven by this aim

Design

Two-group, prospective, randomized pilot study with longitudinal assessments



Sample Size

EPA
N=30

Standard
Care
N=30



Total=60
Roughly 30
from SJCRH &
30 from TCCC

Note: Randomization will ensure balance:

- chemotherapy or fever/neutropenia
- SJCRH or TCC

Sample size based on detecting a meaningful difference in Sleep Efficiency (wrist actigraph) and allowing for 14% missing data rate

Data Collection

| <u>Measures</u> | Day | | | |
|------------------------|-----|---|---|---|
| | 0 | 1 | 2 | 3 |
| <u>Sleep</u> | | | | |
| Actigraph | | X | X | X |
| Parent Diary | | X | X | X |
| Room Entry Checklist | | X | X | X |
| <u>Fatigue</u> | | | | |
| Patient Report | X | X | X | X |
| Parent Report | X | X | X | X |
| Staff Report | | X | X | X |
| <u>Staff Report</u> | | | | |
| Hemoglobin | | X | X | X |
| Hemotocrit | | X | X | X |
| Transfusion Record | | X | X | X |
| Concurrent Medications | | X | X | X |

- Lots of data to be collected longitudinally → missing data and errors are more likely to arise
- Important to accurately and clearly record data

Data Collection (cont.)



- Missing data can make analysis and interpretation difficult
 - Encourage participation of patients and parents at all 3 assessments
 - Check that forms are completely filled out by study participants
 - Verify that all required evaluations, tests, and observations are collected for each patient



Primary Statistical Analyses

- Compute average of sleep measures across multiple time points
 - Actigraph measures, parent diary
- Compare average sleep measures for the 2 groups using t test or nonparametric test with $\alpha=0.10$
- Same approach will be used to compare fatigue measures between 2 groups (secondary Aim 1.2)

Other Statistical Analyses

- Relate the impact of various covariates on sleep and fatigue (Aim 1.3)
 - Demographics, HgB, HCT, sleep interruptions, baseline fatigue, EPA measures
 - Analysis of covariance and mixed model approaches
- Assess the relationship between sleep efficiency and duration with fatigue (Aim 1.4)
 - Plots of sleep variables vs. fatigue with smoothed line fit to the data
 - Regression techniques