Novel Point estimation from a Semiparametric Ratio Estimator (SPRE): Long-term health outcomes from short-term linear data, with application to adults with developmental disability

D. Weissman-Miller, Sc.D., MS, MPH, CE: Affiliate Professor of Biostatistics; School of Occupational Therapy; Brenau University, Gainesville, GA 30501
Catherine Link MS, OTR/L: Shepherd Center, Neuro Rehab
Celina Parkman MS, OTR/L: Wellstar Kennestone, Neuro Rehab

Abstract

Point estimation is particularly important in predicting improvement in individuals who have developmental disabilities (DD). A new health response function based on a model of human response over time estimates long-term health outcomes from a change point in short-term linear regression in a pilot study for adults with DD. These estimations are given by parameters derived from short-term participant data in an (OLS) regression given in a new semiparametric ratio estimator model (SPRE). The response function is a ratio of two-parameter Weibull distributions times a prior outcome value stepping estimated outcomes forward in time. Shape and scale parameters are estimated at the change point. A feature of the SPRE model is that initial treatment response for each single-subject is reflected in long-term response to treatment. Results show that 7 participants had statistically significant results ($p \leq 0.05$); a means analysis had statistically significant ($p=0.00002$) results showing an increase in skills from session 1 to session 12 (15.4%) and smaller increase for predictions (1.4%). Finally, the response function ratio provides a time frame for improvement for participants.

Key Words: disability, point estimation, health response function, semiparametric ratio estimator, SPRE

1. Introduction

Point estimation is particularly important in predicting response to treatment in individuals or small groups of adults with developmental disability (DD). In this analysis, $G_k,\tau(t)$ is a new health response function based on a model of human response over time to estimate long-term health outcomes from a change point in short-term linear regression. The statistical problem considered here can be stated simply: having initial quasi-linear data from participant's primary or secondary datasets, determine the change point from a backwards stepwise regression model. Then, using a ratio of the function $G_k,\tau(t)$ times the prior estimated outcome, determine new point estimates from the change point to the time where the function ratio becomes 1.00. This indicates participant's/participants' stability or no further change. These point estimations taken from the change point are from a new semiparametric ratio estimator (SPRE), using mean response and single-subject design for clinical medicine, therapy in the health sciences, or pilot studies for phase I clinical trials.

Why are we interested in predicting health outcomes for single-subjects or small groups? The motivation in this paper comes from occupational therapy in health sciences where efficacy and duration of treatment has not been able to be satisfactorily statistically determined. Even single-subject experimental studies are often discarded because it is

thought they cannot yield valid results, according to Thompson (2007). The importance to therapy and clinical medicine cannot be overstated, particularly where the medical model emphasizes the systematic use of information about an individual that is broader than genomics to enable a tailored approach to prevention and care. This problem is particular to clinical medicine and therapy in health science where the small number of 5–10 participants in groups, for example, are too few for the results to be analyzed using classical statistics that depend upon larger numbers for statistical significance.

In many cases, statistical estimations can be made from initial data from a least squares regression, or by predicting a trend line, or in time series forecasting when the use of a model to forecast future events is based on known past events. However, health outcomes are often in the form of an exponential cumulative distribution of a survival function, when residuals are not distributed normally (Cleves, Gould, & Gutierrez, 2004) and for time-to-event-data (Hosmer, Lemeshow, & May, 2008) or a parametric Weibull cumulative distribution where past performance may be unavailable or not relevant to the disease outcomes being measured. Therefore, a new method has been developed to predict long-term outcomes in the form of a Weibull distribution using initial data from the patient(s) in a clinical trial or in therapy. A primary analysis is performed on data from a pilot study using the Alert Program® (Williams & Shellenberger, 1996) administered to adults with developmental disabilities to determine the change points and predicted stability or improvement for participants. The Alert Program® is guided by two theories, the Arousal Theory and Sensory Integration, and uses the analogy of an automobile engine to introduce the self-regulation concepts (Williams & Shellenberger, 1996). The body's "engine" can run *high, low,* or *just right*. The therapeutic goal is to combine sensory integration strategies with cognitive awareness to adjust the engine to run at the *just right* level for the desired task.

An analysis on data from a pilot study by Link, Parkman and Frame (2012) provides a well-documented trial to determine the efficiency and accuracy of Alert Program® response estimations. It uses the new SPRE model for the small group assuming that response data is given for each individual and the mean response of 7 participants at any point in time.

The collected data is for 14 participants treated by occupational therapists, where data from The Assessment of Communication and Interaction Skills (ACIS) measure was noted 2 times per week from the start of the trial to the conclusion at 7 weeks. This analysis computes and compares individual point estimates to the mean data starting from a change-point derived from initial data in the first 7.0 weeks. The Link, Parkman, and Frame (2012) analyses are chosen to illustrate the statistical use of SPRE in disability because prevalence and incidence of adults with developmental disabilities have little statistical analytical work in the literature and yet comprise approximately 1.2% to 1.65% of the U.S. population (Developmental Disabilities Assistance Bill of Rights Act of 2000). Of these people, some have deficits in communication and interaction skills and display atypical sensory processing (Koenig & Rudney, 2010). The Alert Program® is a sensory intervention that may be useful for those with these deficits.

2. Method

2.1 Assumptions for the data using the SPRE model

The first assumption is that approximately 14 data points in the primary trial follows an initial quasi-linear form that can be analyzed by ordinary least squares to develop a change point. This is true for a single subject or a small group where the mean indicates the response to treatment. The second assumption is that the outcome values are ordered

data. If the data are not approximately linear, then transforms are used to linearize the data; if both axes are transformed, then values on the time axis are ordered together with the ordered outcome data. The approximate linearity of the initial data is based on a linear relationship implicit in the solution of a convolution integral on which the predictive health response function is based (Weissman-Miller, 1992). A third assumption is that a change point can be determined from initial data before an arbitrary cutoff in time or treatment numbers. A fourth assumption is that long-term data outcomes can be predicted from the change-point using the ratio of $G_k, \tau(t)$ times the prior outcome prediction.

2.2 Identify the change point in the initial patient least squares data

The equations for ordinary least squares regression (OLS) analyses are given by Berk (2004) and Weisberg (2005) where the weight loss data consists of *n* observations. Each observation includes the scalar outcomes response y_i and a vector of predictors (regressors) x_i. In this linear regression model the response variable is a linear function of the regressors:

$$y_i = x_i'\beta + \varepsilon_i, \tag{1}$$

where β is a $p \times 1$ vector of unknown parameters; ε_i's are unobserved scalar random variables (errors) which account for the discrepancy between the actually observed responses y_i and the "predicted outcomes" $x'_i\beta$; and $'$ denotes matrix transpose so that $x'\beta$ is the dot product between the vectors x and β. There are several different frameworks in which the linear regression model can be cast in order to make the OLS technique applicable. The choice of the applicable framework depends mostly on the nature of data that in this paper is observational. Then the regressors are random variables, and the regressors' x_i is random and sampled together with the y_i's from this adult population. Suppose b is a "candidate" value for the parameter $\boldsymbol{\beta}$. The value of b that minimizes $S(b)$ is called the OLS estimator for $\boldsymbol{\beta}$ in (1) where:

$$S(b) = \sum_{i=1}^{n} (y_i - x_i'b)^2 \tag{2}$$

Generally speaking, change-point regression is a regression problem in which the expected value of the dependent variable or response is assumed to have a different functional form in several neighborhoods of the explanatory variable space according to Khodadadi and Asgharian (2008). In the SPRE model, the determination of the change point is a structural change, that for weight loss shows the dynamic nature of the changes. A novel estimation procedure using a backwards-stepwise regression comparatively determines the highest or lowest F statistic for each regression model reduced by reasonably even increments of time or the even number of treatment sessions. The time at the change point is the location of the initial $G_k, \tau(t)$ ratio predictions for long-term estimates, which produces a gradually changing predictive model.

The region of analytical interest for the backward stepwise elimination method is 14 sessions of treatment using a total of 7 weeks of initial pilot study data.

The highest or lowest value of the F statistic in this paper is derived by equations relying on x and y and for this derivation follow the relation given by Wu (2005) for $F = t_1^2$. The data given every session weekly by Link et al. is analyzed for the first 14 data points that are comparable to the modality of weight loss treatment analyses by

Weissman-Miller, Shotwell, and Miller (2012). The derivation for the least squares estimate of OLS regression is given as:

And: $$s_e(\hat{\beta}_1) = \sqrt{\frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2} = \sqrt{\frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

When: $\hat{\beta}_1 \sim N(\beta_1, \sigma^2_{\hat{\beta}_1})$; $\hat{\beta}_0 \sim N(\beta_0, \sigma^2_{\hat{\beta}_0})$;

Then: $$\hat{\beta}_1 = \frac{SXY}{SXX} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} \tag{3}$$

And: $$t = \frac{\hat{\beta}_1}{s_e(\hat{\beta}_1)} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y}) / \Sigma(x_i - \bar{x})^2}{\sqrt{\frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 / \sum_{i=1}^{n}(x_i - \bar{x})^2}} \tag{4}$$

$$RSS = SYY - \frac{(SXY)^2}{SXX} = \sum_{i=1}^{n}(y_i - \bar{y})^2 - \frac{[\Sigma(x_i - \bar{x})(y_i - \bar{y})]^2}{\Sigma(x_i - \bar{x})^2} \tag{5}$$

After algebraic manipulation:, the highest or lowest F statistic is determined by:

$$F = \frac{(SYY - RSS)/1}{\hat{\sigma}^2} = t^2 = \frac{\dfrac{[\Sigma(x_i - \bar{x})(y_i - \bar{y})]^2}{\Sigma(x_i - \bar{x})^2}}{\dfrac{1}{n-2}\Sigma_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2} = \frac{SS\,\mathrm{Re}\,g/1}{\hat{\sigma}^2} = \frac{MS\,\mathrm{Re}\,g}{\hat{\sigma}^2} \tag{6}$$

The results are given in Table 1 for the weight loss treatment with phentermine. These results were computed by the R Development Core Team (2010), R Commander by Fox (2005) and using R Commander through Excel from Heiberger and Neuwirth (2009) also analyzed in this paper by a program written in the R Console. In R Commander, a linear model was used to include the ANOVA results, particularly the F statistic. It should be noted that in this SPRE model, the F statistic is determined comparatively from the regression analyses, where the P value is used to determine statistical significance of the regression analysis to that point. The resulting change point, of the highest or lowest F statistic computed, is then denoted by the value of time, or treatment session number, together with the measured outcome in the dataset.

Determining the highest or lowest F statistic using a nested equations model according to Berk (2004) and Faraway (2005) produces F statistics derived from evaluation of the y relationships alone. In this case, the value of the F statistic continuously decreases from low to high values of say 14 sessions or 7 weeks for this dataset. Another approach taken from Wu (2005), where the highest or lowest F statistic is derived by pooled estimates that depend only upon x, the value of the F statistic continuously increases from low to high values of 7 weeks.

Table (1) Determination of Highest or Lowest F Statistic for a Significant Mean Of 7 Participants in ALERT Program Therapy for DD

Session	R^2	F-statistic	P-value	
2	N/A	N/A	N/A	
3	0.7082	2.427	0.3633	
4	1	NaN	N/A	
5	0.8213	9.19	0.0938	
6	0.7176	10.16	0.0333	
7	0.7567	18.66	0.005	
8	0.768	26.49	0.0009	
9	0.7708	16.82	0.0093	
10	0.791	26.49	0.0013	
11	0.79	11.29	0.0438	
12	0.8294	53.48	0.00002	Change Point
13	0.8286	48.33	0.00004	
14	0.8137	39.3	0.0001	

As shown in Table (1), there is only one highest F statistic in this dataset. In any therapy, the treatment should be carried out past a first mode to be sure that all relevant modes, such as a second mode of response, have been recorded. There is only 1 highest or lowest mode for the mean of 7 participants in the ALERT program.

It can be seen from the analyses in Table (1) that the highest F statistic is at time = 12 sessions from the backward stepwise elimination method, including the variables x (time) and y (outcome), as given in equation (6). Each dataset, starting at an assumed full model at 14 sessions or 7 weeks, is analyzed to determine R^2 and the F statistic. Then the next dataset is analyzed from 13 sessions down to 3 sessions.

In this paper, a change point is defined as the outcome and the time = $\tau(t_i)$ for the highest F statistic, located in this dataset initially at 12.0 weeks, for analysis of the mean, where the participant outcomes with respect to time are no longer linear (Weissman-Berman, & Martin, 2008; Weissman-Miller, 2010; Weissman-Miller & Miller (2011). In medical models of response to treatment, the change point is the data point at which the character of the regression changes, in general, from linear to the shape of an exponential or the Weibull cumulative distribution (Weissman-Miller, 2010). Another example of this character of the regression is shown in a prospective pilot program on the reduction of childhood problems of sensory integration (SI), (Barth, Brooks, & Carroll, 2011) and in a study of weight loss (Weissman-Miller, Shotwell, & Miller, 2012).

2.3 Derivation of the predictive parameter $G_k, \tau(t)$ ratio

The time at the change point determines the location of the initial ratio predictions for long-term estimates from the least squares set of random pairs $\{(Y_t, X_t)\}_{t\in T}$.

The predictive parameter is derived from a Kelvin model in continuum mechanics (Weissman-Berman, 1992) chosen to model human response to disease as given by Weissman-Miller (2010). The Kelvin model is shown in Aklonis and MacKnight (1983) and the generalized Kelvin-Voight model as derived in Mills (1993). The original Kelvin model predicts the relaxation, or change point, of complex materials and can be used to

model the human body's bones and sinews together with the viscous properties such as blood and water). The transformed model is given here as a cumulative Weibull distribution rather than an exponential distribution with the superscript 'k' assumed to be 1.00. This new predictive parameter is given as $G_k,\tau(t)$, the health response function, to determine the shape of the predictive curve, where G_k is the response function and $\tau(t)$, the compliance parameter, that are measures of time:

$$G_k,\tau(t)=\left(1-e^{-\left(\frac{t}{\tau}\right)^k}\right) \tag{7}$$

This equation is used in the SPRE model specifically because it includes $G_k,\tau(t)$ as a function of time and a prior in the form of τ that has been derived initially as the change point in the SPRE model. Finally, this equation is widely used in survival analysis together with the cumulative exponential distribution of a survival function when residuals are not distributed normally (Cleves, Gould, & Gutierrez, 2004) and for time-to-event-data (Hosmer, Lemeshow & May, 2008).

To derive $G_k,\tau(t)$ as a ratio, an assumption is made that a statistical function $\Psi(t)$ relates θ_t *and* $\tau(t)$, outcomes and a function of time, on the initial ordinary least squares data regression line and is a bivariate function. These equations, given here as $\tau(t)=\tau\cdot G_k,\tau(t)$ and $\theta_t=\theta_\tau\cdot G_k,\tau(t)$, are evaluated on $G_k,\tau(t)$ for θ_t which is a population parameter at the change point in the ordinary least squares regression. Here, the parameter τ is derived as the value of time at the change point. The cumulative Weibull distribution in equation (7) is given as a ratio in Weissman-Berman (2009), Miller, Weissman-Berman & Martin (2009), Weissman-Miller (2010) and Weissman-Miller, Shotwell, and Miller (2012). The outcomes beyond τ are point estimates and given as $\hat{\theta}_t$.

At time greater than τ when t_{i+1} is the increased value of time for the point estimates, t_i is the prior time and the ratio is multiplied by the prior outcome θ_{t_i} to determine the new point estimate:

$$\hat{\theta}_t = R\cdot\theta_{t_i} \quad \text{and expanded becomes} \quad \hat{\theta}_t=\frac{G_k,\tau(t_{i+1})}{G_k,\tau(t_i)}\cdot\theta_{t_i} \tag{8}$$

In general, $G_k,\tau(t_{i\pm1})$ depends upon the value of the ordered outcomes. While the denominator of equation 8 on the right hand side is initially given as $\tau(t_i)$, time at the change point, when the participant's data varies widely from session to session, may be expanded to $\tau(t_{i,j})$ for point estimation where the generic $N_{i,j}$ is the number of the jth session in the ith configuration of the test data. In any event, not less than a second highest inflection point in the original test data is defined as $\tau(t_{i,j})$ in the denominator of equation (8). This procedure has been followed for some of the participants in this study of adults with DD (Link, Parkman, & Frame, 2012) to be more inclusive of the participant's response to treatment.

In the Weibull distribution, $\hat{k}$ is the predicted shape parameter and $\tau(t_i)$ is the scale parameter, both derived from the initial least squares data regression line at the change point. Then a ratio of this distribution varies with the value of time, and the right hand side of equations (8) step the point estimations of the outcomes forward in time. The

point estimates continue either until an assumed trial cutoff or until the ratio $R = 1.00$ is determined for no further improvement. These point estimations of the Alert Program® therapy are given as approximately evenly spaced time-dependent treatment outcomes.

2.4 Derivation of the Weibull distribution parameters

The parameter of health outcomes, in this case, is given by θ, where the estimator $\hat{\theta}_t$ is a finite-dimensional semiparametric ratio of the Weibull cumulative distributions times the prior point estimate denoted as θ_{t_i}. This new ratio kernel $G_k, \tau(t)$ is a special case of a 2- parameter Weibull where both the shape and the scale parameter are estimated from the initial least squares regression. A 1-parameter Weibull distribution, reduces model uncertainty and potential bias of the shape parameter 'k' estimated from the slope of the OLS regression line, when τ is given from the change point and the remaining variable is time (Aron, Guo, Mettas, & Ogden, (2009). It is shown by Meng (1993) that the absolute bias ratio (ABR) of a single ratio estimator r_s will be typically small in practice even with moderate sample size. equal to or higher than 0.45 for the health professions,

Given that the shape parameter is estimated and the scale parameter is given at $\tau(t_i)$ from a least squares regression (with an R^2 equal to or higher than 0.45 for the health professions) and the outcome is the ratio times the prior value of the outcome, then the point estimations initiated at some distance from the origin are a shifted Weibull distribution by assumption and the results are point estimations or predictions.

The scale parameter $\tau(t_i)$ is defined as the time associated with the highest F statistic, in this analysis of the dataset as the mean sessions for treatment with the Alert Program®. The estimated shape parameter $\hat{k}$ is given from the slope of the initial least squares regression line, where $\hat{k}$ can be determined from the Weibull distribution as given in Freund and Walpole (1987), Rice (1995), and Tobias and Trindade (1995). The estimated shape parameter $\hat{k}$ is determined from the CDF of the Weibull distribution times $\tau(t_i)$, the value of x (the session number) at $\tau(t_i)$, when it is set equal to the slope $\hat{\beta}_1$ of the least squares linear regression from 0 - $\tau(t_i)$ at the change point, is as follows:

$$\hat{k} = \ln\left|1 - \hat{\beta}_1 \cdot \tau(t_i)\right| \tag{9}$$

In practice, the calculation of the slope should also be taken as the absolute value.
In this derivation of $\hat{k}$, the value of time or the session number is always taken at the change point. In this paper, $\hat{k} = 0.0387$ (for the highest "F" statistic) for the means of the 7 participants. Then point estimates are initiated at the change point.

2.5 Point Estimate Properties

The point estimate given by the SPRE model is unbiased at both time = τ and when the Weibull ratio distributions R = 1.00 from equation (8). The point estimation for the outcome $\hat{\theta}_t$ at τ equals the data outcome at τ. When the ratio = 1.00:

$$F\left(\hat{\theta}_t, R_{1.00}\right) = \theta_{t_\tau} \tag{10}$$

The sequential predictions of $\hat{\theta}_t$ are bounded by unbiased estimators at time = τ and when the ratio = 1.00, then $\left(\hat{\theta}_t - \theta_{t_i}\right) > R = 0$. At the change point and when the ratio $R = 1.00$, the values of the statistic $\hat{\theta}_t$ on the real axis estimate the population parameter θ (Fruend & Walpole (1987).

Furthermore, this estimator is consistent at the beginning and end points of analysis because it is unbiased at the change point and when the ratio $R = 1.00$ as shown in equation (9). Therefore, this estimator is consistent at the upper and lower bounds and the absolute bias ratio of this SPRE ratio will be typically small in practice as given in Meng (1993). The ratio estimator is insensitive to small departures from the results of the statistical assumptions and methodology used to derive the prior k and τ. Therefore, this ratio estimator is also robust.

Using this model, repeated measures may under or over estimate the standard error of a regression coefficient computed from an ordinary least squares regression (Donner, 1984). However, as long as circularity is maintained, when random errors are random, the F statistic derived from both x and y as given in equation (6) will not be affected. At τ, the parameter θ_τ is derived from the initial sample data while the predictive parameter $\hat{\theta}_t$ indicates sequential predictions for that single-subject or small group design. The mean μ is the normal distribution parameter for the response of a small group, given as parameter θ.

3. RESULTS

In this analysis, the residuals at the change points were quasi-random and normal Predictions for each of 7 participants where the data were statistically significant at the change point predictions are computed together with the mean of these 7 participants, and the results for all are graphed in Figure (1). The data is given in Appendix A.

The backwards stepwise regression analyses are carried out on the transformed data. In many therapeutic responses for evaluating statistics in disability, the responses may be high one session and very low the next. In any event, the outcomes will vary widely partly due to comorbidities (Link, Parkman, & Frame, 2012). In this study, the ACIS measurement is taken as a ratio with an initial assessment, The Sensory Profile™, to broaden the scope of the analysis as well as to normalize the outcome data. Then the square root is taken of this outcome ratio to linearize the data (on the 'y' axis). Finally, logarithms are taken of the session numbers (as a function of time on the 'x' axis).

In the means analysis, the semiparametric ratio estimator (SPRE) model predicts skill improvement via the Alert Program® treatment with an increase in skills from session 1 to session 12 of 15.4% and a smaller increase during the predicted period of 1.4%. The relative error of point estimates to the primary data is excellent for treatment where there is an early individual change point and subsequent data from the pilot study. The predictions track through the center of the test data. The SPRE ratio predictions from each change point show clearly that there is significant improvement from the start of the trial to the change point for these participants. The SPRE analysis confirms the same result measured in each participant for which there is an early change point.

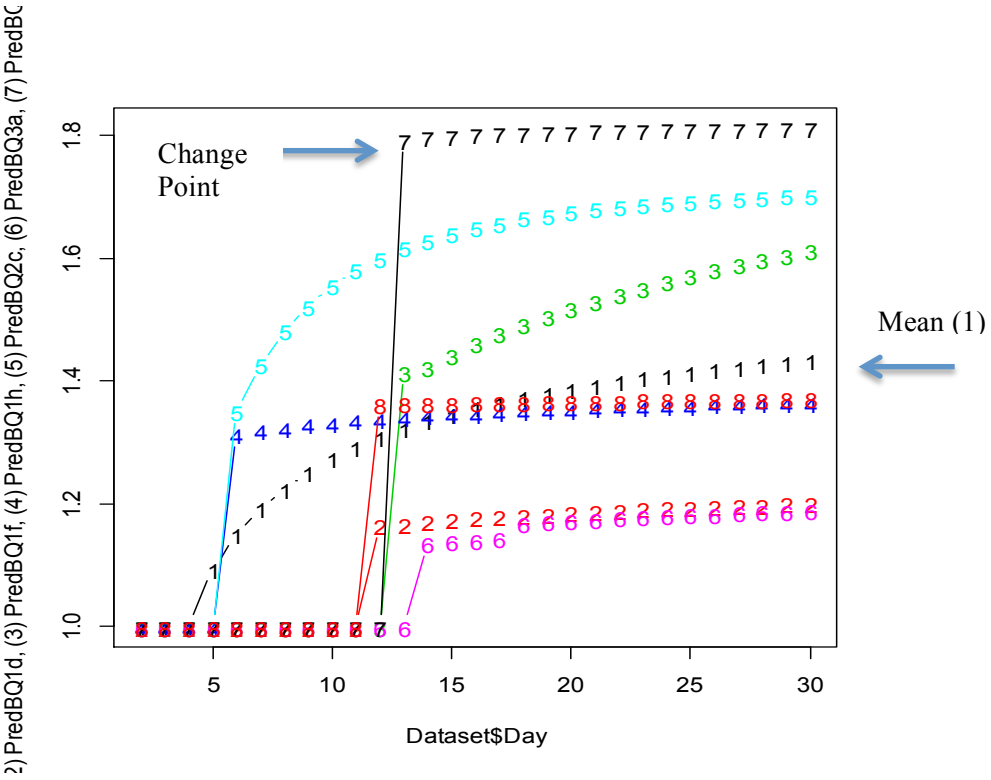

Figure (1) Point estimates of 7 participants with the mean in the ALERT program as measured by ACIS. Used with permission from Link, Parkman and Frame.

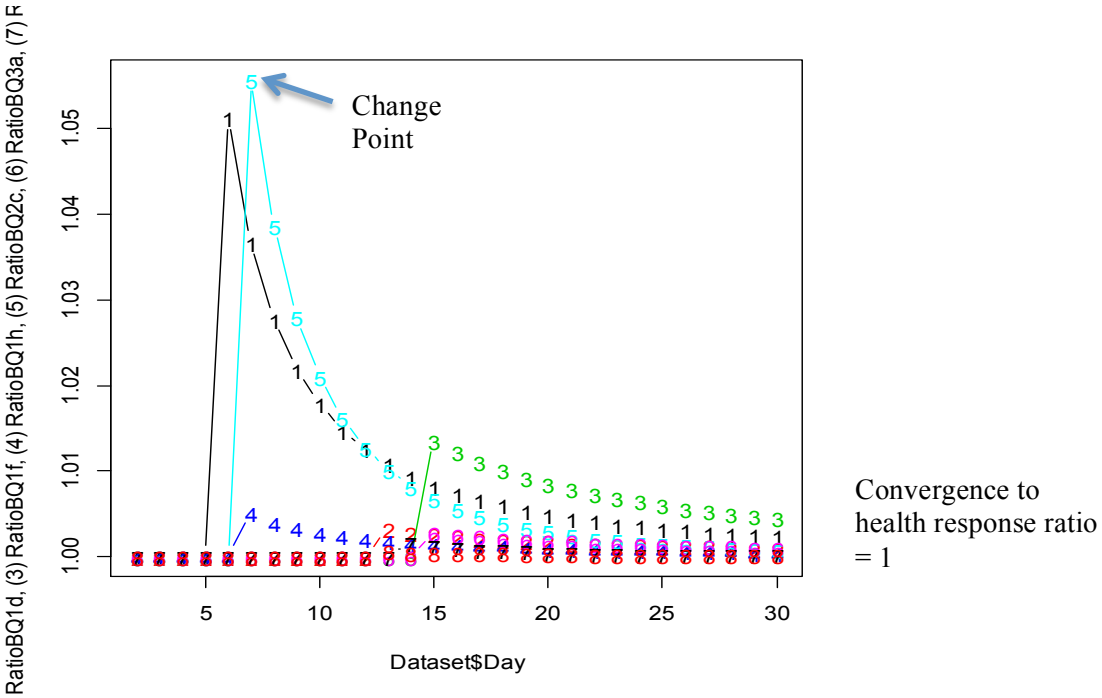

Figure (2) Health response function ratios from the change points plotted versus session day. Used with permission from Link, Parkman and Frame.

The predictions in this paper are based on the calculation of individual change points for 7 participants and the change point for the means analysis and are based on the values of previous events. For this reason, this model is especially useful in predicting the response to treatment for individuals with disabilities. Furthermore, predicting the numbers of treatments until the participant becomes stable reduces waste and controls costs in health care.

4.0 CONCLUSIONS

The relative error of predicted to test data is very good for the long-term test data for treatment of individuals with DD, as shown in Figure (2). This point estimator from the SPRE ratio in equation (8) has been shown to be unbiased at the upper and lower bounds at the change point and when the ratio $R = 1.00$, it and has also been shown to be consistent at the upper and lower bounds of prediction and robust for single subjects or a small group. Furthermore, the absolute bias ratio of this SPRE ratio will be typically small in practice as given in Meng (1993), whether the model is for a small group or the long-term predicted outcomes for a single subject. If the small group has 5–10 participants, then it is proposed that inferences may be made to a similar, larger population provided that the p value at the change point is ≤ 0.05.

REFERENCES

Aron, A., Guo, H., Mettas, A. & Ogden, D. (2009). Improving the 1-parameter Weibull. IEEE. (2009). Retrieved March, 2010, from http://www.reliasoft.com/pubs/2009_RAMS_improving_weibull_bayesian.pdf

Barth, J., Brooks, S., & Carroll B. (2011). The Effect of the Electronic Auditory Stimulation Effect Program on 8 to 12 year-old children who exhibit sensory processing deficits. (Master's Thesis, Brenau University, 2011).

Berk, R.A. (2004). *Regression analysis.* California: Sage Publications.

Cleves, M. A., Gould, W. W., Gutierrez, R. G. (2004). *An Introduction to survival analysis using STATA (Revised ed.).* College Station, TX: State Press.

Donner, A. (1984). Linear regression analysis with repeated measures. *Journal of Chronic Diseases, 37*(6), pp. 441-8.

Faraway, J. J. (2005). *Linear models with R.* New York: Chapman & Hall/CRC.

Fox, J. (2005). The R Commander: A Basic-Statistics Graphical User Interface to R. *Journal of Statistical Software,* September 2005, Volume 14, Issue 9. http://www.jstatsoft.org/

Freund, J. E., & Walpole, R. E. ((1987). *Mathematical statistics (4th ed.)*. Prentice-Hall, Inc.: Englewood Cliffs, N.J.

Heiberger, R. M., & Neuwirth, E. (2009). *R Through Excel*. New York: Springer science + Business Media LLC.

Hosmer, D. H., & Lemeshow, S., May, S. (2008). *Applied Survival Analysis: Regression Modeling of Time-to-event Data (2nd ed.).* Hoboken. N.J.: Wiley-Interscience.

Koenig, K. P., & Rudney, S. G. (2010). Performance challenges for children and adolescents with difficulty processing and integrating sensory information: A Systematic review. *American Journal of Occupational Therapy, 64,* 430–442. doi: 10.5014/ajot.2010.09073

Link, C., Parkman, C., & Frame, H. (2012). The effects of the Alert Program® on the communication and interaction skills of adults with developmental disabilities who display atypical sensory processing during group activities. (Master's Thesis, Brenau University, 2012).

Meng, X.-L. (1993). On the absolute bias ratio of ratio estimators. *Statistics & Probability Letters 18*: p. 345-348.

Miller, R., Weissman-Berman, D., and Martin, L., "A Predictive model for evidence-based practice in occupational therapy: From a single subject to the world", Research paper presentation, April, 2008, American Occupational Therapy Association Annual Conference.

R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-Project.org.

Thompson, C.K. (2007). Single subject controlled experiments in aphasia: The Science and the state of the science. NIH Public Access. Retrieved April 20, 2010 from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1847620/

Weisberg, S. (2005). *Applied Linear Regression*. New Jersey: John Wiley & Sons.

Weissman-Berman, D. (1992). "Viscoelastic Model for an FRP Sandwich Beam", Invited Paper, Second International Conference on Sandwich Construction, University of Florida, Conference Edition, Volume I, March 9-12, 1992: *SANDWICH CONSTRUCTIONS 2*(1), 3-33, EMAS, Engineering Materials Advisory Services, Ltd., 1992.

Weissman-Berman, D., (2009). "Predicting long term response to treatment for prostate cancer based on short term linear regression", Presentation at the 50th Anniversary Celebration of FSU's Statistics Department, April 17-18.

Weissman-Miller, D.(2010). "A Novel Semiparametric Ratio Estimator (SPRE): A Key to Predicting Long-Term Weight Loss in Obesity", JSM Proceedings, 2010.

Wu, B. (2005). Differential gene expression detection using penalized linear regression models: The improved SAM statistics. Retrieved March, 22, 2007, from: http://bioinformatics.oxfordjournals.org/cgi/content/full/21/8/1565

Appendix A – Point Estimation Data

PredMEAN	PredBQ4a	PredBQ1c	PredBQ1d	PredBQ1f	PredBQ1h	PredBQ2c	PredBQ3a
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1.0929	1	1	1	1	1
1	1	1.149	1	1	1.3118	1.3509	1
1	1	1.1912	1	1	1.3186	1.4263	1
1	1	1.2243	1	1	1.324	1.4815	1
1	1	1.2511	1	1	1.3285	1.523	1
1	1	1.2735	1	1	1.3324	1.5551	1
1	1	1.2923	1	1	1.3357	1.5803	1
1.3617	1	1.3086	1.1637	1	1.3386	1.6004	1
1.3627	1.7937	1.3227	1.1675	1.4142	1.3411	1.6167	1
1.3635	1.7967	1.3351	1.1709	1.4232	1.3434	1.6301	1.1349
1.3643	1.799	1.3462	1.1741	1.4427	1.3455	1.6412	1.1381
1.365	1.801	1.3562	1.1769	1.4604	1.3474	1.6506	1.1411
1.3657	1.8026	1.3652	1.1795	1.4766	1.3492	1.6585	1.1438
1.3664	1.804	1.3734	1.1819	1.4917	1.3508	1.6653	1.169
1.3669	1.8053	1.3808	1.1841	1.5056	1.3523	1.6711	1.1713
1.3674	1.8064	1.3876	1.1862	1.5185	1.3537	1.6761	1.1735
1.3679	1.8073	1.3938	1.1881	1.5306	1.3551	1.6805	1.1756
1.3683	1.808	1.3997	1.1899	1.5419	1.3563	1.6842	1.1775
1.3687	1.8087	1.405	1.1916	1.5525	1.3575	1.6876	1.1793
1.3691	1.8092	1.4101	1.1933	1.5626	1.3586	1.6905	1.1811
1.3695	1.8097	1.4148	1.1949	1.572	1.3597	1.693	1.1828
1.3699	1.8102	1.4192	1.1963	1.581	1.3607	1.6954	1.1843
1.3703	1.8106	1.4233	1.1977	1.5895	1.3617	1.6974	1.1857
1.3706	1.811	1.4271	1.199	1.5976	1.3627	1.6993	1.1871
1.3709	1.8114	1.4308	1.2002	1.6053	1.3635	1.701	1.1884
1.3712	1.8118	1.4342	1.2014	1.6127	1.3643	1.7025	1.1897