Adjustment for the Regression to the Mean Effects in Studies with Repeated Measures

by

Quazi Imad Uddin Ibrahim

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Epidemiology

Department of Public Health Sciences
University of Alberta

Abstract

In repeated measures data, large or small values at the initial measurement tend to be followed by values that are closer to the mean at the follow-ups measurements. This tendency is called regression to the mean (RTM). The presence of the RTM effect is inevitable in repeated measures data because of less than perfect correlation (correlation coefficient < 1) between the repeated measurements. Despite the growing evidence of the presence of RTM effects in clinical and public health studies based on repeated measures data, very few studies have evaluated and considered them when interpreting observed changes over time. In intervention studies, an RTM effect is mixed with an intervention effect in observed changes. It is extremely important to separate the RTM effect from the observed change in order to isolate any intervention effect and thus to make valid inferences about the effect of the intervention. In studying changes in outcome variables in repeated measures studies, RTM effects should always be adjusted for the valid interpretation of the changes and unbiased assessment of the intervention effects. The choice of methods to control for the RTM effect should be based on the type (continuous, count) and shape (normal, non-normal distribution) of the outcome variables of interest. A new method of estimating RTM effects for non-normal data using simulation is proposed. The method is a combination of bootstrap sampling from the standardized outcome variable and matrix decomposition of the correlation matrix between the repeated measurements. The method is applied to adjust for the RTM effects in studying changes in mean drinks in a typical week in a study evaluating the impact of a brief alcohol intervention on youth. In the study, mean drinks followed a positively skewed distribution. The proposed method estimated the RTM effects considering the true distribution (positively skewed) of the outcome and in doing so, provided more accurate estimation of the intervention

effects compared to other methods considered in the thesis. The method ensured valid interpretation of the observed changes in the outcome by providing the most accurate estimation of the RTM effect and then removing it from the data. The proposed method could be applied to adjust for the RTM effect in non-normal repeated measures studies.

Preface

This thesis is an original work by Quazi Ibrahim. The research project, of which the data were used in this thesis, received research ethics approval from the University of Alberta Research Ethics Board, Project Name “Youth Transitions Project”, University of Alberta HERO ID No. Pro00003829. No part of this thesis has been previously published.

Dedication

This thesis is dedicated to my family. I am grateful to my family for their continuous support and encouragement. I remember their love, affection and inspiration throughout my life.

Acknowledgements

All praises are due to almighty God. I express my sacred gratitude to the omnipotent and omniscient God, Who gave us the wisdom of creation. Without His blessings this task could not have ended.

I express my cordial gratitude to my supervisors Dr. Ambikaipakan (Sentil) Senthilselvan and Dr. Katerina Maximova. I am heartily thankful to them for their continuing support and encouragement during the preparation of the thesis. Their guidance and co-operation helped me to complete the task. I would like to thank Dr. Cameron Wild for sharing the Online Personalized Assessment-Feedback (PAF) Alcohol Intervention Study data for the proposed research and for serving on my thesis committee. I am thankful to Dr. Irina Dinu, School of Public Health for serving on my thesis committee. I acknowledge the help, co-operation of all my respected teachers, class mates, the members of staff of the School of Public Health of the University of Alberta.

Contents

List of Tables

List of Figures

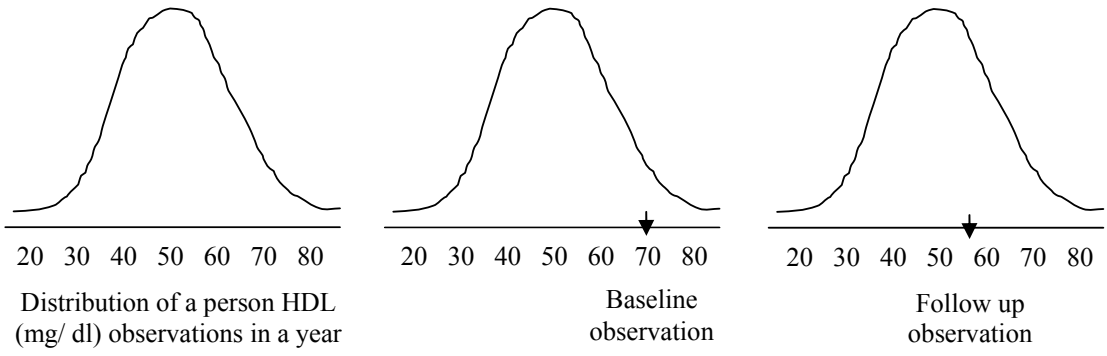
Chapter 1

Introduction

1.1 Background

In repeated measures data, large or small values at the initial measurement tend to be followed by values that are closer to the mean at the measurements taken during the follow-up [1], [2]. This phenomenon is called regression to the mean (RTM) [3]. RTM occurs because individual values are observed with random variation (error) around a true mean. Consider a hypothetical example where observations of a person high density lipoprotein (HDL) in a year follow a normal distribution with mean 50 mg/dL and standard deviation 10 mg/dL. If we observe a much higher HDL value for that individual than his/ her typical value, e.g., 70 mg/dL, it is likely that on follow-up measurement the value would be closer to its mean (50 mg/dL) [3].

Figure 1.1. A hypothetical example of the RTM effects in a person's HDL measurements



RTM occurs at group level as well. Suppose HDL in a population follows a normal distribution with mean 50 mg/dL and standard deviation 12 mg/dL. We select a group of individuals with high HDL values (> 70 mg/dL) from that population based on their initial

measurements. In that group, there would be more individuals with true HDL values below 70 mg/dL than individuals with true HDL values above 70 mg/dL because of random variation in the data and because the mean HDL of the population is 50 mg/dL. On follow-up measurement, the mean HDL value of the group will decrease; subjects with high initial values as a result of random fluctuation tend to approach the population mean of 50 mg/dL [3]. RTM is common in repeated measures studies where repeated measurements are made on the same subjects over time, especially in intervention studies that include subjects that are at higher risk based on high or low value of a clinical characteristic. RTM presents a particular challenge for intervention studies since a change in value due to RTM can erroneously be attributed to the effect of an intervention. This chapter will provide a historical background of the RTM effect; discussions of the concept in epidemiological context; descriptions of different methods of controlling the effects at design and analyses stages; and examples of its applications in various fields. This chapter will conclude with the proposed objectives of the thesis.

1.2 Literature Review

1.2.1 Historical Background

The concept of RTM was introduced over a century ago by Galton [4], [5]. In the experimentation with the growth of peas, it was observed that offspring from tall plants were shorter than either of the parent plants. Similarly, offspring of two shorter plants were taller than either of the parent plants. Galton referred to this phenomenon as “regression towards mediocrity”. This phenomenon had been observed in human’s stature as well with the tendency of children of taller parents to be shorter and vice versa [6].

Galton was the first to provide empirical documentation of RTM [4], [5], [6]. Later it was established as a statistical concept in repeated measures data based on theoretical deductions [7], [8]. Although RTM occurs frequently in repeated measures studies, the effects of RTM are not always described in published reports. RTM is widely misunderstood despite

its simplicity. RTM is often confused with regression which is widely used to explain the least square fitting of lines, curves and surfaces for prediction.

1.2.2 RTM in Epidemiologic Context

RTM is a form of selection bias that affects follow-up studies that involve repeated measurements of individuals over time [9]. In health care, interventions often target people that are at higher risk for adverse outcomes based on clinical measures, costs or utilization of health services. A group of such individuals is selected and invited to participate in an intervention intended to reduce their level of risk. After a period of time, follow-up assessments are done. Any change in the outcome before and after the intervention can be incorrectly attributed to the effect of an intervention. Specifically, initial elevated level of risk may have a reduction in the level of risk without the intervention due to RTM. In observed changes over time, the RTM effect need to be assessed. RTM is a major source of bias in evaluating effects of interventions and may have significant implications for patient care, health service delivery and policy development [10]. In epidemiologic textbooks, RTM is also referred to as 'the regression paradox', 'the regression fallacy', or 'the regression trap' [8]. Policy makers, researchers should be aware of RTM when evaluating the effects of interventions, and take necessary steps to adjust for it. The possibility of RTM should be considered when interpreting observed changes over time to obtain valid conclusions about the effectiveness of an intervention.

1.2.3 Theoretical Framework

Suppose that Y_1 and Y_2 are the first and second measurements of an outcome variable Y, respectively. Let the correlation coefficient between Y_1 and Y_2 be r. If we fit a simple linear regression with Y_2 as the response and Y_1 as the regressor variable, the slope of the regression line would be $\frac{rs_{y_2}}{s_{y_1}}$ where s_{y_1} and s_{y_2} are the standard deviations of Y_1 and Y_2, respectively. One standard deviation change in Y_1 results in r standard deviations change in Y_2. If Y_1 and Y_2 are not exactly linearly related ($r < 1$), for a given value of Y_1, the predicted value of Y_2 is

fewer standard deviations away from its mean compared to that of Y_1. Thus, the RTM effect almost always occur in repeated measures data [7]. In clinical practice, treatments often aim to reduce the level of risk factors such as weight, cholesterol and blood pressure. People with higher values of such measurements are then treated. On repeat measurements in repeated measures study, a mean reduction in the level of the risk factor will be observed. However, this reduction should not be interpreted as a treatment effect. This is because the RTM effect, if present, is mixed in with the treatment effect. If there is no treatment effect, the reduction is inevitable due to less than prefect correlation between the first and second measurements. RTM is a ubiquitous phenomenon in repeated measures data [3]. The RTM effect needs to be separated from the treatment effect for valid evaluation of the treatment or intervention effect.

Typically, negative correlations between baseline values and observed changes at follow ups are considered as an indication of the presence of RTM effects in the data [11], [12]. This method of identifying the RTM effect may not be conclusive and may lead to erroneous conclusion. A spurious high negative correlation between baseline values and changes (follow-up values – baseline values) may be observed due to the common component: baseline values [13]. The common component between both values, which is called mathematical coupling between both values, induces a correlation [14]. If the RTM effect is present in the data, the variance of the variable of interest at follow-ups would shrink compared to the baseline variance [13]. As a result, in presence of the RTM effect, the correlation between observed changes (follow-up – baseline values) and the sum of follow-up and baseline values will be negative. Suppose, X and Y represent baseline and follow-up values, respectively. Then, the expected covariance of the association is $E[(Y-X)*(Y+X)] = E[Y^2] - E[X^2]$. In presence of the RTM effects, follow-up variance would be less than the baseline variance, which implies $E[Y^2] - E[X^2]$ would be negative. Therefore, a plot of observed changes against the sum (or mean) of baseline and follow-up values or the correlation between both values can be used to detect the RTM effects.

The RTM effect can be accounted for either at the design or analysis stage of a study [3]. Studies can be designed to minimize the RTM effect by selecting an appropriate control group, taking two or more baseline measurements. At the analysis stage, the RTM effect can

be estimated from the data through statistical modeling and removed from the observed change in repeated measures values. Each of theses strategies is discussed below, with a particular emphasis on their strengths and weaknesses under different scenarios. Strategies to minimize and to remove RTM effects at the study design and analyses stages are described in the following section with a discussion of their advantages and disadvantages.

1.2.3.1 Strategies to Minimize RTM Effects at the Design Stage

Use of Average of Multiple Measurements for Selecting Subjects

Studies can be designed to have two or more measurements at screening stage. Subjects would be selected based on the average of these measurements [15] [16]. This approach is simple to apply and can minimize the RTM effect by reducing variability in repeated measurements over time.

Two Baseline Measurements: First for Selection and Second for Comparison

Another approach at the study design stage is to take two baseline measurements and use the first measurement to select the subjects and assess the intervention effects from the change from the second measurement [17]. This approach assumes that RTM occurs between the first and second baseline measurements. Suppose that Y_1 and Y_2 are the first and second baseline measurements before an intervention and Y_3 is the first follow-up measurement after the intervention of an outcome variable Y. Let r_{ij} be the correlation coefficient between i^{th} and j^{th} measurements. Then, the regression coefficient for Y_3 on Y_2 having classified on Y_1 is $b_{32(1)} = \frac{r_{31}s_3}{r_{21}s_2}$ where s_2 and s_3 are the standard deviations (SDs) of Y_2 and Y_3, respectively. A necessary condition for the removal of the RTM effect is $b_{32(1)} = 1$, i.e., $r_{31}s_3 = r_{21}s_2$. Therefore, with this approach, the RTM effect would be removed if the correlation between the first and second measurements is exactly the same as that of the first and third measurement under stable SDs ($s_2 = s_3$).

Use of a Control Group

A parallel control group, if it is ethically feasible, can be used to eliminate the RTM effect at the study design stage [18], [19]. Outcomes in both control and intervention groups will be affected alike by RTM if subjects are randomly allocated to the groups. The mean change in an outcome in control group will give an estimate of the RTM effect and the placebo effect. The difference in mean changes in the outcome between the intervention and the control groups will remove the RTM effect and give an unbiased estimate of the intervention effect on the outcome.

Use of a Second Measurement in a Screening and Treatment Program

In a screening and treatment program for blood pressure (BP) reduction in a community, after initial BP measurements, participants with BP greater than a predefined age-specific cut-off value were asked to return for a re-screening within 4 weeks [20]. Suspected hypertensive patients were not referred for medical evaluation before the re-screening. The differences between the first and second BP measurements were taken. Since the “program” was only informing a person about his BP and asking for repeat measurement, the net "program" effect in the observed changes between the first and second BP measurements was assumed to be zero. The average of the observed changes was attributed to the RTM effect.

Selection of subjects based on the average of two or more measurements is the simplest approach and can minimize the RTM effect by reducing intra-individual variance and measurement error. However, this method cannot guarantee the complete elimination of the effect. The method of using the first measurement for selection and the second measurement for assessing the intervention effect completely eliminates the RTM effect if correlations between the first and successive measurements are exactly the same which would be unlikely in practice. The estimation of the RTM effect from a screening and treatment program can be biased. The reduction in the second measurements on re-screening can occur due to other extraneous causes. The most effective way of removing the RTM effect is through the use of a parallel control group. If the balance between the intervention and the control group is

achieved from a randomized controlled trial, the comparison of changes between the groups eliminates the RTM effect.

1.2.3.2 Strategies to Remove RTM Effects at the Analysis Stage

Estimation and Adjustment of the RTM Effect for Normal Data

In repeated measures data, the RTM effect can be estimated and then subtracted from the observed change. Suppose that Y be the variable of interest and K be the cut off point (high) for selection. Assuming $Y \sim N(\mu,\ \sigma^2)$, the RTM effect is estimated as

$$c\sigma(1-\rho) \qquad (1.1)$$

where c is the ratio of the ordinate of standard normal distribution at z [$=(k-\mu)/\sigma$] and area under the standard normal curve > z and ρ is the correlation between repeated measurements in an individual [19]. Thus, the RTM effect increases if within subject variability (σ) increases or the cut-off point (k) for selection is more extreme. It decreases if the correlation between repeated measurements increases. There is no RTM effect if the correlation is 1.

The parameters in equation (1.1) are usually unknown. External estimates of these parameters from related studies can be used [18] or they can be estimated from observed samples. One simple method used in the estimation of these parameters is the method of moments [19], [21] which involves equating sample moments with population moments and then solving those equations. Another widely used method is the maximum likelihood method [22], [23], [24] where the parameters are estimated in such a way that the likelihood of getting the observed data is maximum. When repeated measurements before and after the intervention are available, a model accounting within subject variability in addition to measurement errors can be used to estimate the RTM effect [25].

For normal data, the RTM effect can be estimated through simulation [10], [26] by:

1. simulating two sets of observation from the parent normal distribution, representing the first and second measurements;

2. selecting observations with the first measurement greater than a specified cut-off value determined mainly based on clinical considerations; and
3. calculating the difference between the means of first and second measurements in this group.

The RTM effect is equivalent to the difference between the two means:

$$\text{The RTM effect} = \text{Mean}_1{}^{\text{st}}{}_{\text{measurement}} - \text{Mean}_2{}^{\text{nd}}{}_{\text{measurement}}.$$

The RTM effect can be estimated under different scenario for normal data. It can be estimated when population mean and variance change over time [27]. It can be estimated based on a sequence of observations for different classical tests theory and for autocorrelation models [28].

Under normality assumption, a regression based test was developed to compare the means before and after the intervention in the presence of the RTM effect [29]. A simple linear regression model with the second measurement as the response and the first measurement as the independent variable was fitted to the data. The least square estimate of the intercept is equivalent to the estimated intervention effect controlling for the RTM effect. This approach controls for selection by conditioning on the first measurement, that is, by modeling the expected value of the second measurement given the first measurement. The test of the null intercept is the test for the intervention effect after controlling for the RTM effect. The tendency of the follow-up measurements to be closer to the mean due to RTM was examined in multiple linear and logistic regression models [30]. The shrinkage due to this tendency was found to be a serious problem if the sample size was small and/ or the number of covariates was large. A two-stage approach of estimating the RTM effect for absolute and percent change was developed and its relationship with ANCOVA was shown [31].

Among the methods of estimating RTM effects for normal data, the method of moments is widely used given its simplicity. When certain parameter values are known or reliable estimates are available from external data, the method is computationally

straightforward. However, the method often produces inefficient estimators in case of small samples [32]. The maximum likelihood approach is more advantageous. The maximum likelihood estimators are more efficient compared to the method of moment estimators [32], [33]. The method is adaptable to different models (additive or multiplicative) and different types of sampling (truncated, selected, censored or complete). The major drawback however, is that the method requires intensive computation when analytical solutions are not available. When repeated and replicated measurements are available, the method accounting for within subject variability is more appropriate [32]. Simulation based approaches yield similar estimates as those obtained from the formula- and or model-based approaches. However, with non-normal data, all these methods described above produce biased estimates of the RTM effect.

Estimation and Adjustment of the RTM Effect for Non-normal Data

In repeated measures studies, we frequently observe data that are not normally distributed (i.e. skewed). For example, the serum cholesterol data from the UK Prospective Diabetes Study (UKPDS) [26]; the alcohol consumption data from the cohort study involving students of three tertiary educational institutions in New Zealand are all skewed to the right [34]. In such data, the nature of the RTM effect remains same, an extreme value observed on one occasion tends to be followed by a less extreme observation. However, there are certain differences in the regression. Suppose Y_1 and Y_2 are the first and second measurements, respectively, of a response variable Y which does not follow a normal distribution. The regression of Y_2 on Y_1 is not linear, homoskedastic unlike the normal case [35]. Slope of the regression is not always less than 1. In case of extreme non-normality, the regression shows oscillation. The expected value of latter observation conditional on previous observation, $E[Y_2 | \ Y_1 = y]$, regress to either mode [36] or mean [37] or some other values [38] depending on the distribution of non-normal data. Suppose

$$Y_j = U + \varepsilon_j, \quad j = 1,2,$$

where

1. Y_1 and Y_2 are jointly distributed with mean μ and variance σ^2 and correlation ρ $(0 < \rho < 1)$;
2. U is a subject's true value which is arbitrarily distributed with mean μ and variance $\rho\sigma^2$ and
3. $\varepsilon_j \sim NID\left(0,(1-\rho)\sigma^2\right)$ which are independent of U.

Then the RTM effect can be expressed as

$$E[Y_1 - Y_2 \mid Y_1 = y] = -(1-\rho)\sigma^2 \frac{1}{dy}\ln[g(y)], \tag{1.2}$$

where $g(y)$ is the probability density function of Y and dy is the differential distance in y, that is, infinitesimal change in y [36]. In this case, RTM corresponds to shrinkage towards the mode of the distribution. For $Y_1 > y$, equation (1.2) can be expressed as

$$E[Y_1 - Y_2 \mid Y_1 > y] = (1-\rho)\sigma^2 \frac{g(y)}{1-G(y)}, \tag{1.3}$$

where $G(y)$ is the distribution function of Y. Kernel density estimation and the kernel estimation for hazard rate are used to estimate the RTM effect in equation (1.3) [39]. Both methods are precise in estimating the RTM effect. The methods are applicable if there are data for only two time points.

In another example, suppose Y_1 and Y_2 be the number of visits to a doctor in a population in two time periods, respectively [37]. In that population, for an individual, chance of visiting the doctor at a time period depends on that individual's proneness. Suppose, proneness, λ, in that population follows a gamma distribution with probability density function,

$$f(\lambda) = \frac{1}{\Gamma(\alpha)}\lambda^{\alpha-1}\beta^{\alpha}e^{-\lambda\beta}, \ \ \lambda \geq 0, \ \ \alpha > 0, \ \ \beta > 0.$$

Then, Y_1 and Y_2 follow a bivariate negative binomial distribution of the following form

$$\Pr(Y_1 = y_1, Y_2 = y_2) = \frac{\Gamma(y_1 + y_2 + \alpha)\beta^{\alpha}(\beta+2)^{-(\alpha+y_1+y_2)}}{y_1!\ \ y_2!\Gamma(\alpha)},$$

where $y_1 = 0,1,2,...$; $y_2 = 0,1,2,...$; and $\alpha > 0$ and $\beta > 0$ are the parameters of the distribution. The RTM effect can be expressed as [37]:

$$E[Y_1 - Y_2 \mid Y_1 = y] = \frac{\beta(y - \alpha/\beta)}{\beta + 1}. \quad (1.4)$$

In equation (1.4), RTM corresponds to shrinkage towards the mean of the distribution.

When the distribution of the underlying variable and the contaminating errors are not normal, unknown, under certain regularity conditions the RTM effect is estimated based on local sample means using asymptotic justifications [40].

For non-normal data, the parametric and non-parametric methods can be used to estimate the RTM effect. However, the relative efficiency of these methods is not known.

1.2.4 Examples of Adjusting for the RTM Effects in Epidemiologic Studies

In this section, examples of repeated measures studies are given where the presence of the RTM effect was examined and or an adjustment was made for it, if present.

RTM in Health Care

Substantial evidence of the presence of RTM effects exists in health care studies [10], [41], [42] [43], [44]. However, few studies have evaluated the extent of impact of RTM on study findings or considered it in interpreting the observed change. Results of few existing studies that considered RTM when evaluating intervention effects are discussed below.

In routine clinical practice, patients with rheumatoid arthritis were selected for TNF-alpha inhibitors on the basis of high disease activity scores (DAS). In 35 such patients, selected from three hospitals of London, changes in DAS were studied 9-21 months prior and 1.5-6 months post treatment [45]. The magnitude of the RTM effect was determined in the changes in DAS. The RTM effect was adjusted by regressing the follow-up scores on the initial scores. The estimated intercept from the simple linear regression model was the change

in DAS after adjusting for the RTM effect. The study showed that improvements in DAS from the biological therapy would be overestimated if the RTM effect was not taken into account.

In the Multiple Outcomes of Raloxifene Evaluation (MORE) trial, evidence of an RTM effect was found in changes in biochemical markers of bone turnover in women treated with Raloxifene [46]. The women who experienced extreme changes in biomarker levels after 6 months of treatment had changes in the opposite direction on a subsequent measurement at 1 year. Specifically, among women who had a decrease of 60% or more in the urinary CTX in the first 6 months, 61% had an increase in the next 6 months. Among women who had an increase in the urinary CTX in the first 6 months, 81% had a decrease in the next 6 months. Extreme values at either end were affected by RTM. The RTM effect in change in biomarker level was adjusted by replacing an individual observed baseline value by a 'true' baseline value. The true baseline value for an individual was obtained by taking a weighted average of the observed baseline value and the overall mean of the study population. The weights are the inverse of within individual variance and the overall population variance, respectively. After adjusting the RTM effect, evidence of the effectiveness of Raloxifene as a treatment for urinary CTX was inconclusive, warranting further evaluation.

In the United States, the effect of shell-issue law on states' murder rate was studied with or without adjusting for the RTM effect [47]. A Poisson regression was fitted to obtain slopes for relative murder rates for five years before and five years after the adoption of the law for each of the twenty five states. The RTM effect in murder rate slopes was adjusted by regressing (linear) the post slopes on the pre slopes. The estimated intercept was the adjusted change in the relative murder rate after the law was adopted. In the presence of the RTM effect in the data, the analyses without adjusting for the RTM effect showed a beneficial effect of the law on states' murder rate, but the effect dissipated after adjustment for the RTM effect.

Disease modifying therapies (DMT) are generally accepted as effective therapeutic treatments in reducing relapse rate in multiple sclerosis (MS) patients though some controversies of the treatments in reducing residual disabilities [48]. In a non-treated cohort of 44 relapsing-remitting multiple sclerosis (RRMS) patients, the RTM effect was evaluated in

the relapse rate in order to correctly identify response to DMT [49] [50]. The patients with two or more relapses in the prior two years, and a baseline expanded disability status scale (EDSS) score less than or equal to 5.5 were recruited in Spain in 1994. The patients were monitored for 1 year using a standardized protocol (EDMUS) [51]. In the non-treated cohort, the mean number of relapses decreased from 1.7 to 1.0 in the first year. The absolute mean reduction of 0.7 in the RRMS relapse rate was suspected due to RTM and considered as an estimate of the effect under the assumption of no placebo effect.

In a cohort of male employers from a single industry in India, a five-year trend in blood pressure (BP) was studied using exploratory data analysis (EDA) and regression methods [52]. The decline in the prevalence of hypertension and mean systolic BP over 5 year period without any intervention was considered due to the possible RTM effect.

Disease management is as an important aspect of clinical practice. The effectiveness of a disease management intervention in asthmatic population was studied in Colorado during 2001 to 2003 [53]. Using a non-randomized control group to adjust for the RTM effect, the study showed that the intervention was effective in reducing healthcare costs. The use of a convenience control group in adjusting for the RTM effect was criticized to be biased [54]. A further investigation using a randomized controlled group was suggested to examine the effectiveness of the intervention.

In a double-blind, randomized, controlled trial, the effects of Haloperidol, Trazodone and behavioural therapy on dementia was studied [55]. In the trial, 148 patients with Alzheimer disease were randomly allocated to the four study arms: Haloperidol (n=34), Trazodone (n=37), behavioural therapy (n=41) and placebo (n=36). The main outcome was Cohen-Mansfield Agitation Inventory (CMAI) score [56], [57]. The CMAI comprised of 36 agitated behavior related questions each measured on a 7-point Likert scale ranged from 0 to 6. The score was the sum of all 36 items' score. Changes in the score from the baseline to the end of the study (week 17) in the four groups were studied and compared. No effect of treatment on the outcome was observed. The observed mean change of 6.0 in CMAI score in the placebo group was similar to the estimated RTM effect of 5.88. The RTM effect was computed by the estimated expected change in the outcome from the baseline to the end of the

study in an external validation sample (n=137) where no treatment was administered [58]. The changes observed in the four study groups were considered due to RTM.

RTM in Substance Use Disorder Research

The RTM effect is a persistent phenomenon in Substance Use Disorder (SUD) treatment research, however it has been largely ignored in previous studies [16] [59], [60]. The effect of a brief intervention on alcohol consumption was studied among 967 students from three New Zealand tertiary educational institutions [34]. The possible presence of the RTM effect in the alcohol consumption was explored graphically. Available methods of estimating and adjusting the RTM effect for normal data were not considered due to positive skewness in the alcohol intake data. The presence of RTM effects was found in the reduction of alcohol consumption. A negative association between baseline AUDIT scores and changes in AUDIT scores from baseline to six months was observed. Students who had lower baseline scores tended to have higher follow-up scores. Also, those who had higher baseline scores tended to have lower follow-up scores.

In another repeated measures study in Finland, southern Sweden and Denmark, the effect of a change in alcohol policy on alcohol consumptions was studied [61]. In 2004, in these three Nordic countries, alcohol taxes were lowered by one third and travellers were allowed to import unlimited alcohol for own use. The data collected in 2003 to 2005, that is, before and after the policy change, were examined. A linear regression model of latter alcohol intake on the initial alcohol intake was used to account for the RTM effect. After accounting for the RTM effect, among heavy drinkers, there was no significant change in alcohol consumption after the policy change in all three countries. In southern Sweden, light drinkers raised their alcohol consumption.

RTM in Exercise Outcomes

RTM effects have been often ignored in studies of exercise and health outcomes [62], [63], [64], [65]. Generally, the effectiveness of an exercise program is evaluated by studying

changes in the initial condition of an individual such as blood pressure, ventricular dimensions, ST segmental depression, or serum cholesterol levels at follow-up measurements. Initial high or low values of a clinical characteristic tend to regress towards the mean without any intervention. Thus such studies are susceptible to the RTM effect. The common belief that the extent of aerobic training is inversely related to the initial fitness of the individual was shown to be wrong after accounting for the RTM effect [12].

Despite growing evidence informing clinical and public health practices based on repeated measures data, very few studies evaluated the presence of RTM effects and/ or considered it in the interpretation of the observed change. In most of these studies, the apparent benefit of the interventions dissipated after the adjustment for the RTM effect. This underscores the importance of evaluating the presence of RTM as a source of bias in current research as many interventions may deceivingly appear successful in the presence of RTM. It also highlights the need for the development of new, sophisticated methods to adjust for the RTM effect in intervention studies in order to avoid making misleading conclusions about the effectiveness of an intervention. Better detection of the RTM effect and elimination of this bias from health intervention studies has significant implications, and will prevent unnecessary and costly changes to current clinical and public health practice to implement interventions that are in fact ineffective.

1.2.5 Summary

Although RTM is common in repeated measures data, it is often neglected in analyses and interpretation of results. It is a major source of bias in intervention studies. The RTM effect is often mixed in with any genuine intervention effect in observed changes. It is very important to separate the RTM effect from the observed change in order to isolate any genuine intervention effect, made valid evaluation of the intervention. Different methods have been developed to control for the RTM effect at the design and analysis stages. At the design stage, the best method to eliminate the RTM effect is the use of a parallel control group. At analysis stage, RTM effects are estimated and removed from the observed changes. Various methods of estimating RTM effects were developed for normal and non-normal data. For normal data,

maximum likelihood method; method of moments; method accounting within subject variability; simulation; and regression-based approach were proposed under different situations. All of these methods proposed for normal data have been shown to be biased in cases of departure from normality. For non-normal data, parametric and non-parametric methods (local sample means; kernel density estimation) have been used for estimating the RTM effect. However, the relative efficiency of these methods is not known and has not been systematically investigated. To the best of my knowledge, simulation has not been used for non-normal data in estimating the RTM effect.

1.3 Objectives

The objectives of this thesis are:

1. Develop a simulation-based method of estimating the RTM effect for non-normal (skewed) data.
2. Perform a comparative study of different methods of estimating and adjusting the RTM effect at design and analysis stages including the novel approach.

Chapter 2

Methods

2.1 Repeated Measures Data

In a study, if the outcome data are collected on two or more occasions on each subject or unit in the study population, it is called repeated measures data. Consider a hypothetical example such as a blood pressure (BP) reduction program. BP may be measured on each individual enrolled in the study three times, one at baseline and repeated twice at 3 and 6 months following the intervention. Such data can be represented as

$$Y_{ij},\ \ i = 1,2,\ldots,n;\ \ j = 1,2,3;$$

where n is the total number of individuals enrolled in the study. In general, Y_{ij} is the j^{th} BP measurement on the i^{th} subject. Specifically, when observations are measured at different time points, it is called longitudinal data [1], [2]. The observations from an individual taken at different time points are generally correlated. Methods are developed to take into account such correlation in the statistical analyses of the data. There are mainly two approaches in the correlated data analyses: 1) population average/ marginal approach; 2) conditional approach.

In the marginal approach, average pattern in the outcome in the study population (population average (PA) pattern) is considered. Observations from the same individual tend to behave similarly to the PA. This similar behaviour is addressed by within subject or unit correlation. Generalized Estimating Equation (GEE) is a popular marginal method used in correlated data analyses [66], [67].

In the conditional approach, it is assumed that natural heterogeneity exists between subjects or units. Each subject (or unit) is assumed to have its own intercept and/or slope (random effects). Observations are independent conditional on the subject-specific intercepts and/or slopes. Random effects modeling are used in this approach [66]. In repeated measures studies with more than two waves of data, linear random effects modeling is widely acceptable as an appropriate method in studying association between baseline value of a continuous variable and subsequent changes [14]. In Section 2.2, random effects modeling will be discussed, in brief, as a method of analyzing repeated measures data.

In repeated measures data, the RTM effect is a major source of bias. Chapter 1 reviewed how common RTM effects are. Various methods of adjusting for the RTM effect in the design and analysis stages of a study were discussed. In the design stage, the effect can be accounted for by using an appropriate control group. There are mainly two approaches to control the effect of RTM at the analysis stage. The RTM effect is estimated in the analysis using formulas or simulations and then subtracted from the observed changes. It can also be adjusted for by testing coefficients from the regression models. In the simulation based approaches that have been proposed until now, the RTM effect was estimated only for normally distributed repeated measures data. A new method of estimating RTM effects for non-normally distributed repeated measures data using simulation is presented in section 2.3. As a practical illustration of this new approach, the method will be applied to a study examining the effectiveness of providing online personalized assessment-feedback on alcohol use to young adults. A description of the study and its data will be given in Section 2.4. Finally, in Section 2.5, an analytical plan comparing different methods of adjusting the RTM effect, including the proposed new method, will be presented for the PAF intervention study.

2.2 Random Effects Modeling for Repeated measures Data Analysis

In random effects modeling, the relationship of an outcome variable Y with exposure or regressior variables (Xs; Zs) can be expressed as a generalized linear mixed model of the following form:

Systematic Part

$$f\left(E[Y_{ij} \mid b_{0i}, b_{1i}, b_{2i}, \ldots, b_{qi}]\right) = \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \ldots + \beta_p X_{pij} + b_{0i} + b_{1i} Z_{1ij} + b_{2i} Z_{2ij} + \ldots + b_{qi} Z_{qij} \quad \ldots (2.1)$$

where $i = 1, 2, \ldots,$ n and $j = 1, 2, \ldots, n_i$. Here,

Y_{ij}: j^{th} observation on the i^{th} subject;
X_{lij}: j^{th} observation on the l^{th} fixed regressor (l=1, 2,…, p) for the i^{th} subject;
Z_{mij}: j^{th} observation on the m^{th} random regressor (m=1, 2,…, q) for the i^{th} subject;
β_l: fixed effect of the X_l regressor;
b_{i0}: random intercept for the i^{th} subject;
β_{im}: random effect of the Z_m random regressor for the i^{th} subject and
f: link function, e.g. identity; logit; log links are used for linear, logistic and log linear model, respectively.

Random Part

$$Y_{ij} \mid b_{0i}, b_{1i}, b_{2i}, \ldots, b_{qi} \sim \textit{either Normal, Bernoulli, or Poisson distribution};$$

$$\begin{pmatrix} b_{0i} \\ b_{1i} \\ \vdots \\ b_{qi} \end{pmatrix} \sim \textit{Normal}\ (\underline{0},\ \Sigma).$$

In the model, Xs are fixed regressors whose possible values or levels are only of interest. Gender, age and education levels are examples of fixed regressors. Zs are random factors or

regressors whose levels are from a random sample from a population of levels. Individuals enrolled in a repeated measures study, and families or households from which subjects are chosen for a study are examples of random factors. The random effects model in (2.1) is characterized by its fixed (β's) and random effects (b's). The fixed effects are estimated whereas for random effects, their variances and covariances are estimated. For the estimation, either maximum likelihood or restricted maximum likelihood approach is used [2]. Wald or likelihood ratio tests are used to tests hypotheses regarding β's. If the outcome, Y, is normally distributed, t-tests and F-tests can be used to tests β's. To test hypotheses regarding the random effects (b's) i.e., whether variances or covariances corresponding to random effects are zero, approximate likelihood ratio test or approximate Wald test or approximate mixture test can be used [2]. Random intercepts models are the simplest of random effects models. An example of a random intercept model is given below:

$$Y_{ij} = \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \ldots + \beta_p X_{pij} + b_{0i} + \varepsilon_{ij}; \quad \ldots (2.2)$$

$$\text{The error term, } \varepsilon_{ij} \sim Normal\left(0,\ \sigma^2\right);$$

$$Y_{ij} \mid\ b_{0i} \sim\ \ Normal\ \left(E\left(Y_{ij} \mid\ b_{0i}\right),\ \sigma^2\right);$$

$$b_{0i} \sim\ \ Normal\ \left(0,\ \tau^2\right).$$

Here, ε_{ij} and b_{0i} are independent of each other and

$$E[Y_{ij} \mid\ b_{0i}] = \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \ldots + \beta_p X_{pij} + b_{0i}.$$

2.3 Estimation of the RTM Effect for Non-normal Repeated Measures Data Using Simulation

To date, simulation was used to estimate the RTM effect for normally distributed repeated measures data. In this section, we propose a novel approach of estimating the RTM effect for non-normally distributed repeated measures data using simulation.

Suppose $Y_1=(y_{11},y_{12},...y_{1n})'$ and $Y_2=(y_{21},y_{22},...,y_{2n})'$ be the first and second set of observations of the outcome variable Y in a repeated measures study before implementing the intervention. Here, Y follows a skewed distribution with certain mean, variance, skewness and kurtosis. Also, let the correlation between Y_1 and Y_2 be r. We simulate bivariate skewed data that approximately distributed same as Y with the correlation r. The method is a combination of bootstrap sampling [69] from Y_1 and matrix decomposition [70], [71] of the correlation matrix,

$$R=\begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}.$$

The steps are listed below:

1. Compute $Z=\dfrac{Y_1-\overline{Y_1}}{S_{Y_1}}$, where $\overline{Y_1}$ and S_{Y_1} are mean and standard deviation of Y_1, respectively.
2. Simulate two independent bootstrap samples B_1 and B_2 of size N from Z.
3. Compute eigenvectors and eigenvalues of R. Let $U_1=(u_{11}, u_{12})'$ and $U_2=(u_{21}, u_{22})'$ be the eigenvectors of R with corresponding eigenvalues λ_1 and λ_2, respectively.
4. Obtain the bivariate skewed data using the following equation

$$\begin{bmatrix} Z_1' & Z_2' \end{bmatrix}=\begin{bmatrix} B_1 & B_2 \end{bmatrix}\left[\begin{bmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{bmatrix}\begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix}\right]'. \qquad \ldots(2.3)$$

5. Compute $Y_1' = \overline{Y}_1 + S_{Y_1} Z_1'$ and $Y_2' = \overline{Y}_1 + S_{Y_1} Z_2'$ which are the simulated bivariate 1st and 2nd set of skewed data with the correlation *r*. Both are approximately distributed same as *Y*.

Then, the RTM effect for the skewed data is estimated by selecting observations with the first measurement greater than a specified cut off value and calculating the difference between the means of first and second measurements in this group. This estimate is used in the adjustments of RTM effect in the final results from the multiple regression analysis.

2.4 Online Personalized Assessment-Feedback (PAF) Alcohol Intervention Study

A quasi-experimental study was carried out among a group of 251 graduating high school students in Alberta in 2007-2008. The students were recruited from seven high schools from a pool of 24 schools represented by eleven school boards. They were categorized into the intervention and the non-intervention arms of the study based on whether or not they accessed online personalized assessment-feedback (PAF), a brief intervention designed to reduce drinking among students. Students in the intervention group (n=109) accessed the PAF while students in the control group (n=142) did not. Each enrolled student was given a unique password to access the project website. They completed two baselines and three follow-up surveys by logging into the project website. Students completed the first baseline survey in June 2007 before graduation from high school. The second baseline survey was completed three months later in September 2007. The intervention was made available to the all participants in October 2007 by an email invitation. The intervention was a 'Personalized Drinking Profile' which consisted of 1) a normative feedback pie chart comparing the participant alcohol consumption with that of the reference population of same age and sex [72] and 2) a report on the participant's alcohol problem. A sample report is given in Appendix A.1. It can also be viewed at the following website available to the general public: http://www.CheckYourDrinking.net.

Information on age, gender, mobility, country of origin, marital status, living arrangements, school performance, alcohol drinking and parental support and involvement was collected at the two baselines. Based on the response at the second baseline the students were classified as problem drinkers using the Alcohol Use Disorders Identification Test (AUDIT) score. An AUDIT score is a validated 10-item self-report measure of hazardous and harmful drinking [73], [74], [75]. An AUDIT score of 8 or above for male and 6 or above for female were used to define problem drinkers [76]. The main outcome of interest was the total number of drinks in a typical week or the mean number of drinks per day in a typical week [77]. The mean number of drinks per day was obtained by dividing the total number of drinks reported in a typical week by the number of days reported.

2.5 Analytical Plan

In the PAF intervention study, the primary hypothesis was as follows: *Participants who, at baseline, met criteria for problem drinking and who subsequently accessed the intervention, would exhibit greater reductions in alcohol consumption compared to problem drinkers who did not access the intervention.*

Four groups were defined based on their baseline drinking problem (yes/no) and access to the intervention (accessed or not). Within-subject changes in alcohol drinking was studied for these groups. Average monthly changes in mean drinks/ day were obtained and compared between these groups using a random intercepts model after controlling for marital; living status; and like school score. For each group, the adjusted average change was obtained as a linear combination of the effect estimates (regression coefficients) from the model and tested using Z-test. The RTM effect was not examined and adjusted in testing the hypothesis. This may lead to incorrect conclusions considering the possibility of presence of the RTM effect in observed changes in the repeated measures data.

The main objective is to test the study hypothesis after adjusting for the RTM effects. The proposed new method of estimating and adjusting for the RTM effect for non-normal data

will be used in the analysis of PAF intervention study. This method is appropriate for analyzing the average monthly changes in mean drinks/ day which is skewed to the right.

There are different methods of adjusting for the RTM effects in the design and analysis stages of a study. A comparison of such methods is absent in current studies. In this thesis, a comparison of different methods of adjusting for the RTM effects will be performed. In the PAF study, two baseline data were collected. The first baseline was administered prior to graduation from high school, and the second baseline assessment was administered three months following graduation. It was anticipated that the transition had an impact on the alcohol drinking of the students. Two baseline assessments allowed us to examine two methods of adjusting the RTM effects at design stages. The study hypothesis will be tested applying the following five methods for adjustments of the RTM effects:

1. Use of average of multiple measurements for selecting subjects;
2. Taking two baseline measurements: first for selection and second for comparison;
3. Estimating the RTM effect using the formula in (1.1) by maximum likelihood method assuming bivariate normal distribution;
4. Estimating RTM effects for non-normal repeated measures data using the proposed novel simulation method and
5. Using Poisson regression for the adjustment of the RTM effect.

Methods 1, 2 and 3 were discussed in Chapter 1. Method 4 was described in Section 2.3. In the following section, Method 5 will be illustrated in the context of the study data.

Let,

Y_{ij} : Number of drinks in a typical week for the i^{th} problem drinkers at j^{th} time;

i = 1,2, …., n_i and j=-4,-1, 1, 3, 6 month;

E_{ij}: Expected number of drinks in a typical week for i^{th} problem drinkers at j^{th} time

= Average of number drinks of all subjects in the population at j^{th} time.

The outcome, relative number of drinks (R_{ij}), is defined as

R_{ij} = Number of drinks in a typical week / Expected number of drinks in a typical week

$= Y_{ij}/ E_{ij}$.

The Poisson regression model to account for the RTM effect is

$$\log\ E[Y_{ij}] = \log \mu_{ij} = \log\ \hat{E}_{ij} + \sum_K a_K Z_{ijk} + X_j(\beta_{0i},\ \beta_{1i},\ \beta_{2i})^T, \qquad \ldots(2.4)$$

where

$Z_{ijk} = ij^{\text{th}}$ observation of k^{th} covariate;

$$X = \text{Design matrix} = \begin{bmatrix} 1 & -4 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 3 \\ 1 & 0 & 6 \end{bmatrix};$$

β_{1i}: slope before the intervention for the i^{th} subject;

β_{2i}: slope after the intervention for the i^{th} subject.

One term Taylor expansion gives $e^{\beta} \approx 1 + \beta$ assuming all higher order terms $\frac{\beta^r}{r!} \cong 0;\ r \geq 2$.

Thus, β_1 is the change in relative number of drinks per month in 4 months before the intervention among the problem drinkers. Similarly, β_2 is the change in relative number of drinks per month in 6 months after the intervention in problem drinkers. To asses the impact of intervention among problem drinkers accounting for the RTM effect, the following regression model will be fitted:

$$\hat{\beta}_{2i} = \Delta + \rho\hat{\beta}_{1i} + \varepsilon . \qquad \ldots (2.5)$$

The estimated intercept $\hat{\Delta}$ is the change in relative number of drinks after the intervention among problem drinkers controlling the RTM effect. The hypothesis regarding Δ can be tested using t-tests.

Chapter 3

Results

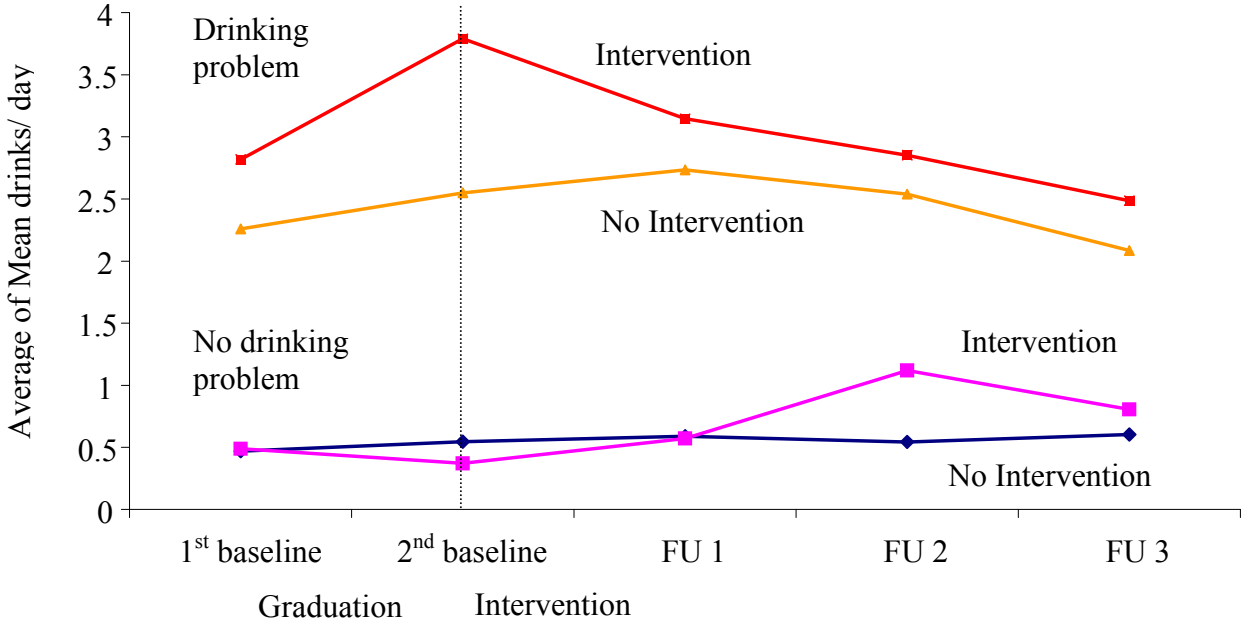
In this chapter, the proposed new method of estimating and adjusting RTM effects for non-normal data will be applied to study changes in alcohol consumption in the PAF intervention study. Four other methods will also be applied: use of multiple measurement for selecting subjects; two baseline measurements: first for selection and second for comparison; estimation of the RTM effect using maximum likelihood method for normal data; and the use of Poisson regression for the adjustment of the RTM effect for count data. At first, the results obtained from the random intercepts model (2.2) without adjusting the RTM effects will be discussed briefly. Then, the results obtained from each proposed method of adjusting the RTM effect will be described and compared with the results obtained without adjusting the effects. The results of different methods of adjusting the RTM effect will be compared at the end.

3.1 Results without Adjusting for the RTM Effects

In PAF study, the intervention was made accessible to the students after one month of the second base line data collection. Of the 251 study participants, 109 (43.4%) accessed the intervention. The impact of the intervention was assessed by studying changes in the alcohol consumption compared to the second baseline data. At the second base line, 79 students were problem drinkers of which 44 (55.7%) accessed the intervention. Among 132 non-problem drinkers, 62 (47.0%) accessed the intervention. In Figure 3.1, mean values of drinks/day were plotted for the intervention and the non-intervention students stratified by their baseline drinking status. Among participants who initially met criteria for problem drinking, those who accessed the intervention had higher intake of alcohol compared to those who did not seek the intervention (3.8 vs. 2.6 drinks/ day, p-value=0.09) at the second baseline with the difference not being statistically significant. A much steeper decrease in the alcohol intake over time was

observed in that group. For the students without the drinking problem, the average drinks were 0.4 and 0.5 per day for the intervention and the non-intervention group at the second baseline and it remained similar during the follow-ups without much change.

Figure 3.1. Average drinks per day in a typical week for the study groups of interest



Within subjects' changes in the alcohol consumption in the four study groups were studied using a random intercepts model of the following form:

$$(Mean\ drink/\ day)_{ij} = \beta_0 + b_{0i} + \beta_1\ Intervention_i + \beta_2\ Problem\ drinking_i + \beta_3\ Time_{ij} + \beta_4\ Intervention_i \times Problem\ drinking_i + \beta_5\ Intervention_i \times Time_{ij} + \beta_6\ Problem\ drinking_i \times Time_{ij} + \beta_7\ Intervention_i \times Problem\ drinking_i \times Time_{ij} + \alpha_1\ living\ arrangement_i + \alpha_2\ school\ liking\ score_i + \epsilon_{ij}, \quad (3.1)$$

where j^{th} observation on i^{th} student was represented by ij subscript and living arrangement and school liking score are the two confounders. For participants in each of these four a priori groups of interest, average adjusted monthly changes in mean drinks/ day was estimated as a linear combination of the effect estimates (coefficients) of the above model and tested using Z-tests. Statistical significance was set at the p = 0.05 level, and confidence intervals were set at 95%. The overall model was significant (Wald χ^2 [df = 10, n_1 = 208 students, n_2=755

observations] = 124.9, p < 0.0001). A three-way interaction between baseline problem drinking status, exposure to the intervention and follow-up time in months was significant (p=0.034, Table 3.1). It indicated that within-subject changes in the mean drinks varied between the study groups from the second baseline to the end of

Table 3.1. Random effects model predicting within-subjects changes in mean number of drinks per day in a typical week

Variable	Adjusted Coefficient (95% CI)	p-value
Main effects – Covariates		
Living arrangements		
Both parents	-	-
Single parent	-0.03 (-0.65, 0.59)	0.92
Others	0.66 (0.02, 1.31)	0.045
Like school score	-0.20 (-0.48, 0.08)	0.16
Main effects - Predictors		
Problem drinking	2.01 (1.25, 2.78)	< 0.001
Intervention	0.02 (-0.63, 0.67)	0.96
Time (month)	0.02 (-0.06, 0.09)	0.69
Interactions		
Problem drinking X intervention	1.06 (0.01, 2.11)	0.047
Problem drinking X Time	-0.11 (-0.24, 0.02)	0.11
Intervention X Time	0.07 (-0.04, 0.18)	0.19
Problem drinking X Intervention X Time	-0.19 (-0.37, -0.01)	0.034
Constant	0.88 (0.12, 1.65)	0.023
Variance components	Estimate (SE)	95% CI
Between subjects variance	2.20 (0.26)	(1.74, 2.78)
Residual variance	1.67 (0.10)	(1.48, 1.88)

follow-up. Among the problem drinkers who accessed the intervention, the adjusted average monthly change (95% CI) in mean drink was -0.21 (-0.30, -0.12) (Table 3.2). The decrease was highly significant (p-value < 0.001). A non-significant decrease was observed in the

problem drinkers without the intervention [mean change (95% CI): -0.09 (-0.20, -0.02); p=0.096]. These average changes in mean drinks between the two groups did not vary (p=0.10) although initial average drinks was much higher among the students who accessed the intervention. In the student without any drinking problem, the mean drinks increased among those who accessed the intervention [mean change (95% CI): 0.09 (0.01, 0.16); p=0.025]. A non-significant average increase [mean change (95% CI): 0.02 (-0.06, 0.09); p=0.689] in mean drinks was also observed among students who did not access the intervention.

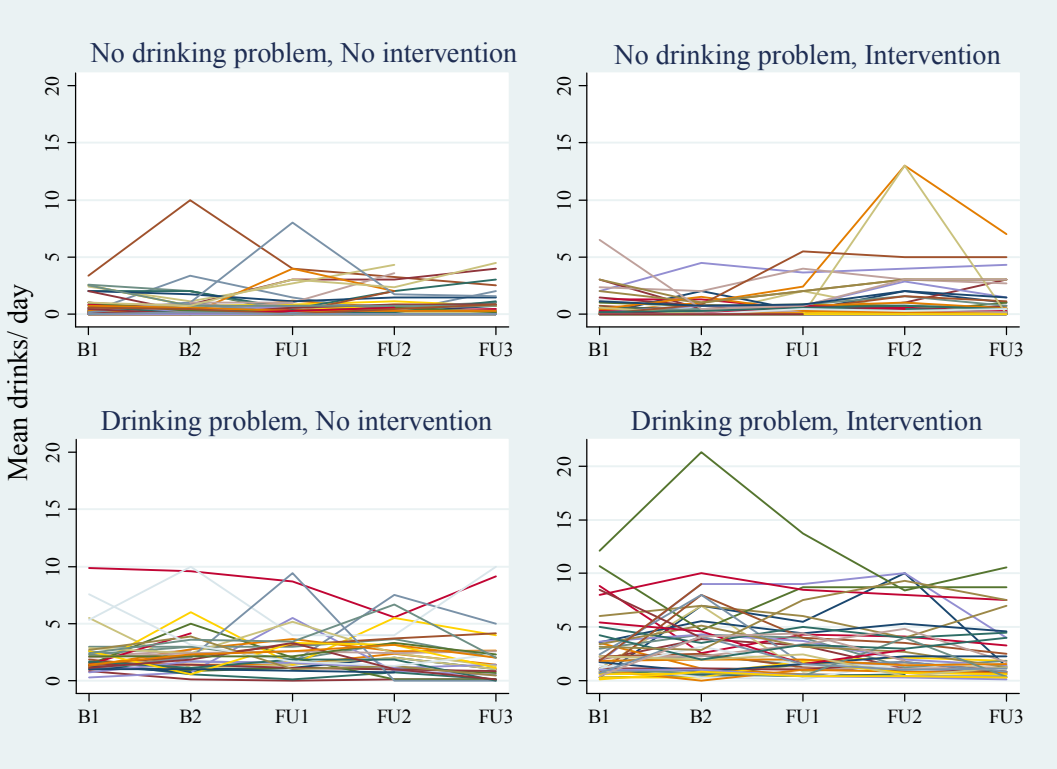
Table 3.2. Average adjusted monthly change in mean drinks/ day in a typical week from a random effects model of drinks/ day in a typical week

Baseline drinking problem	Accessed intervention	Crude mean drinks/ day in a typical week at the second baseline	Average adjusted monthly change (Δ) in mean drinks/day (95% CI)	p-value
No	No	0.55	0.02 (-0.06, 0.09)	0.689
	Yes	0.37	0.09 (0.01, 0.16)	0.025
Yes	No	2.55	-0.09 (-0.20, 0.02)	0.096
	Yes	3.79	-0.21 (-0.30, -0.12)	< 0.001

3.2 Presence of the RTM Effects in the Data

Decrease in alcohol consumptions among problem drinkers not exposed to the intervention and increases in the consumption of non-problem drinkers indicate possible presence of the RTM effects in the repeated measures data [78]. Each participant's alcohol consumption was plotted to detect possible RTM effects by the study groups (Figure 3.2). There were instances where high alcohol intakes were immediately followed by decreased intakes on the repeat measurements on the follow-ups. Also, there were some outlying observations. Each student's alcohol consumption was plotted after removing the outliers in Figure 3.3. The presence of the RTM effects became more notable in Figure 3.3.

Figure 3.2. Average drinks per day in a typical week for each subject in the study groups of interest



The average adjusted monthly changes in mean drinks per day for the study groups with and without removing the outlying observations were presented in Table 3.3. In problem drinkers who accessed intervention, average adjusted mean drinks decreased from 0.21 to 0.17. Also, in non-problem drinkers who accessed the intervention, the average adjusted mean drinks decreased from 0.09 to 0.07. In the subsequent analyses, the outliers were removed from the data and the results obtained adjusting for the RTM effects were presented.

Figure 3.3. Average drinks per day in a typical week for each subject in the study groups of interest without outliers

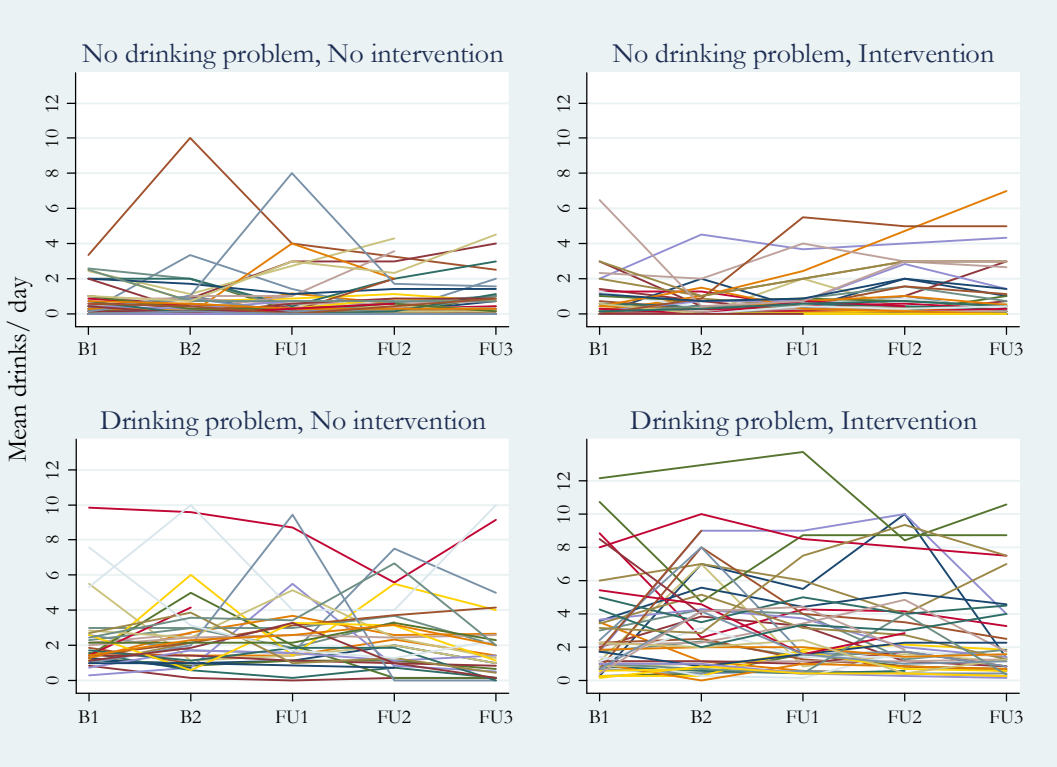


Table 3.3. Average adjusted monthly change in mean drinks/ day in a typical week from a random effects model with and without removing the outliers

Baseline drinking problem	Accessed intervention	Without removing outliers		After removing outliers	
		Δ (95% CI)	p-value	Δ (95% CI)	p-value
No	No	0.02 (-0.06, 0.09)	0.689	0.02 (-0.05, 0.08)	0.63
	Yes	0.09 (0.01, 0.16)	0.025	0.07 (0.004, 0.13)	0.038
Yes	No	-0.09 (-0.20, 0.02)	0.096	-0.09 (-0.18, -0.001)	0.049
	Yes	-0.21 (-0.30, -0.12)	<0.001	-0.17 (-0.24, -0.09)	< 0.001

3.3 Results after adjusting for the RTM Effects

3.3.1. Use of Average of Two Measurements for Selection of Subjects (Method 1)

The average of the first and second baseline AUDIT scores was calculated for each student. This average was used to define problem drinkers. An average AUDIT scores of 8 or above for male and 6 or above for female were considered as problem drinkers [74]. In addition, the average of the first and second baselines' mean drinks/ day was estimated for each student. Changes in mean drinks/ day during the follow-ups compared to the baselines' average mean drinks/ day were studied using model (3.1). This approach minimized the RTM effects by reducing variability; measurement errors in the initial measurements by means of averages, but couldn't remove the effects. In the problem drinkers without the intervention, the mean reduction slightly decreased to -0.08 [95% CI: (-0.16, 0.002)] from -0.09 [95% CI: (-0.18, -0.001)] (Table 3.4) after adjusting for the RTM effects. Similarly, in the non-problem drinkers with the intervention, the average increase in the mean drinks reduced to 0.05 [95% CI: (-0.01, 0.11)] from 0.07 [95% CI: (0.004, 0.13)] (Table 3.4).

Table 3.4. Average adjusted monthly change in mean drinks/ day in a typical week from a random effects model with and without controlling the RTM effects using Method 1

Baseline drinking problem	Accessed intervention	Without controlling the RTM effects		With controlling the RTM effects using Method 1	
		Δ (95% CI)	p-value	Δ (95% CI)	p-value
No	No	0.02 (-0.05, 0.08)	0.633	0.03 (-0.03, 0.09)	0.34
	Yes	0.07 (0.004, 0.13)	0.038	0.05 (-0.01, 0.11)	0.095
Yes	No	-0.09 (-0.18, -0.001)	0.049	-0.08 (-0.16, 0.002)	0.056
	Yes	-0.17 (-0.24, -0.09)	< 0.001	-0.11 (-0.18, -0.04)	0.004

These marginally significant changes in the mean drinks: the decrease (p-value=0.056, Table 3.4) among the problem drinkers without the intervention and the increase (p-value=0.095,

Table 3.4) among the non-problem drinkers with the intervention indicated the presence of the RTM effects even after using Method 1 to control for it.

3.3.2. Two Baseline Measurements: First for selection and second for comparison (Method 2)

In this approach, students were classified as problem and non-problem drinkers based on the first baseline AUDIT score. A male student with the first baseline AUDIT score greater or equal to 8 and a female student with the score greater or equal to 6 were considered as problem drinkers [74]. Then, the changes in the mean drinks/ day from the second baseline were observed during the follow-ups in the four study groups

Table 3.5. Average adjusted monthly change in mean drinks/ day in a typical week from a random effects model with and without controlling the RTM effects using Method 2

Baseline drinking problem	Accessed intervention	Without controlling the RTM effects		With controlling the RTM effects using Method 2	
		Δ (95% CI)	p-value	Δ (95% CI)	p-value
No	No	0.02 (-0.05, 0.08)	0.633	0.03 (-0.04, 0.09)	0.42
	Yes	0.07 (0.004, 0.13)	0.038	0.02 (-0.05, 0.08)	0.62
Yes	No	-0.09 (-0.18, -0.001)	0.049	-0.09 (-0.18, -0.01)	0.029
	Yes	-0.17 (-0.24, -0.09)	< 0.001	-0.10 (-0.18, -0.02)	0.012

using model (3.1) (Table 3.5). Among the problem drinkers who did not access the intervention, an average decrease of -0.09 [95% CI: (-0.18, -0.01), p-value=0.029] in the mean drinks was observed after using Method 2 to control for the RTM effects (Table 3.6). The decrease remained same as compared to the decrease of -0.09 [95% CI: (-0.18, -0.001), p-value=0.049] observed without adjusting for the effect. Similarly, the observed increase in mean drinks among the non-problem drinkers with or without access to the intervention after using Method 2, though statistically non-significant, indicated the presence of the RTM effects. The method failed to remove the RTM effects from the data because the correlation

between the first and second measurement was different from the correlations between the first and each of the follow-ups measurement (Table 3.6).

Table 3.6. Correlations between mean drinks in a typical week

	Baseline 1	Baseline 2	FU 1	FU 2	FU 3
Baseline 1	1	0.599	0.656	0.578	0.734
Baseline 2		1	0.680	0.665	0.645
FU 1			1	0.726	0.748
FU 2				1	0.773
FU 3					1

3.3.3. Estimation of RTM Using Maximum Likelihood (Method 3)

In the study participants, the average mean drinks/ day were 1.35 and 1.45 with standard deviations (SDs) 1.95 and 2.16 at the first and second baselines, respectively (Table 3.7). The second baseline data were collected after the graduation. The alcohol intake was increased among the students in the transitional phase after the graduation. The first baseline data represented typical alcohol consumption among the students. The mean and SD of the first baseline data were considered for the estimation of the RTM effects. It was assumed that in the study participants, mean drinks/ day followed a normal distribution with mean (μ) 1.35 and SD (σ) of 1.95. The correlation between mean drinks in repeated measurements ranged from 0.6 to 0.8 (Table 3.6). The correlation was assumed to be 0.7. The RTM effects was estimated using formula (1.1) for mean drinks greater than 2 (75^{th} percentile, Table 3.7) and for the mean drinks equal to 0 (25^{th} percentile, Table 3.7). The estimation procedures were described below.

Table 3.7. Average drinks/ day in a typical week

						Percentiles			Coeff. of
	N	Min.	Max.	Mean	SD	25	50	75	Skewness
Baseline 1	245	0	12.00	**1.35**	**1.95**	0	0.71	2.00	2.61
Baseline 2	208	0	10.00	1.42	2.16	0	0.57	2.00	2.26
FU 1	203	0	13.71	1.45	2.14	0	0.57	2.00	2.43
FU 2	195	0	10.00	1.39	1.96	0	0.57	2.00	2.19
FU 3	182	0	10.57	1.32	2.01	0	0.54	1.57	2.42

The estimation of the RTM effects for mean drinks/ day > 2

Here, $z = (k- \mu)/ \sigma = (2.00-1.35)/ 1.95 = 0.33$;

c = ordinate of standard normal distribution at 0.33 / area under the standard normal curve > 0.33

$= 0.38/ 0.37 = 1.03$.

Therefore, the estimated RTM effect = $c\sigma(1 - \rho) = 1.03*1.95*(1-0.70) = 0.60$.

The estimation of the RTM effects for mean drinks/ day = 0

Here, $z = (\mu-k)/ \sigma = (1.35-0)/ 1.95 = 0.69$;

c = ordinate of standard normal distribution at 0.69 / area under the standard normal curve > 0.69

$= 0.31/ 0.25 = 1.24$.

Therefore, the estimated RTM effect = $c\sigma(1 - \rho) = 1.24*1.95*(1-0.70) = 0.73$.

The intervention effect was assessed from the second baseline. For each student with mean drinks greater than 2 at the second baseline, the estimated RTM effect, 0.60, was subtracted from their mean drinks. Conversely, for each student with mean drinks equal to 0 at the second baseline, the estimated RTM effect, 0.73, was added to their mean drinks. The

students were classified as problem and non-problem drinkers based on their second baseline AUDIT score. A male student with the second baseline AUDIT score greater or equal to 8 and a female student with the score greater or equal to 6 were considered as problem drinkers [74]. In the updated data, changes in mean drinks during the follow-ups compared to the second baseline were studied using model (3.1). A significant reduction of -0.11 [95% CI: (-0.19, -0.04), p-value=0.004] in mean drinks was observed among the problem drinkers who accessed the intervention (Table 3.8). The reduction was much less compared to the reduction of -0.17 [95% CI: (-0.24, -0.09), p-value < 0.001] in that group without the adjustment of the RTM effects. Among the problem drinkers who did not access the intervention, the observed reduction in mean drinks was reduced to -0.04 [95% CI: (-0.13, 0.05), p-value=0.38] from -0.09 [(95% CI: (-0.18, -0.001), p-value=0.049] after controlling for

Table 3.8. Average adjusted monthly change in mean drinks/ day in a typical week from a random effects model with and without controlling the RTM effects using Method 3

Baseline Drinking Problem	Accessed intervention	Without controlling the RTM effects		With controlling the RTM effects using Method 3	
		Δ (95% CI)	p-value	Δ (95% CI)	p-value
No	No	0.02 (-0.05, 0.08)	0.633	-0.04 (-0.10, 0.02)	0.220
	Yes	0.07 (0.004, 0.13)	0.038	-0.002 (-0.07, 0.06)	0.959
Yes	No	-0.09 (-0.18, -0.001)	0.049	-0.04 (-0.13, 0.05)	0.375
	Yes	-0.17 (-0.24, -0.09)	< 0.001	-0.11 (-0.19, -0.04)	0.004

the RTM effects. In the non-problem drinkers, the changes in mean drinks in the opposite direction was observed after controlling the RTM effects. In this group who accessed the intervention, the observed significant increase of 0.07 [95% CI: (0.004, 0.13), p-value=0.038] in mean drinks was changed to a non-significant decrease of -0.002 [95% CI: (-0.07, 0.06), p-value=0.959] after adjusting for the RTM effects. For the non-problem drinkers without the intervention, an increase of 0.02 [95% CI: (-0.05, 0.08), p-value=0.633] in mean drinks was changed to a decrease of -0.04 [95% CI: (-0.10, 0.02), p-value=0.959] in the consumption.

This method eliminated the RTM effects from the observed changes. However, the method based on the assumption of normality gave bias results because of non-normality in mean drinks. In the study population, mean drinks/ day followed a positively skewed distribution. In the next section, the adjustment for the RTM effects using the new simulation-based method for non-normal repeated measures data were discussed.

3.3.4. Estimation of RTM Assuming Non-normal Distribution (Method 4)

In the study population, mean drinks/ day in a typical week was positively skewed (Table 3.7). At the first baseline, the average mean drinks/ day was 1.35 with SD 1.95 and coefficient of skewness 2.61. The observed correlations in mean drinks/ day between the baselines and follow up measurements ranged from 0.6 to 0.8 (Table 3.6). A correlation of 0.7 between the repeated measurements was considered. Bivariate skewed data following the steps outlined in section 2.3 were simulated. The simulated

Table 3.9. Descriptive statistics for observed first baseline and the simulated data

	Mean drinks/ day at	Simulated data	
Statistic	the first baseline	1st measurement	2nd measurement
N	245	2450000	2450000
Mean	1.35	1.35	1.35
SD	1.95	1.94	1.95
Skewness	2.61	2.00	2.21

data were approximately distributed same as the original study population data (Table 3.9). The correlation between the first and second set of simulated data was 0.7. The RTM effect for mean drinks greater than 2 (75^{th} percentile, Table 3.7) was estimated by selecting the simulated observations with the first measurement greater than 2 and then calculating the difference between the means of first and second measurements in this group. Similarly, the RTM effect for mean drinks equal to 0 was estimated. The estimation procedures were described below.

The estimation of the RTM effects for mean drinks/ day > 2

The RTM effect = [Mean of 1st measurements | 1st measurement > 2] – [Mean of 2nd measurements | 1st measurement > 2]

= 3.85 - 3.41

= 0.44

The estimation of the RTM effects for mean drinks/ day = 0

The RTM effect = [Mean of 1st measurements | 1st measurement = 0] – [Mean of 2nd measurements | 1st measurement = 0]

= 0.02 – 0.87

= - 0.85

For each student with mean drinks greater than 2 at the second baseline, the estimated RTM effect, 0.44, was subtracted from their mean drinks. Conversely, for each student with mean drinks equal to 0 at the second baseline, the estimated RTM effect, 0.85, was added to their mean drinks. The students were classified as problem and non-problem drinkers based on their second baseline AUDIT score [74]. Then, the changes in mean drinks in the four study groups during the follow-ups compared to the second baseline were studied using model (3.1). Among the problem drinkers who accessed the intervention, a significant reduction of -0.13 [95% CI: (-0.21, -0.05), p-value=0.001] in mean drinks was observed (Table 3.10). As observed in method 3, here as well, the reductions in mean drinks were reduced for the both groups of the problem drinkers (with or without access to the intervention) after controlling for the RTM effects. Also, in the both groups of the non-problem drinkers, the increments in mean drinks during the follow-ups were changed to decrements after adjusting for the effects (Table 3.10). This method estimated the RTM effects considering the actual distribution of the data, thus provided more accurate estimation of the effects. The method eliminated the RTM effects from the observed changes and provided a valid interpretation of the changes.

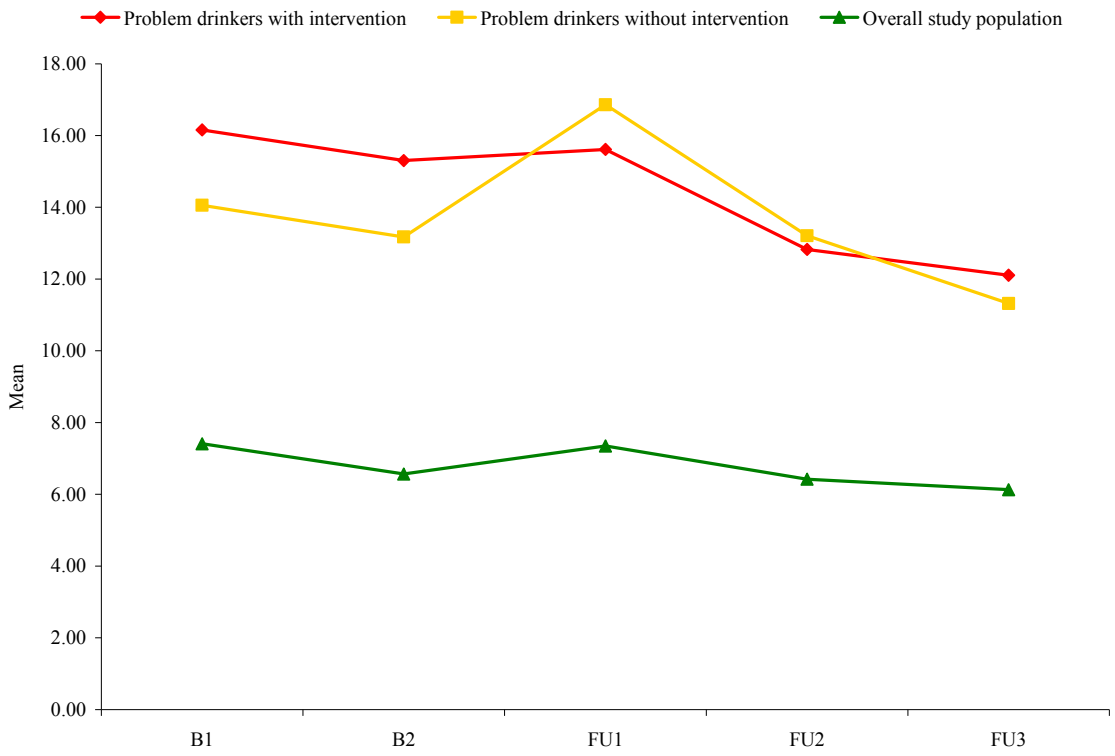
Table 3.10. Average adjusted monthly change in mean drinks/day in a typical week from a random effects model with and without controlling the RTM effects using Method 4

Baseline Drinking Problem	Accessed intervention	Without controlling the RTM effects Δ (95% CI)	p-value	With controlling the RTM effects using Method 4 Δ (95% CI)	p-value
No	No	0.02 (-0.05, 0.08)	0.633	-0.05 (-0.11, 0.01)	0.13
	Yes	0.07 (0.004, 0.13)	0.038	-0.01 (-0.08, 0.05)	0.68
Yes	No	-0.09 (-0.18, -0.001)	0.049	-0.06 (-0.15, 0.04)	0.23
	Yes	-0.17 (-0.24, -0.09)	< 0.001	-0.13 (-0.21, -0.05)	0.001

3.3.5. Adjusting for RTM Using Poisson Regression (Method 5)

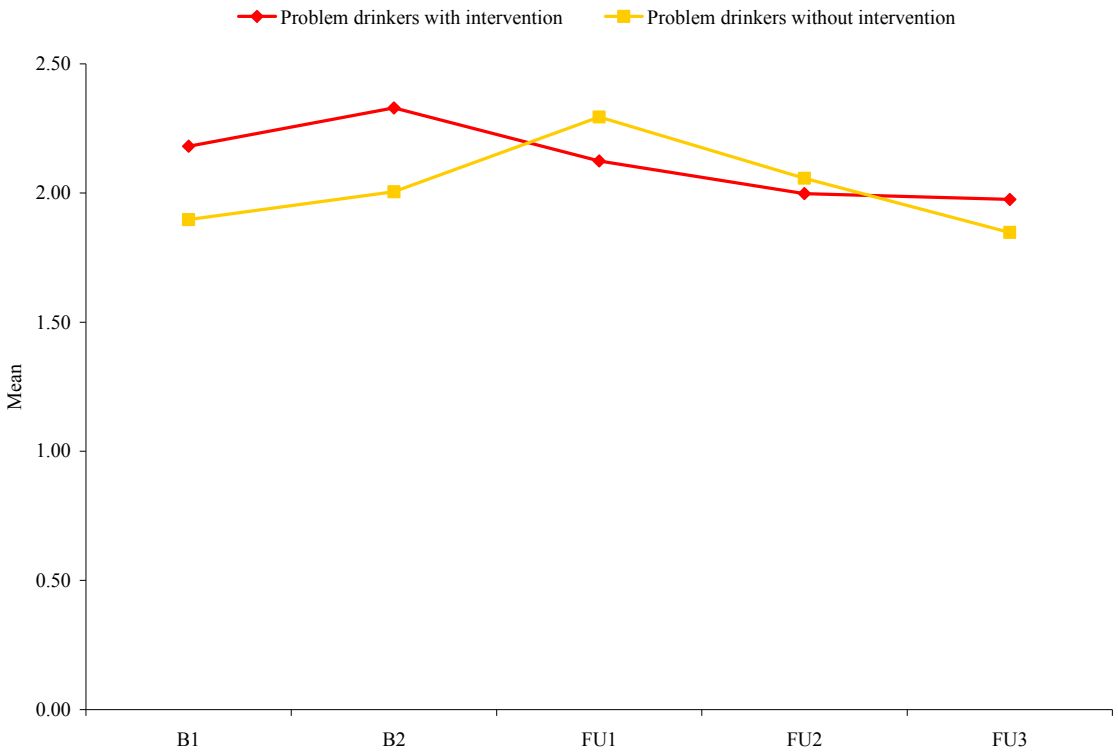
In this section, Poisson regression model was used to adjust the RTM effects in the total number of drinks in a typical week. In Figure 3.4, average total number of drinks was presented for the problem drinkers with and without the intervention and for the overall study population. A general decreasing trend in number of drinks over the study period was observed whether the intervention was accessed or not. To remove the general trend, relative number of drinks was calculated. For each student, the relative number of drinks was defined as the ratio of the total number of drinks in a typical week divided by the expected number of drinks in that week. The expected number of drinks in a typical week at a time point was estimated by the average of number drinks of all subjects in the population at that time.

Figure 3.4. Number of drinks in a typical week among the problem drinkers with and without the intervention and in the overall population



In Figure 3.5, average relative number of drinks was plotted for the problem drinkers with and without the intervention. An increase in the average relative number of drinks was immediately followed by a sharp decrease in the average consumption even without the intervention (Figure 3.5). This indicated the presence of the RTM effects in the relative number of drinks for the problem drinkers.

Figure 3.5. Relative number of drinks in a typical week among the problem drinkers with and without the intervention



For each of the problem drinkers who accessed the intervention, the estimated slopes for relative number of drinks for four month before and six months after the intervention were obtained by fitting the Poisson regression model (2.4). The RTM effects in the slopes of relative number of drinks was adjusted through regressing the post intervention slopes on the pre intervention slopes. The test of the intercept from the simple linear regression model compared the before-after change in the relative number of drinks adjusting for the RTM effects. Similarly, the RTM effects in the relative number of drinks was controlled for the problem drinkers without the intervention. For the non-problem drinkers, the Poisson model (2.4) did not converge. The results obtained with forced finite iterations produced unreliable estimates; thus not reported here. Among the problem drinkers who accessed the intervention, 6% per month increase in the relative number of drinks was observed during 4-month baselines period prior accessing the intervention (Table 3.11). After accessing the intervention, the relative number of drinks decreased by 6% per month in 6-months follow-up period (Table

3.11). Thus, 12% (p-value=0.017) reduction in the relative number of drinks in the follow-ups compared to the baseline was observed. After accounting for the RTM effects, the reduction reduced to 4% (p-value=0.034). Similarly, among the problem drinkers without the intervention, 10% (p-value=0.197) reduction in the relative number of drinks reduced to 4% (p-value=0.085) after adjusting for the RTM effects.

Table 3.11. Average adjusted monthly change in the relative number of drinks in a typical week among the problem drinkers with and without controlling the RTM effects using Poisson regression (Method 5)

	Average change/ month (95% CI) in relative no. of drinks		Ignoring the RTM effect		Controlling for the RTM effect	
	Before intervention	After intervention	$\Delta_{After} - \Delta_{Before}$ (95% CI)	p-value	$\Delta_{After} - \Delta_{Before}$ (95% CI)	p-value
Problem drinkers, accessed intervention (n=44)	0.06 (-0.01, 0.12)	-0.06 (-0.10, -0.02)	-0.12 (-0.21,-0.02)	0.02	-0.04 (-0.08,-0.003)	0.03
Problem drinkers, did not access intervention (n=35)	0.04 (-0.06, 0.14)	-0.06 (-0.13, 0.01)	-0.10 (-0.26,0.06)	0.20	-0.04 (-0.09, 0.01)	0.09

Chapter 4

Discussion and Conclusions

4.1 Discussion

In repeated measures data, the RTM effect is always present due to less than perfect correlation (correlation coefficient < 1) between the repeated measurements. In studying changes in outcome variables over time, the RTM effect should be adjusted for the valid interpretation of the changes. It can be adjusted at the design or at the analysis stage of a repeated measures study. At the analysis stage, the effect can be estimated and removed from the observed changes or it can be adjusted using a regression model of the latter observations on the previous observations. A new method of estimating RTM effects for non-normal data using simulation (Method 4) was proposed. The method was applied to adjust for the RTM effects in studying changes in mean drinks in a typical week in PAF intervention study. The application of the new method in the study data was appropriate as the mean drinks followed a positively skewed distribution. After removing the estimated RTM effects from the data, a random intercepts model was used to study changes in the outcome. To adjust for the RTM effects in the outcome, three other methods were used: use of average of two measurements for selection of subject (Method 1); taking two baseline measurements: first for selection of subjects and second for baseline assessment (Method 2); and estimating the RTM effect using the maximum likelihood method assuming bivariate normal distribution (Method 3). The RTM effects in total number of drinks in a typical week was also studied. Poisson regression was used to adjust for the RTM effect in total number of drinks (Method 5). In this chapter, the strengths and weaknesses of each method in adjusting the RTM effects will be discussed.

In the mean drinks in a typical week, the presence of the RTM effects was evident at individual as well as group levels. At individual levels, there were instances where a high

alcohol intake was immediately followed by a decreased intake on the next follow-up measurement. At group levels, among the problem drinkers who accessed the intervention, a sudden sharp increase in the average mean drinks at the second baseline was immediately followed by a steep decrease in the average intake at the first follow-up. In the problem drinkers who did not access the intervention, initial rises in the mean drinks were followed by a steady decrease in the consumption during the follow-ups. Among the non-problem drinkers who accessed the intervention, the initial low intakes were followed by the increased intakes. In the average adjusted (adjusted for living arrangements and school liking scores) monthly change in the alcohol consumption without controlling for the RTM effects, a significant decrease in the mean drinks in the problem drinkers without the intervention and a significant increase in the outcome in the non-problem drinkers with the intervention indicated the presence of the RTM effects in the data. The estimated average changes in the mean drinks for the four study groups were biased for not adjusting for the RTM effects.

To adjust for the RTM effects at the design stage, Methods 1 and 2 were used. In Method 1, the use of the averages of the two baselines' AUDIT scores to define problem and non-problem drinkers; the baselines' averages of the outcome reduced the variability in the data. Despite reduction in variability, marginally significant changes in the mean drinks: a decrease in the problem drinkers without the intervention and an increase in the non-problem drinkers with the intervention showed that the RTM effects were still present in the data. Method 2 failed to eliminate the RTM effects from the data due to varying correlations between baselines and follow-ups measurements of the outcome. After applying Method 2, a significant average decrease in the mean drinks in the problem drinkers without the intervention and the average increase in mean drinks (statistically insignificant) among the non-problem drinkers with or without the intervention were observed. After using Method 3 to control for the RTM effects, a significant reduction in mean drinks was observed among the problem drinkers who accessed the intervention. The average reduction in the mean drinks among the problem drinkers without the intervention became insignificant. In the non-problem drinkers, the changes in mean drinks in the opposite direction was observed after controlling the RTM effects. In the both groups of non-problem drinkers, reductions in mean drinks, though statistically insignificant, were observed. The method eliminated the RTM effects from

the observed changes. The method based on the normality assumption gave bias estimates of the average changes in the study groups due to substantial non-normality (positive skewness) in the outcome, mean drinks. In the proposed simulation based approach, the RTM effects were estimated considering the true distribution (positively skewed) of the outcome, thus provided more accurate estimation of the effects. After applying the method, a significant reduction in the mean drinks was observed among the problem drinkers who accessed the intervention. In the rest of the study groups, insignificant reductions in the mean drinks were observed. The method eliminated the RTM effects from the observed changes in the outcome. The method ensured valid interpretation of the observed changes in the outcome by providing the most accurate estimation of the RTM effect and then removing it from the data. In Method 5, Poisson regression model (2.4) was used to adjust for the RTM effects in total number of drinks in a typical week. A reduction in the relative number of drinks in the follow-ups compared to the baseline was observed among the problem drinkers with the intervention. Among the problem drinkers without the intervention, a non-significant reduction in the relative number of drinks in the follow-ups compared to the baseline was observed. The model did not converge for the both groups of non-problem drinkers. The forced finite iterations produced unreliable results for the groups. The non-convergence issue was a potential limitation for Method 5.

The RTM effect is a persistent phenomenon in alcohol intervention studies, however it has been largely ignored in previous studies. In few recent studies, the presence of the RTM effects in alcohol consumption was addressed, but none of the studies used appropriate method to control for it. Despite non-normality in alcohol intake data, the RTM effect was estimated assuming normal distribution [16], [61] or its presence was detected using graphs without applying any method to adjust for it [34]. The proposed simulation based approach estimated the RTM effects considering the true distribution (positively skewed) of the alcohol intake thus provided more precise estimation and adjustment for the effect.

4.2 Conclusions

In repeated measures data, large or small values at the initial measurement tend to be followed by values that are closer to the mean at the follow-ups measurements. This tendency is called regression to the mean (RTM). The presence of the RTM effect is inevitable in repeated measures data because of less than perfect correlation (correlation coefficient < 1) between the repeated measurements. It is established as a statistical concept based on theoretical deductions. Despite the growing evidence of the presence of RTM effects in clinical and public health studies based on repeated measures data, very few studies evaluated and considered the effects when interpreting the observed changes over time. An RTM effect is mixed with an intervention effect in observed changes over time. It is extremely important to separate the RTM effect from the observed change in order to isolate any intervention effect and to make valid evaluation of the intervention. In presence of the RTM effects, the estimation of the average changes in the study groups would be erroneous. It can lead to the wrong interpretation of the effectiveness of an intervention if not removed from the data. Any intervention may appear successful due to RTM. It may result in wasteful persuasion of an ineffective intervention. It may have significant implications for patient care, policy development. In studying changes in outcome variables in repeated measures studies, RTM effects should always be adjusted for the valid interpretation of the changes and unbiased assessment of the intervention effects. It can be adjusted at the design or at the analysis stage of a study. At the design stage, the best method to eliminate the RTM effect is the use of a parallel control group. The use of a control group may not always be feasible, ethical especially in the observational studies. The use of average of multiple measurements for selection of subjects into the study may reduce the effects of RTM by reducing the variability in the data. But this method cannot guarantee the removal of the effects from the data. The method of taking two baseline measurements: first for selection of subjects and second for baseline assessment completely eliminates the RTM effect if correlations between the first and successive measurements are exactly the same which is highly unlikely in practice. At the analysis stage, the effect can be estimated and then subtracted from the observed changes or it can be adjusted using a regression model of the latter observations on the previous observations. The choice of the methods to control for the RTM effect should be based on the

type (continuous, count) and shape (normal, non-normal distribution) of the outcome variables of interest. In case of normally distributed outcome variables, the estimation of the RTM effect using the maximum likelihood (ML) method assuming bivariate normal distribution is more appropriate for its efficiency, adaptability to different models (additive or multiplicative) and different types of sampling (truncated, selected, censored or complete). For non-normal data, simulation based or non-parametric methods (local sample means; kernel density estimation) can be used to estimate the RTM effect. A new method of estimating RTM effects for non-normal data using simulation is proposed. The method is a combination of bootstrap sampling from the standardized outcome variable and matrix decomposition of the correlation matrix between the repeated measurements. The method was applied to adjust for the RTM effects in studying changes in mean drinks in a typical week in PAF intervention study. In the study, mean drinks followed a positively skewed distribution. The new method was more accurate in adjusting the RTM effects compared to other methods considered here. For count data, Poisson regression based approach can be used to adjust for the RTM effect.

Bibliography

1. Singer, J. D. and Willett, J. B. (2003). *Applied repeated measures data analysis: modeling change and event occurrence*. New York, NY: Oxford University Press.
2. Kleinbaum, D. G.; Kupper, L. L.; Nizam, and Muller, K. E. (2008). *Applied regression analysis and other multivariable methods*, (4th ed.). Belmont, CA: Duxbury Thomson Brooks/Cole.
3. Barnett, A.G; Pols, J. C. V. D. and Dobson, A.J. (2005). Regression to the mean: what it is and how to deal with it. *International Journal of Epidemiology*, 34, 215-220.
4. Galton, F. (1877). Typical Laws of Heredity. *Nature*, 15 (389), 512-514.
5. Galton, F. (1877). Typical Laws of Heredity. *Nature*, 15 (390), 532-533.
6. Galton, F. (1886). Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15, 246-263.
7. Bland, J. M. and Altman, D. G. (1994). Regression towards the mean. *British Medical Journal*, 308(6942), 1499.
8. Stigler, S. M. (1997). Regression towards the mean, historically considered. *Biometrics*, 29(1), 121-130.
9. Ruck, A. and Sylven, C. (2006). Improvement in the placebo group could be due to regression to the mean as well as to sociobiologic factors. *The American Journal of Cardiology*, 97 (1), 152-153.
10. Linden, A. (2013). Assessing regression to the mean effects in health care initiatives. *BMC Medical Research Methodology*, 13(119), 1-7.
11. Vickers, A. J. and Altman, D. G. (2001). Analysing controlled trials with baseline and follow-up measurements. *British Medical Journal*, 323, 1123–1124.
12. Shephard, R. J. (2003). Regression to the mean. a threat to exercise science? *Sports Medicine*, 33(8), 575-584.
13. Nevill, A.; Holder, R.; Atkinson, G. and Copas, J. (2004). The dangers of

reporting spurious regression to the mean. *Journal of Sports Sciences*, 22, 800–802.

14. Chiolero, A.; Paradis, G.; Rich, B. and Hanley, J. A. (2013). Assessing the relationship between the baseline value of a continuous variable and subsequent change over time. *Frontiers in Public Health*, 1(29), 1-8.
15. Gardner, M. J. and Heady, J. A. (1973). Some effects of within-person variability in epidemiological studies. *The Journal of Chronic Disease*, 26, 781-795.
16. Gmel, G.; Wicki, M.; Rehm, J. and Heeb, J. (2007). Estimating regression to the mean and true effects of an intervention in a four-wave panel study. *Addiction*, 103, 32-41.
17. Ederer, F. (1972). Serum cholesterol changes: effects of diet and regression toward the mean. *Journal of Chronic Disease*, 25, 277-289.
18. Davis, C. E. (1976). The effect of regression to the mean in epidemiologic and clinical studies. *Journal of Epidemiology*, 104(5), 493-498.
19. James, K. E. (1973). Regression toward the mean in uncontrolled clinical studies. *The American Statistician*, 62(4), 289-295.
20. Shepard, D. S. and Finison, L. J. (1983). Blood pressure reductions: correcting for regression to the mean. *Preventive Medicine*, 12, 304-317.
21. Pearson, K. (1936). Method of moments and method of maximum likelihood. *Biometrika*, 28, 34-59.
22. Aldrich, J. (1997). R. A. Fisher and the making of maximum likelihood 1912-1922. *Statistical Science*, 12(3), 162-176.
23. Cohen, A. C. (1955). Restriction and selection in samples from bivariate normal distributions. *Journal of the American Statistical Association*, 50, 884-893.
24. Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A*, 222, 309-368.
25. Johnson, W. D. and George, V. T. (1991). Effect of regression to the mean in the presence of within-subject variability. *Statistics in Medicine*, 10, 1295-1302.
26. Yudkin, P. L. and Stratton, I.M. (1996). How to deal with regression to the mean in intervention studies. *Lancet*, 347, 241-243.
27. Chinn, S. and Heller, R. F. (1981). Some further results concerning regression to

the mean. *American Journal of Epidemiology*, 114(6), 902-905.

28. Nesselroade, J. R.; Stigler, S. M. and Baltes, P. B. (1980). Regression toward the mean and the study of change. *Psychological Bulletin*, 88(3), 622-637.
29. Mee, R. W. and Chua, T. C. (1991). Regression toward the mean and the paired sample t test. *The American Statistician*, 45(1), 39-42.
30. Copas, J. B. (1997). Using regression models for prediction: shrinkage and regression to the mean. *Statistical Methods in Medical Research*, 6, 167-183.
31. Chuang-Stein, C. and Tong, D. M. (1997). The impact and implication of regression to the mean on the design and analysis of medical investigations. *Statistical Methods in Medical Research*, 6, 115-128.
32. Lin, H. M. and Hughes, M. D. (1997). Adjusting for regression toward the mean when variables are normally distributed. *Statistical Methods in Medical Research*, 6, 129-146.
33. Senn, S. J.; Brown, R. A. and James, K. E. (1985). Estimating treatment effects in clinical trials subject to regression to the mean. *Biometrics*, 41(2), 555-560.
34. McCambridge, J.; Kypri, K. and McElduff, P. (2014). Regression to the mean and alcohol consumption: A cohort study exploring implications for the interpretation of change in control groups in brief intervention trials. *Drug and Alcohol Dependence*, 135, 156– 159.
35. Chesher, A. (1997). Non-normal variation and regression to the mean. *Statistical Methods in Medical Research*, 6, 147-166.
36. Das, P. and Mulder, P. G. H. (1983). Regression to the mode. *Statistica Neerlandica*, 37, 15-20.
37. Senn, S. (1990). Regression: a new mode for an old meaning? *The American Statistician*, 44(2), 181-183.
38. Schmittlein, D. C. (1989). Surprising inferences from unsurprising observations: do conditional expectations really regress to the mean? *The American Statistician*, 43(3), 176-183.
39. John, M. and Jawad, A. F. (2010). Assessing the regression to the mean for non-normal populations via kernel estimators. *North American Journal of Medical Sciences*, 2(7), 288-292.

40. Muller, H.; Abramson, I. and Azari, R. (2003). Nonparametric regression to the mean. *Proceedings of the National Academy of Sciences*, 100(17), 9715-9720.
41. Bland, J. M. and Altman, D. G. (1994). Some examples of regression towards the mean. *British Medical Journal*, 309, 780.
42. Morton, V. and Torgerson, D. J. (2003). The effect of regression to the mean on decision making in health care. *British Medical Journal*, 326, 1083-1084.
43. Morton, V. and Torgerson, D. J. (2005). Regression to the mean: treatment effect without the intervention. *Journal of Evaluation in Clinical Practice*, 11(1), 59-65.
44. Zhang, X. and Tomblin, J. B. (2003). Explaining and controlling regression to the mean in repeated measures research designs. *Journal of Speech Language and Hearing Research*, 46(6), 1340-1351.
45. Greenwood, M. C.; Rathi, J.; Hakim, A. J.; Scott, D. L. and Doyle, D. V. (2007). Regression to the mean using the disease activity score in eligibility and response criteria for prescribing TNF-a inhibitors in adults with rheumatoid arthritis. *Rheumatology*, 46(7), 1165-1167.
46. Chapurlat, R. D.; Blackwell, T; Bauer, D. C. and Cummings, S. R. (2001). Changes in biochemical markers of bone turnover in women treated with Raloxifene: influence of regression to the means. *Osteoporosis International*, 12(12), 1006-1014.
47. Grambsch, P. (2008). Regression to the mean, murder rates, and Shall-Issue laws. *The American Statistician*, 62(4), 289-295.
48. Goodin D. S. (2004). Disease-modifying therapy in MS: a critical review of the literature. Part I: analysis of clinical trial errors. *Journal of Neurology*, 251(5), v3-v11.
49. Martinez-Yelamos, S; Martinez-Yelamos, A.; Martin Ozaeta, G.; Casado, V.; Carmona, O. and Arbizu, T. (2006). Regression to the mean in multiple sclerosis. *Multiple Sclerosis*, 12(6), 826-829.
50. Filippini G., Munard L., Incovaia B., Ebers G. C., Polman C. and D'Amico et al. (2003). Interferons in relapsing remitting multiple sclerosis: a systematic review. *Lancet*, 2003, 961, 545-552.

51. Confavreux C., Compston D. A. S., Hommes O. R., McDonald W. I., Thompson A. J. (1992). EDMUS, a European database for multiple sclerosis. *Journal of Neurology, Neurosurgery and Psychiatry*, 55, 671-676.
52. Tongia, R.; Gupta, R.; Agarwal, M. and Gupta, V. P. (2005). Five-year blood pressure trends and regression-to-the mean in an industrial population. *Journal of the Association of Physicians of India*, 53, 693-696.
53. Tinkelman, D. and Wilson, S. (2004). Asthma disease management: regression to the mean or better? *The American Journal of Managed Care*, 10(12), 948-954.
54. Levin-Scherz, J.; Wilson, T.; Linden, A.; Wilson, S. and Tinkelman, D. (2005). Correspondence regarding "asthma disease management: regression to the mean or better?" *The American Journal of Managed Care*, 11(3), 136-137.
55. Cummings, J. L.; Tractenberg, R. E.; Gamst, A.; Teri, L.; Masterman, D. and Thal, L. J. (2004). Regression to the mean: implications for clinical trials of psychotropic agents in dementia. *Current Alzheimer Research*, 1(4), 323-328.
56. Cohen-Mansfield, J. and Billig, N. (1986). Agitated behaviors in the elderly. I. A conceptual review. *Journal of the American Geriatrics Society*, 34(10), 711-721.
57. Koss, E.; Weiner, M.; Ernesto, C.; Cohen-Mansfield, J.; Ferris, S. H. and Grundman, M. et al. (1997). Assessing patterns of agitation in Alzheimer's disease: patients with the Cohen-Mansfield agitation inventory. *Alzheimer Disease and Associated Disorders*, 11(2), S45-S50.
58. McDonald, C. J. and Mazzuca S. A. (1983). How much of the placebo 'effect' is really statistical regression? *Statistics in Medicine*, 2, 417-427.
59. Finney, J. W. (2008). Regression to the mean in substance use disorder treatment research. *Addiction*, 103(1), 42-52.
60. Stout, R. L. (2008). Regression to the mean in addiction research. *Addiction*, 103(1), 53.
61. Ripatti, S. and Makela, P. (2008). Conditional models accounting for regression to the mean in observational multi-wave panel studies on alcohol consumption. *Addiction*, 103(1), 24-31.
62. Atkinson. G. and Taylor, C. (2011). Normalization effect of sports training on blood pressure in hypertensive individuals: Regression to the mean? *Journal of*

Sports Sciences, 29(6), 643-644.

63. Atkinson, G.; Taylor, C. E. and Jones, H. (2010). Inter-individual variability in the improvement of physiological risk factors for disease: gene polymorphisms or simply regression to the mean? *The Journal of Physiology*, 588.6, 1023-1024.
64. Shephard, R. J. (2003). Need to evaluate regression to the mean. *Medicine and Science in Sports and Exercise*, 35(5), 886.
65. Rejeski, W. J.; Brawley, L. R. and Norris, J. L. (2003). When study designs control for regression to the mean. *Medicine and Science in Sports and Exercise*, 35(5), 887.
66. Liang, K. and Zeger, S. L. (1986). Repeated measures data analysis using generalized linear models. *Biometrika*, 73(1), 13-22.
67. Hanley, J. A.; Negassa, A.; Edwardes, M. D. deB. and Forrester, J. E. (2003). Statistical analysis of correlated data using generalized estimating equations: an orientation. *American Journal of Epidemiology*, 157(4), 364-375.
68. Sashegyi, A. I.; Brown, K. S. and Farrell, P. J. (2000). Application of a generalized random effects regression model for cluster-correlated repeated measures data to a school-based smoking prevention trial. *American Journal of Epidemiology*, 152(12), 1192-1200.
69. Efron, B. and Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Boca Raton, FL: Chapman and Hall/CRC.
70. Vale, C. D. and Maurelli, V. A. (1983). Simulating multivariate nonnormal distributions. *Psychometrika*, 48(3), 465-471.
71. Kaiser, H. F. and Dickman, K. (1962). Sample and population score matrices and sample correlation matrices from an arbitrary population correlation matrix. *Psychometrika*, 27(2), 179-182.
72. Wild, T.C.; Wolfe, J. and Currie, C. (2006). The Alberta youth experience survey (TAYES): technical report. *Edmonton: Addiction and Mental Health Research Laboratory*, University of Alberta.
73. Allen, J. P.; Litten, R. Z.; Fertig, J. B. and Babor, T. (1997). A review of research on the alcohol use disorders identification test (AUDIT). *Alcoholism: Clinical and Experimental Research*, 21, 613-619.

74. Conigrave, K. M.; Saunders, J. B. and Reznik, R. B. (1995). Predictive capacity of the AUDIT questionnaire for alcohol-related harm. *Addiction*, 90, 1479-1485.
75. Saunders, J. B.; Aasland, O. G.; Babor, T. F., De La Fuente, J. R. and Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption— II. *Addiction*, 88, 791-804.
76. Reinert, D.F. and Allen, J.P. (2002). The alcohol use disorders identification test (AUDIT): A review of recent research. *Alcoholism: Clinical and Experimental Research*, 26, 272-279.
77. Romelsjo A.; Leifman H. and Nystrom S. (1995). A comparative study of two methods for the measurement of alcohol consumption in the general population. *International Journal of Epidemiology*, 24, 929-936.
78. Heather, N. (2014). Interpreting null findings from trials of alcohol brief interventions. *Frontiers in Psychiatry*, 5(85), 1-9.

Appendix A

A.1 A sample report

Consider a hypothetical example of a student, Adam Smith, who participated in the PAF Alcohol Intervention Study. When the intervention was made available to all students in October 2007 by an email invitation, he got the intervention. After completing the AUDIT items, he received the intervention as a personalized drinking profile. The profile consisted of a normative pie chart and a report of his drinking problem. The profile is given below.

Final Report for

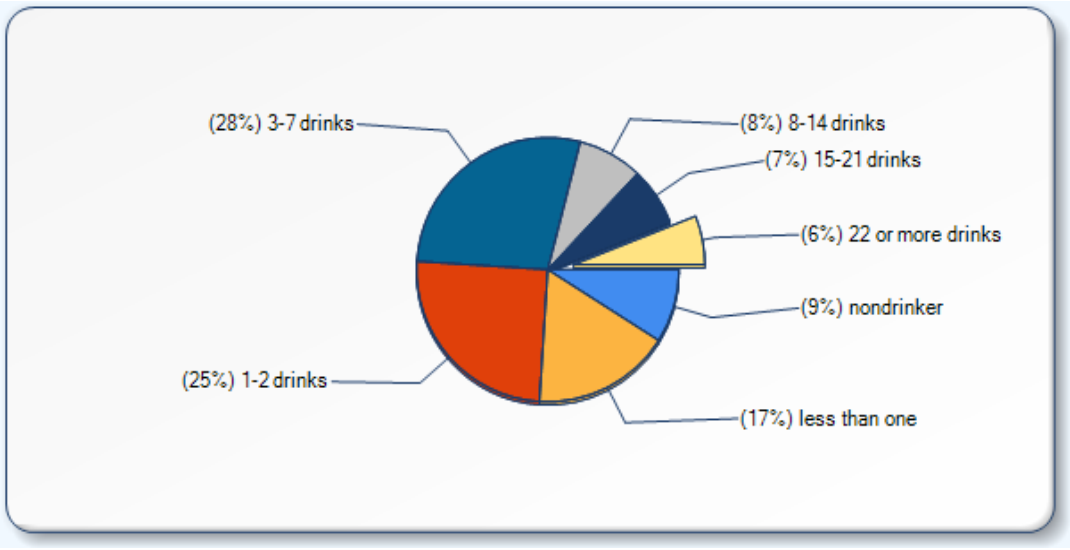
Adam Smith

The average number of drinks you reported consuming per week was 26.

How do you compare to males your age from Canada?

The highlighted slice of the pie chart below is where your drinking fits compares to other males in your age range from Canada.

Average drinks per week for males aged 18 - 24 from Canada



Within the last twelve (12) months:

- You reported drinking on approximately 100.00% of days in the last year.
- You reported that you drank a total of 1,352 drinks in the last year.

This also means that:

- You spent approximately $4,056.00 in the last year, depending on where you drank (at home, in a bar, etc.).
- You consumed (on average) 300 calories from alcohol on days that you drink. Based on the total amount of drinking you had enough alcohol to add roughly 38 pounds or 17.27 kilograms to your weight in the last year. Note: One drink has about 100 calories and 3,500 calories roughly equals 1 extra pound of weight.
- You also reported that within the past year, the greatest number of drinks you had on one occasion was 12 drinks.

Your Drinking Patterns

The following graph outlines how your weekly alcohol consumption rates compare to other males in your age range from Canada.

Drinks per day for males aged 18 - 24 from Canada