METHODS FOR ESTIMATING MEDIATION EFFECT IN SURVIVAL ANALYSIS: DOES WEIGHT LOSS MEDIATE THE UNDERNUTRITION-MORTALITY RELATIONSHIP IN THE OLDER ADULTS?

by

YANHUI SUN

DAVID L. ROTH, CHAIR
INMACULADA ABAN
GARY R. CUTTER
JULIE L. LOCHER
CHRISTINE S. RITCHIE

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

BIRMINGHAM, ALABAMA

2010
METHODS FOR ESTIMATING MEDIATION EFFECT IN SURVIVAL ANALYSIS: DOES WEIGHT LOSS MEDIATE THE UNDERNUTRITION-MORTALITY RELATIONSHIP IN THE OLDER ADULTS?

YANHUI SUN

BIOSTATISTICS

ABSTRACT

The influence of undernutrition on mortality and other adverse outcomes through the mechanism of unintentional weight loss in older adults is often assumed, but the analytic methods to test these mediation mechanisms are not well-developed, and the need for methodological advances in this area motivated this program of research. We first examined the test-retest reliability and predictive validity of self-reported caloric intake as a measure of undernutrition. Acceptable reliability was observed, consistent with previous reports, but the evidence for predictive validity was inconsistent and self-reported caloric intake deficiency was not found to be related to observed weight loss. We then extended the existing mediation methods in survival analysis by conducting a simulation study to further investigate the properties of two mediation effect calculation methods under the condition when censored data is present. Our findings from examination of product of coefficients method did not show a clear pattern in terms of bias with different specifications of hazard rate, mortality rate and amount of censoring. However we did find point estimates with increasing hazard rate have shown the smallest standard error and mean square error, followed by the constant and decreased hazard rates. The comparison between two mediation effect methods showed there is evidence that two methods can lead to substantially different estimates and inferential conclusions under the impact of hazard rate, mortality rate and amount of censoring. Generally
speaking, the product of coefficient method performs better than the other under most of scenarios with moderate sample size, and two methods become less distinguishable when sample size increases to be 1000. Further, we also applied this improved method to a population of older adults and our findings indicated that the causal relationship between certain risk factors and mortality are mediated by weight loss. Finally, we have contributed a novel input to the research of examining mediation effects in the context of survival analysis with censored data and have made recommendation regarding the choice between two methods under difference scenarios.
DEDICATION

I would like to dedicate this dissertation to my parents, Cailing Sun and Chunshan Sun, who have taught me the importance of education, honesty and hard work.

I would also like to dedicate this dissertation to my husband, Haiying Wan, who has played a crucial role throughout my doctoral program. I am eternally grateful and appreciative for all the support he has provided me with the best of everything.
ACKNOWLEDGEMENTS

This dissertation could not have been written without Dr. David Roth, who not only served as my advisor but also encouraged and challenged me throughout my academic program. Throughout my doctoral work he was always there for me and put up with me despite his busy schedule. I am really grateful to him for patiently going through the several thousand drafts of this dissertation and never failing to ensure that the revised version was better.

I am also very grateful to my very accomplished doctoral committee and wish to thank Dr. Inmaculada Aban, Dr. Gary Cutter, Dr. Julie Locher and Dr. Christine Ritchie for their continual support and encouragement. Dr. Locher has played a crucial role in supporting my very first published paper. Dr. Aban has provided me with specific guidance in developing complicated mathematical derivations. Dr. Ritchie has given me valuable suggestions in her area of expertise. Dr. Cutter has supported me in numerous ways, and I will forever cherish our very interesting conversations on statistics, my research, and life in general.

I am extremely grateful to the Department of Biostatistics for giving me the opportunity to pursue a dream. I would like to thank all my teachers at UAB for helping me understand the intricacies of statistics. The administrative staff at the Department of Biostatistics, especially Della Daniel, deserves special commendation for their invaluable assistance.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>v</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xii</td>
</tr>
<tr>
<td>MOTIVATION AND LITERATURE REVIEW</td>
<td>1</td>
</tr>
<tr>
<td>Motivating Example</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>6</td>
</tr>
<tr>
<td>Undereating in Older Adults</td>
<td>8</td>
</tr>
<tr>
<td>Unintentional Weight Loss in Older Adults</td>
<td>14</td>
</tr>
<tr>
<td>Association between Undereating, Unintentional Weight Loss,</td>
<td>16</td>
</tr>
<tr>
<td>and Mortality in Older Adults</td>
<td></td>
</tr>
<tr>
<td>Existing Methods to Estimate and Test Mediated Effects in</td>
<td>17</td>
</tr>
<tr>
<td>General</td>
<td></td>
</tr>
<tr>
<td>Existing Method to Estimate Mediated Effects in Survival</td>
<td>22</td>
</tr>
<tr>
<td>Analysis</td>
<td></td>
</tr>
<tr>
<td>Proposed Methods</td>
<td>25</td>
</tr>
<tr>
<td>RELIABILITY AND PREDICTIVE VALIDITY OF CALORIC INTAKE</td>
<td>32</td>
</tr>
<tr>
<td>MEASURES FROM THE 24-HOUR DIETARY RECALLS OF</td>
<td></td>
</tr>
<tr>
<td>HOMEBOUND OLDER ADULTS</td>
<td></td>
</tr>
<tr>
<td>A SIMULATION STUDY OF MEDIATION EFFECT ESTIMATION IN</td>
<td>48</td>
</tr>
<tr>
<td>SURVIVAL ANALYSES WITH CENSORED DATA</td>
<td></td>
</tr>
<tr>
<td>THE MULTIVARIATE DETERMINANTS OF UNINTENTIONAL</td>
<td>85</td>
</tr>
<tr>
<td>WEIGHT LOSS AND POTENTIAL MEDIATION EFFECT OF WEIGHT</td>
<td></td>
</tr>
<tr>
<td>LOSS ON MORTALITY IN COMMUNITY-DWELLING OLDER ADULTS</td>
<td></td>
</tr>
<tr>
<td>SUMMARY AND FUTURE RESEARCH</td>
<td>112</td>
</tr>
<tr>
<td>LIST OF GENERAL REFERENCES</td>
<td>119</td>
</tr>
</tbody>
</table>
APPENDICES

A  RELIABILITY AND VALIDITY OF THE 24-HOUR DIETARY RECALL IN HOMEBOUND OLDER ADULTS...........................................127

B  WEIGHT LOSS AND MORTALITY IN THE COMMUNITY-DWELLING OLDER ADULTS (MOBILITY AMONG OLDER AFRICAN AMERICANS AND WHITES).........................................128
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Tables</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>RELIABILITY AND PREDICTIVE VALIDITY OF CALORIC INTAKE MEASURES FROM THE 24-HOUR DIETARY RECALLS OF HOMEBOUND OLDER ADULTS</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Descriptive statistics for subgroups of participants</td>
</tr>
<tr>
<td>2</td>
<td>Generalized linear and logistic regression models of weight loss for the subgroup of 52 participants</td>
</tr>
<tr>
<td>3</td>
<td>Logistic regression models for multivariate effects of predictors on adverse outcomes for the subgroup of 143 participants</td>
</tr>
<tr>
<td>THE MULTIVARIATE DETERMINANTS OF UNINTENTIONAL WEIGHT LOSS AND POTENTIAL MEDIATION EFFECT OF WEIGHT LOSS ON MORTALITY IN COMMUNITY-DWELLING OLDER ADULTS</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Descriptive statistics for two subgroups and whole sample of participants</td>
</tr>
<tr>
<td>2</td>
<td>Generalized linear model for multivariate determinants of weight loss for the subgroup of 350 participants</td>
</tr>
<tr>
<td>3</td>
<td>Cox proportional hazard model for multivariate effects of predictors on mortality for the subgroup of 350 participants</td>
</tr>
<tr>
<td>4</td>
<td>Cox proportional hazard model for multivariate effects of predictors including weight change on mortality for the subgroup of 350 participants</td>
</tr>
<tr>
<td>5</td>
<td>Mediation effects of weight loss on risk factors-mortality relationship</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>MOTIVATION AND LITERATURE REVIEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Models showing total effect and mediated effect of X on Y ....................... 28</td>
</tr>
<tr>
<td>2</td>
<td>Model showing relationship between undereating and UWL .......................... 29</td>
</tr>
<tr>
<td>3</td>
<td>Model showing relationships among undereating, UWL, and mortality .......... 30</td>
</tr>
<tr>
<td>4</td>
<td>Model showing relationship among potential elements, UWL, and mortality ..... 31</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure</th>
<th>A SIMULATION STUDY OF MEDIATION EFFECT ESTIMATION IN SURVIVAL ANALYSES WITH CENSORED DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Models showing total effect and mediated effect of X on Y .................................................. 71</td>
</tr>
<tr>
<td>2</td>
<td>Distribution of bias of $ab$ from true value at different hazard rates .................................. 72</td>
</tr>
<tr>
<td>3</td>
<td>Distribution of standard errors from point estimate of $ab$ at different hazard rates .................. 73</td>
</tr>
<tr>
<td>4</td>
<td>Distribution of mean square errors from point estimate of $ab$ at different hazard rates .................. 74</td>
</tr>
<tr>
<td>5</td>
<td>Distribution of bias of $ab$ from true value at different mortality rates .................................... 75</td>
</tr>
<tr>
<td>6</td>
<td>Distribution of standard errors from point estimate of $ab$ at different mortality rates .................. 76</td>
</tr>
<tr>
<td>7</td>
<td>Distribution of mean square errors from point estimate of $ab$ at different hazard rates .................. 77</td>
</tr>
<tr>
<td>8</td>
<td>Distribution of bias of $ab$ from true value at different censoring percents ............................. 78</td>
</tr>
<tr>
<td>9</td>
<td>Distribution of standard errors from point estimate of $ab$ at different censoring percents ............ 79</td>
</tr>
</tbody>
</table>
THE MULTIVARIATE DETERMINANTS OF UNINTENTIONAL WEIGHT LOSS AND POTENTIAL MEDIATION EFFECT OF WEIGHT LOSS ON MORTALITY IN COMMUNITY-DWELLING OLDER ADULTS

1 Path models showing relationship among risk factors, unintentional weight loss and mortality ........................................................................................................................................111
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EER</td>
<td>Estimated Energy Requirements</td>
</tr>
<tr>
<td>GLM</td>
<td>generalized linear model</td>
</tr>
<tr>
<td>MLE</td>
<td>maximum likelihood estimator</td>
</tr>
<tr>
<td>MSE</td>
<td>mean square error</td>
</tr>
<tr>
<td>OLS</td>
<td>ordinary least squares</td>
</tr>
<tr>
<td>SE</td>
<td>standard error</td>
</tr>
<tr>
<td>UWL</td>
<td>unintentional weight loss</td>
</tr>
</tbody>
</table>
MOTIVATION AND LITERATURE REVIEW

1 Motivating Example

Unintentional weight loss which is related to morbidity, mortality and quality of life is associated with significant health problems for older adults [1-3]. Unintentional weight loss can make illness complicated and slow down the recovery procedure from treatment. Unintentional weight loss is also often assumed to be an effect of undernutrition and is related to functional decline, cognitive impairment, and medication usage [4-6]. Although there are various definitions for undernutrition, a general and popular definition of undernutrition in older adults often is related to protein and calorie deficiencies which are insufficient to maintain current body weight [6-8]. Previous studies have shown that unintentional weight loss has resulted in adverse clinical outcomes, increased morbidity and mortality in older adults [9-13]. However, not many studies in the literature have been conducted to determine the influence of undereating on mortality or other adverse outcomes such as hospitalization, institutionalization, and health decline, through the mediated effect of unintentional weight loss in older adults.

The relationship between two variables, X and Y, including conditions that X as a possible cause of Y has been investigated a lot by most research. A simple scenario of mediation is that it adds another variable, M, to this X→Y relation, where X predicts M, and as a result, M predicts Y; therefore, the relation of three variables becomes X→M→Y. Figure 1 describes the elements from a single mediator model.
Let’s assume variable X predicts variable Y. Part 1 of Figure 1 shows the simple relation between X and Y and indicates that a unit change of X is associated with \( t \) units change of Y. For instance, let Y be the score scaled 0-120 for the TOEFL exam, and X be the number of days spent studying for the exam, then \( t \) is the scores expected to obtain for every day of study. Since the research in this proposal will involve a binary dependent variable, an example for dichotomous event is given here. For instance, let Y be a binary dependent variable, with 1 indicating mortality in older adults, and 0 indicating survival over some time period; X be an indicator of undernutrition such as the caloric intake discrepancy calculated by taking the difference between a participant’s self-reported mean daily caloric intake and their Estimated Energy Requirements; then under the model of logistic regression, \( t \) is the log odds of mortality that one expects to see increase for each unit increase of this undernutrition indicator.

Part 2 of Figure 1 describes a model after introduction of a mediator variable, M. Continuing with the example discussed earlier and assuming the dependent variable is continuous, if M is the hours of a new mentorship, then we could determine that \( a \) hours of the new mentorship is provided for every day of study and \( b \) scores are obtained for every hour of mentorship which is provided. Thus, \( ab \) scores are gained through new mentorship provided for each day of study. In the example of dichotomous case, if M is the unintentional weight loss in kilograms, then we could determine that \( a \) kilograms of weight are lost for each unit change in undernutrition and that \( b \) log odds of mortality is increased for each kilogram of weight loss. Therefore, \( ab \) log odds of mortality are changed through weight loss resulting from undernutrition.
Here the effect $t$ in part 1 is called the total effect, $t'$ in part 2 is called the direct effect, and $ab$ is called the indirect or mediated effect. When total effect $t$ and mediated effect $ab$ are equal to each other, it implies that the complete mediation happens. When it is the case, the path $t'$ is zero, as no direct effect exists. If the mediated effect is smaller than the total effect but both have the same sign, it implies that the partial mediation happens. When it is the case, the value of path $t'$ is less than $t$ but not equal to zero. When either complete or partial mediation occurs, both mediated effect $ab$ and direct effect $t'$ have the same sign.

Another situation we would like to point out is where the direct effect and mediated effect have different signs in the model [14-16], i.e., when the inconsistent mediation occurs. When this happens, knowing the relation between predictors $X$ and dependent variable $Y$ is not enough to interpret the results, as there are cases where an overall $X$ to $Y$ relation is not significant, but the mediation exists [17]. For example, let $X$ be the failure of task, $M$ be the solution seeking, and $Y$ be the resulted stress. From Figure 1, path $a$ is positive (failure of task makes one seek solution), but path $b$ is negative (solution seeking reduces one’s stress level) and, therefore, the indirect effect of $ab$ is negative. If solution seeking is held constant, the failure of task is positively associated with stress level, which is confirmed by $t' > 0$. The total effect $t$ is the sum of $ab$ and $t'$. Under the circumstance that a person has a good solution seeking skills, $t'$ and $ab$ may have a similar magnitude but opposite signs. If this happens, the total effect $t$ may be close to zero. It clearly shows that only considering the simple effect of $X$ on $Y$ can sometimes obscure the complexity and fail to detect the real causal relations among these variables. The mediation analysis taking into account inconsistent mediation effect can be
conducted to better describe more complicated causal relations between variables. In the model where more than one mediator and different signs are present, considering inconsistent mediation could explain why a relation is expected to be interesting but is weak in magnitude [18].

Another term which is almost always mentioned with mediator as a pair is moderator. As described from the classic paper written by Baron and Kenny [19], the different impacts that a moderator and a mediator can have on the relationship of the other variables is that a moderator mostly affects the direction and strength of the relationship, while a moderator mostly explains the relationship of the other variables and a mediator is a variable that explains the relationship. Let’s take the same example regarding the relation between undernutrition and mortality. Race might be a moderator variable since white and African American participants might have a different relation between undernutrition and mortality. Unintentional weight loss, on the other hand, could be a valid mediator since it explains why undernutrition predicts mortality. If this is the case, then the relation between undernutrition and mortality disappears when the effect of weight loss is removed. In this proposed study, we have included some covariates into the analysis models that might be moderators; however, our study focuses mostly on the properties of mediation effects.

There are two ways to calculate the mediated effect in a single-mediator model, $ab$ or $t-t'$ [20]. Under the assumption that the dependent variable is continuous and the residual variance follows a normal distribution, MacKinnon et al. [20] has used ordinary least squares regression to show the algebraic equivalence of $ab$ and $t-t'$ as the measures of mediation. However, in mediation analyses with binary dependent variables, the
equivalence of $ab$ and $t-t'$ does not always hold [20]. In particular, in a survival analysis with time-to-event outcomes, Tein and MacKinnon [21] examined the estimates of the mediated effect $ab$ and $t-t'$ in the log-survival time (i.e., LIFEREG procedure) and log-hazard time (i.e., PHREG procedure) models. With the assumptions of no censored event, time-invariant mediating variable, and no ties among the events, they found that the approximate equivalence of two mediation methods does apply to log-survival time survival analyses, but not to log-hazard time survival analyses (Cox’s proportional hazard model). The differences between two methods were close to zero in the LIFEREG procedure. However, in the PHREG procedure, the differences between $ab$ and $t-t'$ were not negligible. Specifically, the mediation estimates for $|t-t'|$ was consistently smaller than those for $|ab|$, although the difference decreased as the sample size got larger.

In this proposed study, we would like to extend the existing methods on both its assumptions and model to further examine the estimation of mediation effects in the context of survival or time-to-event models, although it is well-known that there are many attractive features about Cox’s proportional hazard model. For instance, the probability distribution of the survival times is not required in Cox's model, which makes it more robust [22]. However, Tein and MacKinnon’s [21] simulation show a very stable scenario for the LIFEREG procedure. The extension of their studies by introducing the censored data is a necessity to add a further complexity to the model but not to involve too many unstable parameters that could compromise the results. Examining whether the conclusions of Tein and MacKinnon’ study apply to data that have censored data, especially focusing more on log-life survival-time model will be a good contribution to the literature.
2 Statement of the Problem

Due to the special health, nutrition, and mortality interests in community-dwelling older adult populations, the primary goals of this research are to advance methods for testing mediation relationships relating nutrition, weight change, and mortality variables. Consider a longitudinal observational study to assess the relationship between undereating, unintentional weight loss, and mortality or hospitalization in homebound older adults using a questionnaire interview and 24-hour dietary recalls. Weight loss is obtained by taking the difference of a participant’s measured weight at two time periods. As part of this program of research, there are three specific aims to be achieved:

1. To examine the reliability and validity of self-report caloric intake deficiency measures from 24-hour dietary recalls for a vulnerable population of homebound older adults. As shown in Figure 2, we will examine the predictive validity of 24-hour dietary recalls by evaluating the effect of caloric intake deficiencies at baseline on observed weight loss over the subsequent 6 months. This validation aspect of the study also consists of the $a$ path of our undereating $\rightarrow$ UWL $\rightarrow$ mortality mediation model. In addition, the test-retest reliability of self-reported caloric intake will be evaluated as well to examine the consistency of the measurement from baseline to 6 months.

2. To conduct a statistical simulation of the effect of undereating on mortality through the mediating effect of weight loss in the context of survival models. It is expressed in Figure 3. A range of possible effect sizes for $a$ and $b$ will be specified based on the research literature and the existing studies, and the simulations will yield evaluations for the mediated effect in analyses with censored events. The property of $ab$ will be first examined in terms of bias, standard error, and mean square error. The two
methods will be also compared to the true value of $ab$ under different specifications of mortality rate, hazard rate, and percent of censoring data. Overall, based on those assumptions, simulations will be conducted for each of parameter combinations, sample sizes, and choice of other parameters of our interest. The patterns of bias, SE, and MSE will be summarized and compared over those methods.

3. To examine the multivariate determinants of weight loss besides caloric intake measures and, furthermore, reexamine the relationships between those elements and mortality to check if the mediated effect of weight loss exists (as shown in Figure 4). This model will apply to a UAB Study of Aging dataset. The dataset consists of 1000 community-dwelling African American and white adults 65 years old or older. Participants were administered a baseline questionnaire regarding their overall health status in their homes. A follow-up assessment was conducted every 6 months over the telephone interviews. The study is ongoing. Our research is interested in the observed unintentional weight loss over the initial 4 years and impact on mortality between 4 years and 8.5 years.

As stated earlier, the assumptions and models used to conduct simulation studies of the mediated effect in survival analysis found in the literature are going to be evaluated and extended. This situation motivates this research. In order to achieve these aims, there are several issues that must be addressed. The relationships among undereating, unintentional weight loss and adverse outcomes such as mortality and hospitalization must be examined. The measurement of undereating and weight loss in previous studies must be reviewed. While estimating the mediated effect, the standard error of those estimates must be evaluated as well.
3 Undereating in Older Adults

3.1 The Measurements of Undereating

Undernutrition has significant health, social, and economic consequences in older adults [23-25]. Undernutrition in older adults is very prevalent, and studies show that undernutrition occurs in up to 20% of community-dwelling older adults, 50% of hospitalized older patients, and 11% of medical outpatients [26-27]. Due to the fact that the agreement on “normal” nutritional requirements in healthy older adults has not been reached by the experts and the limitation of understanding the impact of illness on nutritional status, it is difficult to detect and diagnose the undernutrition problem in older adults [6,28-29].

Although food intake can be reduced by age-related problems, nutrient digestion is rarely affected by aging itself. The difference in basic nutritional requirements is quite small between younger and older adults, although older adults have less leaner body mass and less activity, which might lead to a slightly lower requirements of energy [30]. However, a person's body mass, height, and body fat distribution is affected by the natural aging procedure; therefore, the signs of undernutrition in children and younger adults may not apply to older adults [31].

Although it is not easy to diagnose undernutrition, a combination of dietary assessment, anthropometric measurements, and laboratory tests has been used as a promising tool to assess undernutrition [6]. For example, in a prevalence study of undernutrition in older subjects, after carefully reviewing 18 studies, Clark et al. [32] summarized that about 20% of homebound older adults and up to 60% of long-term institutional patients have been diagnosed with undernutrition problem. Among these 18
studies, each study has at least two criteria to assess undernutrition, including serum albumin or transferring, and anthropometric measures, such as BMI value or percentile, weight for height, skinfold thickness, and mid-arm circumference, etc. In particular, 13 studies have used serum biomarker as a criterion for undernutrition. But there are some limitations in the use of serum albumin. As serum albumin decreases gradually due to its relatively long half-life, there are almost always signs of undernutrition associated with lower blood level [6]. However, because reduced albumin can cause many other problems in the ill older adults, some researchers have used serum transferring as an indicator to reflect the protein status [33]. For instance, Liu and his colleagues [2] examined the association between protein-energy undernutrition and mortality using a serum biomarker. Despite the varieties of assessment of undernutrition, however, in the literature we find out that the 24-hour dietary recalls were used in most of the studies to obtain a participant’s mean daily caloric intake in order to determine whether a subject is undereating or not.

Payette [34] reported in a cross-sectional study of older participants receiving community services that low levels of energy intake were observed in more than 70% of participants, and a high prevalence of deficient vitamin and mineral intake was observed as well, indicating that older adults who have restricted mobility or resources and are receiving home health services may have high risk to experience undernutrition. Other studies reporting that inadequate caloric intake more often happened in female, African American ethnicity older adults with less education and low income [35-36] used the 24-hour dietary recalls to assess undereating.
The 24-Hour dietary recall is a very popular dietary assessment data collection method. The participant is asked by personal interview or telephone to recall what he or she has eaten and drunk in the preceding 24 hours.

There are several advantages to the 24-hour recall [37]. Thompson and Subar note that high levels of physical functioning and literacy of participants are not required because responses are recorded by an interviewer. This reduces the potential for both nonresponse bias and respondent burden. Further, because the recall period is immediate, participants are able to remember most of what has been consumed in the preceding day. A final advantage is that the likelihood that these self-reports will interfere with typical dietary behavior is minimized because the food has been consumed before dietary recalls occur.

It should be pointed out that there is an important issue related to the measurement of dietary recall, called “reactivity” that needs further evaluation. High reliable methods of measuring food intake such as dietary recalls and food diaries involve either direct observation or self-report measure. One of the distinct advantages of these approaches is that they are direct and behavior specific. However, on the other hand, this creates a major disadvantage as well; that is, the subjects know the amount of food is the variable being investigated and thus potentially highly reactive [38]. If subjects know they will be asked to recall the food intakes (which is true in 24-hour dietary recall), they may underestimate or overestimate (to make they seem to eat the way they are supposed to eat) instead of accurately estimating food intake. In general, any behavior that is knowingly monitored is likely to be affected by the monitoring.
The other weakness associated with the approach that individuals may not reliably report their food intake because of knowledge, memory problems or factors associated with the interview situation [37]. Special problems among the elderly could, but not necessarily, be present due to the frequency of chronic illness, impaired memory, and other factors. Some research shows that the validity of dietary information collected from both older adults and younger adults are comparable to each other and that recall can be enhanced by supplementing standard methods with memory strategies and probes [39-40]. In a cohort study of more than 120,000 men and women ages 55 to 69 conducted to assess the validity of the food frequency questionnaire, Goldbohm et al. [39] evaluated the representativeness of the study subgroup (59 men and 50 women) for the entire cohort. Correlation coefficients for most nutrients and caloric intake were only slightly modified when the results were extrapolated to the cohort at large, indicating the good representativeness of the subgroup. In another study based on a single 24-hour dietary recall followed by 14 consecutive days of food records, Drewnowski et al. [40] recruited 24 younger people ages 20-30 years and 24 older people ages 60-75 years to assess the dietary variety. It was found that older adults did not consume less varied diets than younger subjects. Therefore, it specifically addressed the question that aging is not necessarily associated with the consumption of monotonous and nutritionally inadequate diets and the dietary information for younger and older adults is comparable.

The reliability of dietary recall as a measurement of undereating has been investigated in different populations, as well, and has shown promising results. Dawber et al. [41] found a correlation of two measurements agreed against a varied diet separated by 2 years of caloric intake, 0.92 calculated by the same nutritionist and 0.89 by different
nutritionist, while McDann et al. [42] found a much lower correlation of 0.59 for caloric intakes. In another study by Reshef et al. [43], it was found the correlations between two measurements separated by seven and half months in a group aged 40 years and over with both sexes were between 0.68 and 0.89, depending on the gender and race of participants. In a more recent study conducted by the Epic group of Spain [44], a biological marker of nitrogen excretion and nitrogen intake was used to examine the reliability of four repetitive 24-hour dietary recalls for a group of people aged 35-60 years with both sexes, and a coefficient of 0.52 was reported.

In this proposed study, caloric intake discrepancy and undereating are defined as the difference between a participant's mean daily caloric intake and his/her Estimated Energy Requirements (EER). Caloric intake discrepancy is the actual difference measured as a continuous variable and undereating is a dichotomized variable. EER is computed using a formula according to the Institute of Medicine [45].

For females, the EER formula is:

\[
\text{Energy(kcal)} = 354.1 - (6.91 \times \text{Age}[\text{y}]) + \text{PhysicalActivityCoefficient} [1 \text{forsedentary}] \\
\times (9.36 \times \text{Weight}[\text{kg}] + 726 \times \text{Height}[\text{m}]) ;
\]

For males, the formula is:

\[
\text{Energy(kcal)} = 661.8 - (9.53 \times \text{Age}[\text{y}]) + \text{PhysicalActivityCoefficient} [1 \text{forsedentary}] \\
\times (15.91 \times \text{Weight}[\text{kg}] + 539.6 \times \text{Height}[\text{m}]) .
\]

As is mentioned in the prevalence studies of undernutrition for elderly subjects [32] in this section, there are a lot of other studies using BMI to screen for the presence of undereating, as well [46-50]. BMI calculated from weight (kg) divided by the square of height (m) is an index of weight-for-height. Generally speaking, a BMI of < 18.5 kg/m² is
considered underweight; however, for older adults, a BMI of <22 indicates illness or malnutrition according to the Nutrition Screening Initiative guidelines [51]. BMI is a reliable estimate of weight adequacy for height and has considerably contributed to numerous studies related to nutrition. But undernutrition is a lot more complicated topic, especially for older adults with already compromised health status. The values of BMI will vary considerably depending on age, gender, and ethnicity. Therefore, in a population of persons, such as older subjects, whose have different body composition from that of healthy younger adults, BMI might not be an ideal tool [52]. BMI has a very stable time dimension, i.e., people with normal weight may be still undereating and starting to lose weight without being captured by the fluctuation of BMI. Thus, BMI is a highly reliable measurement, but its insensitiveness of recognizing recent but significant weight losses and not being able to distinguish between loss of lean body mass and loss of fat mass makes BMI a limited, questionable measure of assessing undernutrition.

3.2 The Consequences of Undereating in Older Adults

Consequences associated with undernutrition in older adults are severe. They include increased mortality and morbidity, functional decline, health decline, and increased possibility of hospitalization and institutionalization [1-3]. Homebound older adults may have even more consequences from undernutrition as insufficient caloric consumption and illness may adversely affect each other [45,53].

In the general population of older adults, many factors have been reported in studies associated with inadequate caloric intake, including poverty, lack of access to community resources, cognitive impairment, reduced social support, functional
impairment, poor mental health and poor oral health [1,25,54-56]. Those issues exist for the homebound older adults as well and, therefore, increase their likelihood of undernutrition.

Previous research that has been conducted with older adults finds that inadequate caloric intake measured by 24-hour dietary recalls is very common, ranging between 71% and 93% of individuals [35-36,57-58], and is associated with deterioration of disease [36], female, African American, low income, less education, not eating breakfast [35], and multi-morbidity [59].

4 Unintentional Weight Loss in Older Adults

Unintentional weight loss (UWL) is another common issue among older adults and is associated with increased mortality, significant adverse clinical outcomes and progressive morbidity. It is well-known that either intentional or unintentional weight loss without exercise results in the loss of both muscle and fat. It especially reduces physical function in older adults because people naturally lose muscle as they age [60]. In a large cohort study Ensrud et al. [61] reported that women of over 65 years old who experience voluntary weight loss of at least 5% body weight within 4 to 5 years have a higher risk of hip-bone loss and hip fracture. However, from the literature for studying mortality, morbidity, and adverse clinical outcomes, we find either that some studies have excluded participants with intentional weight loss from recruitment [9,62], or that most studies - as discussed below - have made a clear distinction between unintentional weight loss and intentional weight loss if the necessary question is asked. Payette and his colleagues [34], after examining published literature from 1990 to 2003 addressing the
prevalence, determinants, and consequences of undernutrition among 65 years or older adults, have summarized that a high prevalence of unintentional weight loss as well as very low energy intakes have been reported in this population. In clinical practice, it is reported that as many as 27% of frail community-dwelling older adults over 65 years experienced unintentional weight loss of at least 5 kg over the last 2 years of the study [13]. Although as many as one in four older adults do not have obvious identifiable medical causes, unintentional weight loss of about 5% of body weight is often a symptom of one or more diagnosed or undiagnosed illnesses, such as depression, gastrointestinal illnesses, dementia, and cancer [10,63] Oral health problems have also been reported to contribute to significant unintentional weight loss of 10% body weight in the past 12 months among community-dwelling older adults [64].

Previous studies have reported numerous factors associated with unintentional weight loss of at least 5 kg or 5% from 3 to 10 years in older people. They include older age [65-66], disability [67], poorer health status [66], previous admission to hospital [65], widowhood [65-66], low education level [66], smoking [61,66], lower body mass index [66-67], and cognitive impairment [68].

Unintentional weight loss increases mortality in older adults. Unintentional weight loss of 5kg or 5% of body weight can be followed by an increase in mortality as high as 38% within 2.5 years [9-10,52] and higher likelihood of admission to an institution [13,69]. Mortality can be greatly increased after more than 4% of body weight loss occurs within 1 year, or more than 10% of body weight loss occurs within 5 to 10 years [10-11,70-72]. As for older adults who have been hospitalized, the risk of
complications can be increased by unintentional weight loss of 5 kg or more in the past 12 months [12].

5 Associations between Undereating, Unintentional Weight Loss, and Mortality in Older Adults

From the literature, numerous studies have shown there are associations between undereating and mortality, and unintentional weight loss and mortality, in older people. This encourages us to think how much undereating and unintentional weight loss can work mutually together to severely compromise older adults’ health, specifically resulting in mortality and hospitalization. Locher et al. [56] observed in a longitudinal study that the odds of experiencing mortality in older adults with unintentional weight loss were 1.67 times as the odds of those without weight loss. In a 5-year study in seniors, it was reported that unintentional weight loss resulting in a BMI change of at least 2.0 units is significantly predictive of mortality [73]. Lee et al. [74] pointed out that unintentional weight loss may be easier to continue and more difficult to resolve than other weight changes. The participants who had short-term weight fluctuations or instability were more inclined to report poor health status, regardless of their ability to resolve their previous weight changes, than were the participants with stable weights or with no short-term weight changes [74]. This observation warrants continued attention, because it suggests that weight fluctuation or unintentional weight loss could indicate the health problems among older people.

In order to understand the causes of mortality and other adverse outcomes in older adults, and probably have a better chance of improving the older adults’ overall health status and their remaining life, it is necessary to investigate how the causal effect of
undereating on mortality or other adverse outcomes can be explained by the intervening effect of undereating on unintentional weight loss.

In this research, the psychometric properties of 24-hour dietary recall will be evaluated. Again, with both caloric intake discrepancy and unintentional weight loss being derived, and adverse outcomes observed from the longitudinal study, the effects of undereating on mortality or other adverse outcomes through the intervening effect of weight loss will be investigated.

6 Existing Methods to Estimate and Test Mediated Effects in General

What follows is a description of literature relevant to the problem in this dissertation. It is not intended to be exhaustive, but rather to familiarize the reader with various methodologies related to estimating and testing mediated effects.

Equations (1)-(3) are used to estimate the single-mediator model of Figure 1.

\[ M = i_3 + aX + e_1 \]  
\[ Y = i_1 + tX + e_2 \]  
\[ Y = i_2 + t'X + bM + e_3 \]  

Where \( i_1 \) and \( i_2 \) and \( i_3 \) are intercepts, \( Y \) is the dependent random variable, \( X \) is the independent random variable, \( M \) is the mediating variable, \( t \) is the coefficient between \( X \) and \( Y \), \( t' \) is the coefficient between \( X \) and \( Y \) adjusted for \( M \), \( b \) is the coefficient between \( M \) and \( Y \) adjusted for \( X \), \( a \) is the coefficient between \( X \) and \( M \), and \( e_1 \), \( e_2 \), and \( e_3 \) are residuals.
Three different general approaches or methods have been proposed to test mediated effects: (1) The causal steps approach; (2) the difference of coefficients approach as defined by $t-t^*$; and (3) the product of coefficients as defined by $ab$.

6.1 Causal Steps Approach

The causal steps approach described thoroughly by Baron and Kenny [19, 75] and Judd and Kenny [76-77] has been widely used to assess mediation. Mediation can be established with four steps. First, a relation of X to Y is required. Second, a relation of X to M is required. Third, the mediator M is required to be related to Y after adjusting for predictor X in Equation (3). Fourth, the coefficient $t'$ is estimated and tested to determine if complete mediation exists. Because of the possibility of inconsistent mediation, in which the mediated effect has an opposite sign of direct effect so that it makes the total effect disappear, the necessity of testing the overall association between X and Y has been questioned by some researchers [6,78-79]. Others also have pointed out that the test of X to Y could have more power if a valid mediation is present; therefore, it is not appropriate to hold off the consideration of mediation until the relation between X and Y is verified [18]. There are several limitations associated with the causal steps method, since its purpose was not to build a statistical test for the relation of $X \rightarrow M \rightarrow Y$. Firstly, a joint test of the conditions of testing $a=0$, $b=0$ and $t'=0$ as defined by Baron and Kenny is not provided by the causal steps methods. Second, an estimate of the indirect effect of X on Y is not provided. Third, the standard errors are not produced to construct confidence limits [19,75]. In addition, when multiple mediators are present in the model, the causal steps methods are not able to evaluate each of the mediating variables.

18
separately [78,80]. Finally, as mentioned earlier in the case that an inconsistent mediation is present, the significant relation requirement between X and Y might fail to detect a valid mediation effect, as the opposite signs may cancel each other out [16].

6.2 Difference in Coefficients Approach

Another approach to assess mediated effects seems more straightforward, since the calculation is basically the difference between the adjusted and unadjusted regression coefficients \( (H_0: t-t'=0) \) developed by Freeman and Schatzkin [81]. A standard error formula is shown in Equation (4)

\[
\sigma_{\text{Freeman-Schatzkin}} = \sqrt{\sigma_i^2 + \sigma_{i'}^2 - 2\sigma_i\sigma_{i'}\sqrt{1 - \rho_{XI}^2}}
\]  

(4)

In this equation, \( \rho_{XI} \) is the correlation between X and the mediated variable M, \( \sigma_i \) and \( \sigma_{i'} \) are the standard errors of \( t \) and \( t' \), respectively. Test significance is conducted by comparing the division of \( t-t' \) over the standard error from Equation 4 and to the \( t \) distribution.

A standard error formula of \( t-t' \) based on the standardized variables was derived by McGuigan and Langholtz [82-83]:

\[
\sigma_{\text{McGuigan-Langholtz}} = \sqrt{\sigma_i^2 + \sigma_{i'}^2 - 2(\rho_{it}\sigma_i\sigma_{i'})}
\]  

(5)

Again, a significant test is performed by comparing the division of point estimate of \( t-t' \) over the standard error from Equation (5) to the \( t \) distribution.

Other methods comparing the correlations of X and Y before and after the adjustment for mediator are not described in detail at this point. The basic idea of
conducting a significant test is to compare the division the difference of correlations over
the standard error to the standard normal distribution [84].

Each of the difference in coefficients methods in this section provides an estimate
of the mediating effect, and the significance is tested by dividing the difference by the
standard error. As mentioned earlier, one of the limitations with the difference in
coefficients methods is that it does not generalize the tests to detect the significance in the
models with multiple mediators [83].

6.3 Product of Coefficients Approach

This approach is to test the significance by comparing the division of the products
of coefficients $ab$ over its standard error to a standard normal distribution.

Sobel [85] derived an approximate formula to calculate standard error:

$$\sigma_{ab_{\text{first}}} = \sqrt{a^2 \sigma_b^2 + b^2 \sigma_a^2}$$

(6)

Because this formula was based on the first-order Taylor series, it is also called
the first-order Taylor method and has become the most commonly used formula. In this
equation, $\sigma_a$ is the standard error of $a$, and $\sigma_b$ is the standard error of $b$.

In order to test for significance, the division of $ab$ over its standard error $\sigma_{ab_{\text{first}}}$ is
compared to a standard normal distribution ($H_0: ab=0$).

The standard error of $ab$ based on both first order and second order Taylor series
was derived by Aroian [86]:

$$\sigma_{ab_{\text{second}}} = \sqrt{a^2 \sigma_b^2 + b^2 \sigma_a^2 + \sigma_a^2 \sigma_b^2}$$

(7)
The significance is tested to compare the division of \( ab \) over \( \sigma_{\text{absecond}} \) to a standard normal distribution \( (H_0: ab=0) \).

Goodman et al. [87-88] derived the standard error of \( ab \) as shown in Equation (8):

\[
\sigma_{\text{abunbiased}} = \sqrt{a^2\sigma_b^2 + b^2\sigma_a^2 - \sigma_a^2\sigma_b^2}
\] (8)

Notice the only difference between Goodman and Sobel’s formulas is that an extra term of product of variances is subtracted in Goodman’s formula, which should make the two methods very similar except with \( \sigma_{\text{abfirst}} \) being slightly larger than \( \sigma_{\text{abunbiased}} \). Again, in order to test for significance, the point estimate of \( ab \) is divided by its standard error, and this value is then compared to a standard normal distribution.

One limitation of testing the significance of the product of coefficients is that the distribution of \( a \) and \( b \) is often asymmetric distributed with high kurtosis rather than normally distributed [89-91]. Therefore, the \( ab/\sigma_{ab} \) method, which is used to compare to the normal distribution, may have low power [92]. In order to make this method theoretically more accurate, MacKinnon and colleagues [83] proposed three alternative variants after carefully investigating the statistical theory of the random variables products [89,91,93]:

1. Empirical distribution of \( ab/\sigma_{ab} (H_0: ab/\sigma_{ab} = 0) \) and its critical values are derived based on a wide range of values of \( a \) and \( b \) from simulations.

2. Distribution of the product of two standard normal variables, \( z_a z_b (H_0: z_a z_b = 0) \), where \( z_a = a/\sigma_a \) and \( z_b = b/\sigma_b \). If \( a \) and \( b \) are assumed to be
normal, the significance of \( z_a z_b \) can be tested by comparing the value of \( z_a z_b \) to the critical values derived from the distribution of two normal random variables product [91].

3. Asymmetric confidence limits for the product of \( a \) and \( b \). The asymmetric confidence limits are constructed based on the nonnormal distribution of \( ab \). Two \( z \) statistics, \( z_a = a / \sigma_a \) and \( z_b = b / \sigma_b \), are computed, and then used to find critical values determining the lower and upper limits from the tables in Meeker et al. [93]. The formula \( CL = ab \pm (\text{critical value}) \sigma_{ab} \) is used to calculate the confidence limits. A significant mediating effect is indicated if zero does not fall into the confidence interval.

One advantage of the product of coefficients methods over the difference in coefficients methods is that the mediating effect is hypotheses to measure the causal relations between variables; therefore, it can be extended to the models with more than one mediator [94]. However, there are limitations associated with this method as well. For instance, the nonnormal distribution of two normally distributed variables might take efforts to make necessary adjustments. In addition, as it can be seen from the formulas, the hypothesis test involving the product of variables is complicated [83].

7 Existing Method to Estimate Mediated Effects in Survival Analysis

As stated earlier, two methods of calculating mediated effects are employed. The difference of coefficients method calculated as \( t-t' \) in Equations (2) and (3), and the product of coefficients method calculated as \( ab \) in Equations (1) and (3). Under the assumption that the dependent variable is continuous and the residual variance follows normal distribution, the two methods, using ordinary least square regression, yield
identical estimates of mediation effect, $t-t'=ab$ [20]. However, it has been shown that, for mediation analysis with binary dependent variables, $ab$ is not equal to $t-t'$ due to the fact that the residual variance is a constant in the logistic regression so the coefficients of $t$ and $t'$ come from equations on different scales [20]. Tein and MacKinnon [21] also examine the hypothesis whether $t-t'$ equals to $ab$ with survival analyses when the outcomes are time until an event occurs. In the simulation study, they assumed that there was no censored event, the mediating variable was time invariant, and there were no ties among the events. They simulated survival time, $Z_i$, following a Weibull distribution, based on equation [95]

$$Y = \log Z_i = u + t X_i + bM_i + \sigma W$$

Where the dependent variable $Y$ is the log function of $Z_i$, the time of the event for subject $i$. $X_i$ and $M_i$ are the independent and mediating variables, respectively. The residual term is the product of $W$, which has a 2-parameter extreme value distribution, and $\sigma$, which is the scale parameter determining the shape of the hazard rate. They estimated the Weibull distribution of event times with a log-survival time model and a log-hazard time model, the two most flexible procedures to estimate regression models with covariates and censored data. The log-survival time of the model can be estimated by LIFEREG procedure in SAS as

$$\log T_i = t_0 + t_1 x_{i1} + \ldots + t_k x_{ik} + \varepsilon_{i1}$$

$$\log T'_i = t'_0 + t'_1 x_{i1} + \ldots + t'_k x_{ik} + bM_i + \varepsilon_{i2}$$

$$M_i = a_0 + a_1 x_{i1} + \ldots + a_k x_{ik} + \varepsilon_{i3}$$

where $x_{i1} \ldots x_{ik}$ are the $k$ covariates for subject $i$, $M_i$ is the mediator, $\varepsilon_{i1}$, $\varepsilon_{i2}$ and $\varepsilon_{i3}$ are the error terms or residuals, $t_0 \ldots t_k$ are the regression coefficients relating covariates
to log survival time for subject $i$, $t_0 \ldots t_k$ are the coefficients relating covariates to log survival time adjusted for the mediator for subject $i$, $a_0 \ldots a_k$ are the regression coefficients relating covariates to the mediator for subject $i$, and $\sigma$ is the scale parameter.

The log-hazard form of the model is estimated by the PHREG procedure in SAS as

$$\log h(t) = \alpha(t) + \beta_1 x_{i1} + \ldots + \beta_k x_{ik}$$  \hspace{1cm} (13)

$$\log h(t) = \alpha(t) + \beta'_1 x_{i1} + \ldots + \beta'_k x_{ik} + b'M_i$$  \hspace{1cm} (14)

$$M_i = \alpha(t) + a'_1 x_{i1} + \ldots + a'_k x_{ik}$$  \hspace{1cm} (15)

where $\alpha(t) = \log \lambda_0(t)$ and $\lambda_0(t)$ is a baseline hazard function for any individual whose covariates all have values of 0. The interpretations of the rest parameters are similar as they are in the log-survival function. For the Weibull model, the relationship between the parameters in Equations (10) and (13) can be equated as $\beta_k = -t_k / \sigma$ [22].

The results show that, under the condition of no censored data, the assumption that mediated effects calculated by the difference in coefficients method ($t-t'$) (or ($\beta - \beta'$) as expressed in log-hazard function) and those calculated by the product of coefficients method ($ab$) (or $a'b'$ as expressed in log-hazard function) are identical does apply to log-survival time-survival analyses but not to log-hazard time-survival analyses. In the LIFEREG procedure, the differences between two methods were only close to zero. However, the differences between $ab$ and $t-t'$ from the PHREG procedure were quite different, with the mediation estimates for $|t-t'|$ being consistently smaller than those for $|ab|$, and the difference decreased as the sample size got larger.
As stated previously, our proposed research is interested in examining mediation effects in the context of survival analysis, particularly as it relates to unintentional weight loss being a mediator of certain nutritional risk factors for mortality. As shown in the literature, mediation analyses have been studied extensively with normally distributed continuous dependent variables. Current methodology also exists to examine the equivalence of two calculations of mediated effects, assuming no censored data and time-invariant mediator in survival analysis, it is necessary to improve this existing method by improving its assumption and model as well to estimate the mediated effect.

To facilitate the research, it is necessary to develop procedures which will perform the calculations for estimating mediated effects with a high degree of accuracy. Mediating variables are often assumed to be continuous and follow a normal distribution, whereas risk factor variables can take continuous or categorical values. In the simulation, we also assume that there are censored cases. Currently in this proposal, we plan to use SAS to do the simulation, so the details of the simulation with censored cases from SAS will form the topic of Chapter 3.

There are several ways in which the existing method can be improved. We believe that assuming the censored data in the simulation might lead us to a situation closer to reality. Note an observation is said to be censored if we have some information about an individual survival time and we know that the target event has not occurred prior to a censoring event, such as the end of the study or the participant being lost to follow-up. Right censoring, in particular, refers to that situation where the target event does not occur until after the period of observation of the study. It is especially common in studies
of mortality. In Tein and MacKinnon’s study [21], the fact of no censored data made Equation (2) \(Y = i + tX + e\) be readily estimated with ordinary least squares by simply creating the dependent variable as the logarithm of the event time (i.e., \(Y = \log Z\)). However, if we include censored data in the analysis, we would consider the shape of the distribution of the error term. With this more general situation, we will be able to extend the study to reflect research situations more likely encountered in applied analyses. The performance of examining the equivalence of two methods to calculate mediated effects under each of these modifications will be reviewed. We also propose to apply this improved method to a population of older adults to investigate the relationship between undereating and mortality through the mediating effect of unintentional weight loss.

The dissertation is organized as a collection of three manuscripts. The first manuscript, entitled “Reliability and Predictive Validity of Caloric Intake Measures from the 24-Hour Dietary Recalls of Homebound Older Adults,” examines whether homebound older adults provide reliable and valid measures of total caloric intake in 20-hour dietary recalls. The predictive validity of this measure on observed weight change represents the \(a\) pathway of our mediation model. The second manuscript, entitled "A Simulation Study of Mediation Effect Estimation in Survival Analyses with Censored Data," conducts simulations to examine the effect of undernutrition on mortality through the mediating effect of weight loss in the context of survival models. Our simulations will yield evaluations for the mediated effect in terms of bias, standard error, and mean square error in analyses with censored events. The patterns of the evaluated parameters from point estimates will be summarized and compared across different methods. The third manuscript, entitled "The Multivariate Determinants of Unintentional Weight Loss and
Potential Mediation Effect of Weight Loss on Mortality in Community-Dwelling Older Adults, "examines additional predictors of observed weight loss over 4 years for a group of community-dwelling older adults and examines the relationships between those elements and mortality to check if the mediated effect of weight loss exists. An integrative discussion section then highlights recommendations for further observational studies in the areas of nutrition and weight loss in vulnerable older adult populations.
Figure 1. Models showing total effect and mediated effect of $X$ on $Y$

Part 1

- Independent Variable $X$ → $e_1$ → Dependent Variable $Y$
- Caloric Intake Discrepancy → Log Odds of Mortality

Part 2

- Independent Variable $X$ → $a$ → Mediating Variable $M$ → $b$ → Dependent Variable $Y$
- Caloric Intake Discrepancy → $e_2$ → Log Odds of Mortality
- Unintentional Weight Loss → $e_3$ → Mediating Variable $M$
Figure 2. Model showing relationship between undereating and UWL.
Figure 3. Model showing relationships among undereating, UWL, and mortality.
Figure 4. Model showing relationship among potential elements, UWL, and mortality
Introduction

Insufficient caloric intake in older adults is a serious problem and is related to unintentional weight loss, functional decline, morbidity, mortality, and quality of life [1-12]. It is important that valid, reliable methods be used to assess caloric intake, especially in homebound older adults, who are particularly vulnerable to experiencing undereating.

Twenty-four hour dietary recalls are a widely used methodology for collecting data on eating behaviors and measuring caloric intake, including in homebound older adults [13-17]. There are several advantages to the 24-hour recall [18]. Thompson and Subar note that because responses are recorded by an interviewer, high levels of physical functioning and literacy of participants are not required. This reduces the potential for both nonresponse bias and respondent burden. Further, because the recall period is immediate, participants are able to recall most of the foods and beverages consumed in the preceding day. A final advantage pointed out is that the likelihood that these self-reports will interfere with typical dietary behavior is minimized because the food is consumed before dietary recalls occur.

There are also potential disadvantages to the use of 24-hour dietary recalls that frequently call into question their reliability or validity, especially among older adults [18]. Because of knowledge, memory problems, or factors associated with the interview situation, individuals may not reliably report their food intake [18]. Some research shows
that the validity of dietary information collected from both older adults and younger adults are comparable to each other; and that recall can be enhanced by supplementing standard methods with memory strategies and probes [19,20]. Studies conducted specifically with older adults generally have found that 24-hour recall estimates are similar to observed intakes [21,22]. Similar to younger populations, those with lower actual intakes tend to over-report caloric intake, whereas those with higher actual intakes tend to under-report [21,22]. We are unaware of any research that has examined the reliability or validity of self-reported caloric intake among homebound older adults—a population in which dietary recalls are frequently used. Assessment of reliability and validity is necessary to estimate the extent of bias of the 24-hour dietary recall in this population.

Reliability refers to the extent to which a measure yields the same results under similar circumstances; while validity refers to the extent to which a measure actually measures what it is designed to measure [23]. Test-retest reliability is commonly used to estimate the reliability or stability of a measure by comparing the results of the same measure at two points in time. Test-retest reliability is an especially important consideration when data will be collected longitudinally to assess changes in intake over time. Predictive validity is a type of validity especially important in assessing 24-hour dietary recalls. Predictive validity evaluates whether a measure predicts an outcome that it would be expected to predict. In this study, we would expect that not consuming enough calories would predict subsequent adverse events such as weight loss.

The purpose of this study was to evaluate the 6-month test-retest reliability and predictive validity of caloric intake measures from the 24-hour dietary recalls of
homebound older adults. The study focused on 1) the test-retest reliability of self-reported caloric intake in order to examine the consistency of the measurement from baseline to 6-months, and 2) the predictive validity of 24-hour dietary recalls by evaluating the effect of caloric intake deficiencies at baseline on observed weight loss and the occurrence of adverse events during the 6-month follow-up interval.

2 Methods

2.1 Sample and Design

Homebound older adults were enrolled in a longitudinal study designed to assess eating behaviors and the factors and health outcomes associated with those eating behaviors. The study protocol was reviewed and approved by the UAB Institutional Review Board. The sample and design are described in-depth in other papers [17,24,25]. Briefly, the eligibility of being included in the study was that participants were community dwelling, receiving Medicare home health services, able to communicate verbally with interviewers in English, free of significant cognitive impairment, free of terminal illness, and not receiving nutrition through a feeding tube.

Two hundred and thirty participants were assessed at baseline, and 173 participants completed a 6-month follow-up assessment. 24-hour dietary recalls were collected at both assessments, and the actual weight of the participants was measured at both assessments for those who were able to stand.

Two sub-samples were drawn for the analyses. The first consisted of 52 participants who had actual measured weight at both interviews and who had not been hospitalized between the two assessments. Because recent hospitalization was found to
be a strong predictor of undereating in the baseline analysis [17], participants who were hospitalized were excluded from this sub-sample in order to eliminate this known source of variability in the reliability analysis. This sub-sample was used to evaluate the test-retest reliability of caloric intake and the predictive validity of caloric intake deficiencies on actual observed weight loss over a 6-month period. This group was selected because these participants should have had relatively stable diets across the time period, making them good candidates for examining test-retest reliability.

The second sub-sample consisted of 143 participants who had measured weight at both interviews or had any adverse outcomes (i.e., weight loss of 2.5% or more of baseline body weight, hospitalization, long-term care institutionalization, or mortality) between baseline and the 6–month assessment. This sub-sample was used to evaluate predictive validity only. Because weight loss is not the only adverse outcome that participants may experience as a consequence of undereating, this larger sub-sample may be more a more sensitive and comprehensive way to evaluate predictive validity because of its inclusion of other important competing hazards.

2.2 Procedure and Measures

24-hour dietary recalls. During structured interviews in the participants’ homes, 24-hour dietary recalls were conducted using standard protocols [26]. Interviewers were formally trained in the 24-hour recall methods at the Nutrition Coordinating Center at the University of Minnesota. The recalls were conducted using standardized probing questions, two-dimensional food models to estimate portion size, and a multiple-pass methodology. The interviewer inspected refrigerators and kitchen storage space to better
determine foods actually eaten by participants and the materials in which they were prepared or consumed; and, if necessary, caregivers provided supplementary information. Participants were contacted by telephone two more times over the next 2 weeks to obtain two additional 24-hour dietary recalls, one of which was obtained for a weekend day. Mean daily caloric intake was obtained by taking the average of the three 24-hour dietary recalls.

*Caloric intake discrepancy and undereating.* Caloric intake discrepancy and undereating were defined as the difference between a participant’s mean daily caloric intake and his/her Estimated Energy Requirements (EER) at baseline. Caloric intake discrepancy is the actual difference measured as a continuous variable and undereating was a dichotomized variable (described below). EER is computed using a formula according to the Institute of Medicine [27]. For females, the formula is:

\[
\text{Energy (kcal)} = 354.1 - (6.91 \times \text{Age}[y]) + \text{Physical Activity Coefficient}[1 \text{ for sedentary}] \times (9.36 \times \text{Weight}[kg] + 726 \times \text{Height}[m]);
\]

For males, the formula is:

\[
\text{Energy (kcal)} = 661.8 - (9.53 \times \text{Age}[y]) + \text{Physical Activity Coefficient}[1 \text{ for sedentary}] \times (15.91 \times \text{Weight}[kg] + 539.6 \times \text{Height}[m]).
\]

Depending upon the analyses, the difference between mean daily caloric intake and EER was analyzed as a continuous variable reflecting caloric intake discrepancy or was coded as a dichotomous variable, undereating (yes/no), indicating whether enough caloric was consumed by a person to maintain his/her current body weight.
Weight Loss. Weight change was derived by taking the difference between a participant's 6-month weight and their actual baseline weight. Weight change was analyzed both as a continuous variable in kilograms and as a dichotomous variable, indicating whether a participant experienced weight loss that was equal to or greater than 2.5% of their baseline weight.

Other Adverse Outcomes. Adverse outcomes included the participant experiencing any one of the following events between baseline and the 6-month assessment: hospitalization, institutionalization (either nursing home or assisted living facility), mortality, or loss of weight of 2.5% or more of baseline body weight. This was coded as a dichotomous variable.

Control Variables. We controlled for body mass index (BMI), comorbidity, age, gender, and ethnicity in our analyses. BMI was assessed by obtaining height and weight on all participants who were able to stand (55% of 230 participants) and categorized according to NHLBI Clinical Guidelines [28]. Self-report of height and weight was obtained for those who were unable to stand. Comorbidity was assessed using the Charlson Comorbidity Index [9-31]. Age, gender, and ethnicity were assessed by self-report.

2.3 Statistical Methods

Data analyses were performed using SAS version 9.1. Descriptive statistics were completed first to characterize the sub-samples.
Test-retest reliability was assessed with the Pearson product-moment correlation between the means of the 24-hour dietary recalls at baseline and the 6-month assessment using the first sub-sample of 52 participants.

Predictive validity of the 24-hour dietary recall was tested initially in a Pearson’s correlation analysis by comparing caloric intake discrepancy and weight loss, and then, by generalized linear models (GLM) and logistic regression models that assessed the effects of undereating on weight loss of 2.5% or more, while controlling for other variables. These analyses were performed for the sub-sample of 52 participants. The numeric value of observed weight change (6-month weight minus baseline weight) was used in the GLM procedure, whereas a categorical weight loss variable of 2.5% or more of baseline body weight was used in the logistic regression models.

Next, logistic regression analyses were conducted to assess the multivariate effects of undereating on adverse outcomes, while controlling for other variables. These analyses were performed using the second analytic sub-sample of 143 participants.

3 Results

3.1 Descriptive statistics

Table 1 presents baseline characteristics for the two analytical samples. The subgroup of 52 participants was similar to the group of 143 participants. However, the group of 143 had a higher average caloric intake discrepancy than the group of 52. At the 6-month follow-up, 25% of the 52 participants and 37.8% of the 143 had weight loss of 2.5% or more of baseline body weight. Additionally, for the 143 participants, 28 had been hospitalized, 5 were institutionalized, and 56 died.
3.2 Test-Retest Reliability

Pearson’s correlation coefficient shows a statistically significant linear relationship between caloric intake at baseline and 6 months for the subgroup of 52 (r = 0.59, p<0.0001).

3.3 Predictive Validity

A low and non-significant correlation was observed between caloric intake discrepancy (baseline caloric intake minus EER) and weight loss for the subgroup of 52 (r = 0.08, p=0.58). Table 2 presents the relationships between weight loss and undereating for the subgroup of 52 participants. In both regression analyses, undereating was not associated with observed weight loss.

Table 3 presents the multivariate effects of predictors on adverse outcomes for the subgroup of 143 participants. While controlling for other variables that might affect outcomes, participants who were undereating were 3.5 times more likely to experience an adverse outcome (p=0.009) compared with those who were not undereating.

4 Discussion

Our findings indicate that there exists substantial test-retest reliability over a 6-month time period for total caloric intake measures obtained from homebound older adults using 24-hour dietary recalls (r=0.59). This relationship was true for participants who were not hospitalized and who had two measured weights at both baseline and the 6-month assessment. These findings are consistent with previous reports in the literature of
young and middle-aged adults, including studies reporting test-retest correlations ranging from 0.40 to 0.92 [32-35].

Concerning predictive validity, we hypothesized that a relationship would exist between caloric intake discrepancy (caloric intake minus EER) and observed weight loss in homebound older adults. Because of their frailty and vulnerable health status, the criteria of actual measured weight and non-hospitalization were used to identify a subgroup that we speculated may have been a more reliable group in which to study the predictive validity of self-reported caloric intake. Our findings, however, revealed a low correlation between caloric intake discrepancy and actual weight loss in the sub-sample of 52 participants. Further, there was no association between caloric intake discrepancy and weight loss in either the logistic regression or GLM analyses for this sub-sample. Although the odds ratio of 2.06 in the logistic regression analysis was in the predicted direction, the p-value was not significant.

Because those who were unable to be weighed or who had been hospitalized were excluded from the smaller sub-sample, our power to detect predictive effects may have been limited by the decision to exclude those who were more frail or more ill. This assumption was confirmed by our analyses conducted on the larger sub-sample of 143, which combined mortality with weight loss, hospitalization, and institutionalization as a more complex composite adverse outcome measure. In these analyses, undereating was clinically and statistically associated with adverse outcomes.

This study is the first to examine the reliability and predictive validity of caloric intake measures from the 24-hour dietary recalls of homebound older adults. While the reliability results are encouraging, the validity analyses were limited by the relatively
small sample size. Our other analyses and findings have suggested that future studies might use larger sample sizes, with better control of caloric intake underreporting, to examine the relationship between weight loss and caloric intake discrepancy in longitudinal studies with more repetitive observations. Future studies might also investigate further the effects of undereating on mortality or other adverse outcomes through the intervening effect of undereating on unintentional weight loss.
References


12. Institute of Medicine, Committee on Nutritional Services for Medicare Beneficiaries. The role of nutrition in maintaining health in the nation’s elderly: Interventions and assessments can help beneficiaries. Washington, DC: National Academy Press; 2001.


Table 1. Descriptive statistics for subgroups of participants

<table>
<thead>
<tr>
<th>Variables</th>
<th>Subgroup (Mean (SD) or N (%))</th>
<th>N=52&lt;sup&gt;a&lt;/sup&gt;</th>
<th>N=143&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td>81.0 (8.5)</td>
<td>79.5 (8.7)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>38 (73.1)</td>
<td>106 (74.1)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>14 (26.9)</td>
<td>37 (25.9)</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>23 (44.2)</td>
<td>49 (34.3)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>29 (55.8)</td>
<td>94 (65.7)</td>
<td></td>
</tr>
<tr>
<td>Charlson Comorbidity Index</td>
<td>3.1 (2.2)</td>
<td>3.7 (2.7)</td>
<td></td>
</tr>
<tr>
<td>Weight at Baseline</td>
<td>72.7 (19.8)</td>
<td>76.7 (22.6)</td>
<td></td>
</tr>
<tr>
<td>Weight at 6-month</td>
<td>73.7 (20.0)</td>
<td>75.0 (22.3)</td>
<td></td>
</tr>
<tr>
<td>Caloric Intake at Baseline</td>
<td>1672 (493.9)</td>
<td>1527 (469.8)</td>
<td></td>
</tr>
<tr>
<td>Caloric Intake at 6-month</td>
<td>1653 (464.6)</td>
<td>1633 (494.4)</td>
<td></td>
</tr>
<tr>
<td>Body Mass Index at Baseline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>3 (5.8)</td>
<td>10 (7.0)</td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>26 (50)</td>
<td>62 (43.4)</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>12 (23.1)</td>
<td>33 (23.1)</td>
<td></td>
</tr>
<tr>
<td>Obese Class I</td>
<td>2 (3.9)</td>
<td>15 (10.5)</td>
<td></td>
</tr>
<tr>
<td>Obese Class II or III</td>
<td>9 (17.3)</td>
<td>23 (16.1)</td>
<td></td>
</tr>
<tr>
<td>Body Mass Index at 6-month</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>2 (3.9)</td>
<td>6 (5.0)</td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>24 (46.2)</td>
<td>56 (46.7)</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>12 (23.1)</td>
<td>24 (20.0)</td>
<td></td>
</tr>
<tr>
<td>Obese Class I</td>
<td>5 (9.6)</td>
<td>18 (15.0)</td>
<td></td>
</tr>
<tr>
<td>Obese Class II or III</td>
<td>9 (17.3)</td>
<td>16 (13.3)</td>
<td></td>
</tr>
<tr>
<td>Caloric Intake Discrepancy (kcal/day)</td>
<td>-84 (559.6)</td>
<td>-275 (614.6)</td>
<td></td>
</tr>
<tr>
<td>UnderEating</td>
<td>31 (60)</td>
<td>97 (68)</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Participants who had actual measured weight at both baseline and 6-month and who had not been hospitalized between baseline and six-month

<sup>b</sup> Participants who either had any adverse outcomes between baseline and 6-month or had measured weight at both times
Table 2. Generalized linear and logistic regression models of weight loss for the subgroup of 52 participants

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Generalized Linear Model</th>
<th>Logistic Regression Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>p</td>
</tr>
<tr>
<td>Undereating (Yes vs. No)</td>
<td>-0.316</td>
<td>0.808</td>
</tr>
<tr>
<td>BMI at Baseline</td>
<td>-0.083</td>
<td>0.378</td>
</tr>
<tr>
<td>Gender (Male vs. Female)</td>
<td>0.803</td>
<td>0.550</td>
</tr>
<tr>
<td>Ethnicity(AA vs. White)</td>
<td>2.405</td>
<td>0.046*</td>
</tr>
<tr>
<td>Age</td>
<td>-0.017</td>
<td>0.802</td>
</tr>
</tbody>
</table>

* p<0.05
Table 3. Logistic regression models for multivariate effects of predictors on adverse outcomes \(^a\) for the subgroup of 143 participants

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate</th>
<th>Odds-ratio</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undereating (Yes vs. No)</td>
<td>1.251</td>
<td>3.49</td>
<td>1.37-8.91</td>
<td>0.009**</td>
</tr>
<tr>
<td>Gender (Male vs. Female)</td>
<td>-0.393</td>
<td>0.68</td>
<td>0.25-1.80</td>
<td>0.432</td>
</tr>
<tr>
<td>Ethnicity (AA vs. White)</td>
<td>-1.126</td>
<td>0.32</td>
<td>0.13-0.81</td>
<td>0.016*</td>
</tr>
<tr>
<td>Age</td>
<td>-0.019</td>
<td>0.98</td>
<td>0.93-1.03</td>
<td>0.453</td>
</tr>
<tr>
<td>Charlson Comorbidity Index</td>
<td>0.096</td>
<td>1.10</td>
<td>0.93-1.30</td>
<td>0.253</td>
</tr>
<tr>
<td>BMI Category</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;18.5 vs.18.5-25</td>
<td>0.718</td>
<td>2.05</td>
<td>0.39-10.89</td>
<td>0.400</td>
</tr>
<tr>
<td>25-30 vs.18.5-25</td>
<td>0.252</td>
<td>1.29</td>
<td>0.46-3.61</td>
<td>0.633</td>
</tr>
<tr>
<td>30-35 vs.18.5-25</td>
<td>1.047</td>
<td>2.85</td>
<td>0.51-15.80</td>
<td>0.231</td>
</tr>
<tr>
<td>&gt;35 vs.18.5-25</td>
<td>-0.581</td>
<td>0.56</td>
<td>0.16-1.97</td>
<td>0.366</td>
</tr>
</tbody>
</table>

\(^a\) Adverse outcomes included death, weight loss of 2.5% or more of baseline body weight, hospitalization, longer-term care institutionalization

\(^*\) p<0.05

\(^**\) p<0.01
A SIMULATION STUDY OF MEDIATION EFFECT ESTIMATION IN SURVIVAL ANALYSES WITH CENSORED DATA

1 Introduction

Mediation occurs when an independent variable causes a mediator and as a result, the mediator causes a dependent variable [1-2]. In the case when the mediation estimate process is available, mediation analysis can incorporate the information in the study [2]. Mediating variables can be psychological, such as attitudes; behavioral, such as social skills; or biological, such as serum cholesterol levels. Mediators must be able to show some change over time in response to the independent variable. In our proposed study, we are interested in modeling the influence of undernutrition among the elderly on mortality through the mediated effect of unintentional weight loss.

Figure 1 part 1 shows the elements of the two-variable relations, with \( t \) representing the total effect between \( X \) and \( Y \). Part 2 shows the elements of the single-mediator model after introducing a mediator \( M \), where \( a \) is the coefficient between \( X \) and \( M \), \( b \) is the coefficient between \( M \) and \( Y \) when \( X \) is held constant, and \( t' \) is the coefficient between \( X \) and \( Y \) when \( M \) is held constant. \( e_2 \) and \( e_3 \) are error terms or residuals.

Equations (1)-(3) show how to estimate the coefficients in the single mediator model from Figure 1.

\[
M = i_1 + aX + e_1 \quad (1)
\]
\[
Y = i_2 + tX + e_2 \quad (2)
\]
\[
Y = i_3 + t'X + bM + e_3 \quad (3)
\]
In these equations, \( i_1 \) and \( i_2 \) and \( i_3 \) constant intercepts, \( Y \) is the dependent random variable, \( X \) is the independent random variable, \( M \) is the random mediating variable, \( t \) is the constant coefficient between \( X \) and \( Y \) before accounting for the mediator (total effect), \( t' \) is the constant coefficient relating \( X \) to \( Y \) after adjusting for the mediator (indirect effect), \( b \) is the constant coefficient between \( M \) and \( Y \) after adjusting for \( X \), \( a \) is the constant coefficient between \( X \) and \( M \), and \( e_1 \), \( e_2 \) and \( e_3 \) are random error or residual terms. When \( Y \) is a continuous random variable, it is typical to assume that the residual terms \( e_i \)s have normal distributions with a mean and variance that is constant over observations and that \( e_i \)s are independent of each other and across observations. When \( Y \) is a dichotomy with possible values of 1 or 0, the assumptions of homoscedasticity of variance and normality distribution for the residual terms do not apply any more, while the rest of the assumptions still hold [3].

There are two ways to calculate the mediated effect in a single-mediator model, as either the product of coefficients \( ab \) or the difference of coefficients \( t - t' \) [4]. The difference of coefficients method corresponds to the reduction in the predictor before and after adjusting for the mediator. The product of coefficients method describes a process regarding to what extent the predictor affects the mediator, \( a \), and to what extent the mediator affects the outcome variable, \( b \) [5].

As described above, the \( t - t' \) and \( ab \) are two computational methods to calculate the mediated effect. These two methods of mediation are theoretically equal when the dependent variable and \( M \) are continuous. The proof, as detailed in a paper by MacKinnon and his colleagues [6], is based on computing the \( \text{Cov}(Y,X) \) and \( \text{Cov}(M,X) \) using Equations (1), (2), and (3). Simultaneously solving these three equations, the
authors showed that \( t-t'=ab \) with the assumption that \( X \) is independent of \( e_1, e_2 \) and \( e_3 \). Note that this result states that the parameters are equal but the estimators are not necessarily equal, because the estimators are random, but should be close to each other if this model holds.

For multilevel models with binary dependent variables, \( ab \) is not always equal to \( t-t' \) due to the fact that the residual variance is not a constant in the logistic regression so the coefficients of \( t \) and \( t' \) come from equations on different scales [4,7]. MacKinnon and colleagues [8] conducted a simulation study in the model with binary outcomes, and their findings indicated that the asymptotically unbiased estimates can be directly produced by the \( ab \) method, while the \( t-t' \) method must be scaled properly before it is used to estimate mediated effect.

In order to examine whether \( t-t' \) equals to \( ab \) with survival analysis in which the dependent variable is the time to the occurrence of an event, Tein and MacKinnon [9] examined the estimates of the mediated effect \( ab \) and \( t-t' \) in the log-survival time (i.e., LIFEREG procedure) and log-hazard time (i.e., PHREG procedure) models. With the assumptions of no censored cases, time-invariant mediating variable and no ties among the events, they found that the approximate equivalence of two mediation methods does apply to log-survival time survival analyses but not to log-hazard time survival analyses (Cox’s proportional hazard model). In the LIFEREG procedure, the differences between two methods were approximately equal to zero. However, the differences between \( ab \) and \( t-t' \) from the PHREG procedure were quite different, with the mediation estimates for \( |t-t'| \) being consistently smaller than those for \( |ab| \). And the difference decreased as the sample size got larger. In reviewing the assumptions used to show the equality of \( ab \) and \( t-t' \), the
assumptions of the models in Equation (2) and (3) also apply to the log-survival time model. Therefore, we see that \( ab = t - t' \) theoretically for log-survival time model as well. However, in log-hazard time model, where we are modeling the hazard function it may well be the case that this equivalence no longer applies.

Because censored survival data are very common and censoring requires special treatment, it is necessary to conduct a simulation study to extend Tein and MacKinnon’s [9] study to assess the estimates of mediated effects in the context of survival analysis with certain proportions of censored cases. Censoring comes in different forms and occurs for different reasons. There are left censoring and right censoring. Left censoring is most likely to occur when one begins observing a sample at a time when some of the individuals may have already experienced the event. Right censoring is far more common than left censoring and happens because observation is terminated before the target event occurs [3].

Note that if \( \hat{a}, \hat{b}, \hat{t}, \) and \( \hat{t}' \) are the maximum likelihood estimators (MLE) of \( a, b, t \) and \( t' \), respectively, then by invariance property of MLE, \( \hat{a} \hat{b} \) and \( \hat{t} - \hat{t}' \) are the MLE of \( ab \) and \( t - t' \). Furthermore, under the ordinary least squares (OLS) with normal error terms, \( \hat{a}, \hat{b}, \hat{t}, \) and \( \hat{t}' \) are each normally distributed, as well as \( \hat{t} - \hat{t}' \). Because the product of two normal random variables is not itself normally distributed in most circumstances [10-11], \( \hat{a} \hat{b} \) does not necessarily have a normal distribution. Even in the case where the mean of two normal random variables is zero, their product is symmetrically distributed with a kurtosis of six [12]. When the product is nonzero, the distributions continue to have
excess kurtosis and are skewed. However, as the ratio of at least one variable’s mean to its standard error increases, the distribution approaches normality [13-14].

The purpose of this proposed study is to investigate the properties of the estimates of $ab$ under the log-survival model. There are several methods for computing the standard errors of a product of two random variables [6]: (Method 1) first-order Taylor series [15], (Method 2) second-order Taylor series formula [16], (Method 3) unbiased standard error formula [17], and (Method 4) empirical standard errors generated through resampling methods [6,9,18]. These methods have been applied to computing confidence intervals and testing the statistical significance of mediated effects (i.e., $z = \alpha \beta / se_{\alpha \beta}$) with OLS [6] and using survival analyses with no censored data [9]. The results indicate, for both kinds of data, there are no significant differences in the standard errors among three formulaic methods. However, with the survival data, the empirical $se_{\alpha \beta}$ tends to be larger than those from the three formulaic methods, while with the OLS estimator, results from the three methods are consistent with the empirical standard errors. It is unknown whether these conclusions apply to survival analysis with censored data. MacKinnon et al. [19] showed evidence that because of the nonnormal and sometimes of the high kurtosis distribution of coefficients $a$ and $b$, the $ab / \sigma_{ab}$ methods from the first-order Taylor series have lower power to test the significance of $ab$. Thus, we will also examine the consistency and accuracy of the first order Taylor series method [15], the unbiased formula [17] and the empirical standard error method for the standard errors using survival analyses.

As we plan to examine the equivalence of $ab$ and $t-t'$ with censored data, we would like to compare the properties of two methods in particular the log-survival time
model. The reason is that results from Tein and MacKinnon’s [9] simulation show a very stable scenario for LIFEREG procedure. The extension of their studies by introducing the censored data is a necessity to add a further complexity to the model but not to involve too many unstable parameters that could probably compromise the results. Therefore, in this paper, the property of $ab$ will be examined in terms of bias, standard error, and mean square error. In addition, the two methods will be compared to the true value of $ab$ under different specifications. Bias is defined as the difference between the point estimate and the true value of $ab$. Mean square error is computed as the sum of squared standard deviation and squared bias, which is a measure of both precision and accuracy of the estimate. The mortality rates, hazard rates, and censoring rates are manipulated in order to examine how the changes affect the characteristics of $ab$ and $t-t'$.

2 Overview of Simulation Study

The aim of this article was to conduct a simulation study to extend Tein and MacKinnon’s [9] study to assess the estimates of mediated effects in the context of survival analysis with censored data. The primary focus of our study was to investigate the property of $ab$ in terms of bias, standard error (SE) and mean square error (MSE). The characteristics of $ab$ were examined in different scenarios considering mortality rates, hazard rates and censoring. The mediated effects calculated from $ab$ and $t-t'$ were compared to the true value of $ab$. The comparison between $ab$ and $t-t'$ and the different methods of estimating the standard errors were also investigated under each of the different scenarios.
We hypothesize that both \( ab \) and \( t-t' \) would produce similar estimates and have smaller standard errors as the amount of censoring decreases. A main question being investigated by the simulation was whether the choice of \( ab \) or \( t-t' \) could lead to different inferences and conclusions when applied to the censored data and whether the behavior of two methods would change with the different hazard and mortality rates.

3 Methods

All statistical simulations and analyses were conducted by the SAS (Version 9.1) programming language. Data were simulated following a single-mediator pathway where the relationship between the independent X and dependent variable Y was mediated by a third variable M. The independent variable, X, and mediating variable, M, were simulated from a standard normal distribution with mean 0 and variance 1. Survival time, Z, was simulated based on the Weibull distribution using the equation [20],

\[
Y = \log Z_i = u + t' X_i + bM_i + \varepsilon
\]

where Y is the log of \( Z_i \), the time to event for subject \( i \), and \( X_i \) and \( M_i \) are the independent and mediating variables, respectively. The residual term is the product of \( \varepsilon \) and \( \sigma \), where \( \varepsilon \) follows a 2-parameter extreme value distribution, and \( \sigma \) is the scale parameter determining the shape of the hazard rate [3,20]. The extreme value distribution is a unimodal distribution and not symmetrical. The hazard rate determined by scale parameter can be increasing, decreasing or constant, which makes the Weibull model very flexible. A special conversion between the Weibull model and the accelerated failure-time model makes the Weibull model different from the other parameter models, as it has a unique characteristic of a proportional hazards representation [20].
We estimated the Weibull model with a log-survival time model with covariates and censored data. The log-survival time of the model can be estimated by LIFEREG procedure in SAS as

\[
M_i = a_0 + a x_i + \varepsilon_{i1}
\]

(5)

\[
\log Z_i = t_0 + tx_i + \sigma \varepsilon_{i2}
\]

(6)

\[
\log Z_i = t_0' + t' x_i + b M_i + \sigma \varepsilon_{i3}
\]

(7)

where \(x_i\) is the covariate for subject \(i\), \(M_i\) is the mediator, \(\varepsilon_{i1}\), \(\varepsilon_{i2}\) and \(\varepsilon_{i3}\) are the measurement errors, \(t_0\) and \(t\) are the coefficients relating the covariates to the log survival time for subject \(i\), \(t_0'\) and \(t'\) are the coefficients relating the covariates to the log survival time adjusted for the mediator for subject \(i\), \(a_0\) and \(a\) are the coefficients between \(X\) and \(M\) for subject \(i\), and \(\sigma\) is the scale parameter. Equation (5) is a simple, linear regression model, so it is typical to assume that \(\varepsilon_{i1}\) has a normal distribution. However, the log survival time here in Equations (6) and (7) allow for non-normal distributions. Because of the possibility of a non-normal error term, the distribution of the variable \(Z\) needs to be specified for the LIFEREG procedure [9]. The parameters are estimated with maximum likelihood. Models with log-survival time as dependent variable examine mediation processes of increase or decrease in the expected survival time.

The data were generated based on Equation (4), where \(X\) and \(M\) followed a standard, normal distribution and were generated using the RANNOR function with zero as the seed. The values of the parameters specified in the model were derived from multiple sources, including a population-based, longitudinal study of mobility and mortality among community-dwelling older adults in the University of Alabama at
Birmingham Study of Aging, where 30% of the 1000 subjects were observed to be deceased in 6 years as of this date.

3.1 Varying Hazard Rates

The cumulative density function (PDF) of the Weibull distribution is

\[ F(z) = 1 - \exp(-\lambda z^\alpha) \]  

where \( \alpha \) is the shape parameter and \( \lambda \) is the scale parameter. In this case, when \( \alpha > 1 \), the hazard rate is increasing; when \( \alpha < 1 \), the hazard rate is decreasing; and when \( \alpha = 1 \), the hazard rate is constant. The intercept term in Equation (4) was determined using the formula \( -\ln(\lambda)/\alpha \) \[20\]. The residual term was determined by taking the product of \( \sigma \) which was derived from formula of \( (\sigma = \alpha^{-1}) \) \[20\] and a random data generated from the extreme value variable with location parameter equal to \( -(\ln(\lambda)/\alpha \text{-intercept})/\sigma = 0 \), and the mode scale parameter was equal to \( 1/(\sigma \times \alpha) = 1 \) \[9\].

In each simulation, \( \alpha \) takes the values of 0.5, 1, and 2, from which \( \lambda \) takes the values of 0.1456, 0.0594, and 0.0099, respectively. The intercept and residuals sets are \((u=3.85, \sigma=2)\), \((u=2.82, \sigma=1)\), and \((u=2.31, \sigma=0.5)\), respectively. A dummy variable was generated from a Bernoulli distribution using RAND(‘BERNOULLI’, \( p \)) that will be associated with each failure time where 1 means that observation is not censored and 0 means the observation is censored. Therefore, \( p \) denoted the proportion of observation that is not censored. The censoring proportion was set at 70% and the mortality rate at 30% for this set of simulation studies.
3.2 Varying Mortality Rates

In order to change the mortality rate, the probability of 30% death, 50% death and 80% death in 6 years were varied in the simulation. We set the censoring at 70% and \( \alpha = 2 \) so that the hazard rate increases. The use of Equation (8) in this simulation set, a probability of 30% death in 6 years, is given by

\[
0.3 = 1 - \exp(-\lambda 6^\alpha)
\]  
(9)

The corresponding \( \lambda = -\ln(1 - \text{mortality rate}) / 6^\alpha \) associated with the three different mortality rates when \( \alpha = 2 \) are 0.0193, 0.0447, and 0.0029. The intercept and residuals sets are \((u=1.98, \sigma = 0.5), (u=1.55, \sigma = 0.5), \) and \((u=2.92, \sigma = 0.5)\), respectively.

3.3 Varying Percent Censored

The probability of censoring \((1 - p)\) from Bernoulli distribution was varied with values of 0, 20%, 50%, and 70%. With \( \alpha = 2 \) (increasing hazard rate) and mortality rate of 30%, the corresponding intercept \( u = -\ln(\lambda)/\alpha \) is 2.31 (\( \lambda = -\log(1 - \text{mortality rate}) / 6^\alpha \)) and \( \sigma = 1/\alpha \) is 0.5.

There are some common parameters set up for all the three sets of simulations described above. For each simulation, only one parameter combination is considered for \( a, b, \) and \( t' \). A total of four combinations of parameters were generated: \((a=0.14, b=0.14, t'=0.14), (a=0.14, b=0.26, t'=0.14), (a=0.14, b=0.39, t'=0.14), \) and \((a=0.14, b=0.59, t'=0.14)\). Values 0.14, 0.39 and 0.59 were chosen because they corresponded to Cohen’s [21] criteria for different effect sizes ranging from small (2% of the variance), medium (13% of the variance), to large (26% of the variance), respectively. The value 0.26 was also chosen as it was approximately halfway between the values from small and medium...
effects and was used in the previous study by MacKinnon et al. [6]. We decided to choose one combination of parameters a time rather than take the averages across various parameter combinations because of the following reasons. First, the aim of our paper was to investigate the properties of $ab$ and $t-t'$ after manipulation of hazard rates, mortality rates, and censoring rates. The true value of $ab$ can be calculated in one combination as well as in various combinations. Second, taking the average of all combinations has the risk of taking into account of many elements and variability, which could make the already complicated scenario even more complex. Three different sample sizes of 50, 200, and 1000 were chosen, which were found in social science studies as from small to large sizes [9]. For each simulation condition, 1000 replications were conducted. The following is the example of the SAS code used to generate the data.

data one;
   do i=1 to 1000;
      do sample=1 to 50;
         do a=0.14;
            do b=0.14,0.26,0.39,0.59;
               do c=0.14;
                  xbernoulli = RAND('BERNOULLI',0.5);
                  u=ranexp(0);
                  w=0-1*log(u);
                  x=rannor(0);
                  x = (round(x,.001));
                  err=rannor(0);
                  m=a*x+err;
               m = (round(m,.001));
                  event_hazard=exp(2.31+c*x+b*m+0.5*w);
                  event = enroll+event_hazard;
                  event = (round(event,.001));
                  if xbernoulli=1 then indicator=1;
                  else if xbernoulli=0 then indicator=0;
               output;
                  end;
               end;
            end;
        end;
   end;
run;
4 Simulation Results for the Property of \( ab \)

The simulated data based on Equation (4) were analyzed with a log-survival time model (see Equations (5)-(7)) using LIFEREG procedure in SAS. The sample values of the population parameters are represented by \( a, b, t \) and \( t' \).

4.1 Varying Hazard Rates

Figures 2 (a)-(d) show the property of point estimate \( ab \) in terms of the bias from the true value. Censoring is set at 70% and mortality at 30%. Figures show different patterns regarding the change of hazard rates when the parameter combination changes. Although there is no obvious pattern among the relations in the bias among the hazard rates, a point estimate with a constant hazard rate seems to always start with the largest bias relative to increasing and decreasing hazard rates when \( n=50 \). However, when the sample size is 200 or 1000, the bias decreases significantly. This is expected when an estimator is consistent which is the case for MLEs.

Figures 3 and 4 show the property of point estimate \( ab \) in terms of SE and MSE. There is a clear pattern in the relation among the hazard rates regardless of parameter combinations. Point estimates with decreasing hazard rate have the largest SE and MSE, followed by the constant hazard rate, and the estimates with the increasing hazard rates have the smallest values SE and MSE. As sample size increases, both SE and MSE decrease as well as the differences in the SE among the different hazard rates. Figures 3 and 4 show the relationship, since the rest of the figures with different parameter combinations follow a very similar pattern.
4.2 Varying Mortality Rates

Figures 5 (a)-(b) show the property of point estimate \( ab \) in terms of the bias from the true value. Here we set censoring at 70% and assumed an increasing hazard rate with \( \alpha=2 \). Similar to the relation in the hazard rates, there is no obvious pattern in the relationship between the bias and the change of mortality rate. Again we see that the bias decreases as sample size increases as expected.

Figures 6 and 7 show the property of point estimate \( ab \) in terms of \( SE \) and \( MSE \). Here we see that the mortality rate does not seem to make a difference in the \( SE \) and \( MSE \) of the estimate of \( ab \). Again, both \( SE \) and \( MSE \) decrease as sample size increases.

4.3 Varying Percent Censored

Figures 8 (a)-(d) show the property of point estimate \( ab \) in terms of bias. There is no obvious relationship between bias and censoring rates except that those values seem to decrease as sample size increases.

Figures 9 and 10 show the property of point estimate \( ab \) in terms of \( SE \) and \( MSE \). Both \( SE \) and \( MSE \) from different censoring rates are very close to each other, with fewer censoring cases showing a little bit smaller \( SE \) and \( MSE \). Again, both \( SE \) and \( MSE \) decrease as sample size increases. The patterns of \( SE \) and \( MSE \) are very similar for the other parameter combinations of \( a \) and \( b \), so the graphs are only present for one parameter combination.
4.4 Comparison of Standard Errors

We compared the standard errors of product ab between Goodman [17], Sobel [15] formulas and the empirical data in different scenario of hazard rates, mortality rates, and percents of censored data. The standard errors calculated from both formulaic methods are given as
\[
\sigma_{ab, \text{Goodman}} = \sqrt{a^2 \sigma_b^2 + b^2 \sigma_a^2 - \sigma_a^2 \sigma_b^2} \quad [17]
\]
and
\[
\sigma_{ab, \text{Sobel}} = \sqrt{a^2 \sigma_b^2 + b^2 \sigma_a^2} \quad [15],
\]
respectively, where \(\sigma_a\) and \(\sigma_b\) are the standard errors of \(a\) and \(b\). \(\sigma_a\) and \(\sigma_b\) are computed in our simulation as the standard errors of random variables \(a\) and \(b\) within a sample size. There is a very clear pattern in the relationship between the three methods. For instance, Figure 11 is an example of the comparison with small parameter effect size when 70% of censoring rate, 30% of mortality rate, and increasing hazard rate are all fixed. Notice standard errors from Goodman and Sobel are very close to each other, with Sobel’s slightly larger than Goodman. This is due to the fact that an extra positive term is subtracted from the Goodman standard error formula. Figure 11 also shows that both formulaic methods result in smaller standard errors than the empirical method. However, we see that the estimates are already very close to each other, even when \(n=200\), and are practically identical when \(n=1000\).

5 Comparisons between the Two Methods of Estimating Mediation Effects

5.1 Varying Hazard Rates

Figure 12 is a set of boxplots comparing the point estimates of ab and t-t’ to the true value of ab in the scenario of different hazard rates. The line and the “+” symbol inside the box represent the median and mean, respectively. Each box is bounded by its 1st and 3rd quartiles, and the vertical lines outside the box extend to the maximum and
minimum values. For \( n=50 \), the maximum values for the two methods are outside the range of the graphs. The true value of \( ab \) is the horizontal line in the graph. This graphs were chosen for the parameter combination of \((a=0.14, b=0.14, \text{ and } t'=0.14)\) only as they were representative of the pattern within the same hazard rate. For example, under the condition of an increasing hazard rate and sample size less than 1000, means of both point estimates are higher than the true value of \( ab \); however, \( t-t' \) shows a larger variability than \( ab \). The smaller sample sizes are associated with positively skewed distribution. However, as \( n \) increases, the means and medians for both methods get closer to the true value, the variability decreases and the distribution gets closer to being symmetric. This pattern holds for both the constant hazard rate and the decreasing hazard rate, except that the decreasing and constant hazard rates exhibit higher variability with decreasing hazard rate showing the most variability. Although in most of the scenarios regarding the sample sizes and hazard rates, the point estimate of \( t-t' \) is larger than \( ab \), the difference between two methods is getting smaller, especially at sample sizes as large as 1000.

5.2 Varying Mortality Rates

Figure 13 is a set of boxplots comparing the point estimates of \( ab \) and \( t-t' \) to the true value of \( ab \) when mortality rates are varying at 30\%, 50\%, and 80\%. Again, these graphs were picked for the parameter combination of \((a=0.14, b=0.14, \text{ and } t'=0.14)\) only as they were representative of the pattern within the same mortality rate. Figure 13 shows different characteristics of \( ab \) and \( t-t' \) at different scenarios of mortality rate. For example, estimates from \( t-t' \) show a relatively larger variability than the point estimates of \( ab \) with
mortality rate of 30% and 50%. However, the variability of the two methods is close at a mortality rate of 80%. Variability of $ab$ shows a better symmetric pattern than that of $t-t'$ within the same sample size when the mortality rate is smaller than 80%. As the sample size increases, the variability of estimates from both methods drops substantially; however, $t-t'$ still has a relatively higher variability than $ab$. In most of the scenarios regarding the sample sizes and mortality rates, the point estimate of $t-t'$ is larger than $ab$.

5.3 Varying Censoring Rates

Figure 14 is a set of boxplots comparing the point estimates of $ab$ and $t-t'$ to the true value of $ab$ when the percent of censoring varies in the model. These graphs from one parameter combination of $(a=0.14, b=0.14, \text{ and } t'=0.14)$ were representative of the pattern within the same censoring rate even with other parameter combinations. Figure 14 shows that the point estimate of $t-t'$ has a larger variability than that of $ab$ when censoring is present. Figure 14 (a) also shows that the difference regarding the symmetry and variability between the two methods are smaller when no censoring is present than the remaining scenarios of censoring. As sample size $n$ increases, the variability of the estimates drops substantially. In most of the scenarios for varying sample sizes and percent of censored, the point estimates of $t-t'$ tend to be larger than $ab$ but the difference between two methods is getting smaller, especially at a sample size as large as 1000. It is interesting to note that the means even for $n=50$ for both methods are very close to the true value.
6 Discussions

The purpose of this paper was to examine the point estimates of mediated effects in the context of survival analysis with censored data. With OLS, the mediated effect can be assessed with the difference in coefficients method, \( t-t' \), or the product of coefficients methods, \( ab \), and the two yield identical results. It has also been shown in the literature that under the condition of no censored data and no time-variant mediators, the differences in the two methods are negligible using the LIFEREG procedure, although there are consistent differences between \( ab \) and \( t-t' \) for the log-hazard model [9].

This study first examined the property of \( ab \) in terms of bias, SE and MSE. The changes of mortality rate, hazard rate, and amount of censoring were particularly manipulated to examine how these affect the behavior of \( ab \). Our findings did not show any clear pattern in terms of bias with various hazard rates, mortality rates, or amount of censoring, although bias always decreases as sample size increases. When it comes to SE and MSE, our findings found clear relationships across the different scenarios. Specifically, point estimates with decreasing hazard rate have the largest SE and MSE, followed by the constant hazard rate, and the estimates with the increasing hazard rates have the smallest values of SEs and MSEs. Fewer censoring cases are associated with somewhat smaller SE and MSE. Mortality rate does not have much effect on SE and MSE, as the values are all very close. For all those scenarios, both SE and MSE decrease substantially as sample size increases.

The comparison of standard errors between two formulaic methods and an empirical method shows a clear pattern regardless of the different specifications of the hazard rate, mortality rate and amount of censoring. That is, the SEs from Goodman and
Sobel are very close to each other with Sobel’s are always slightly larger than Goodman, which is due to the fact that an extra term is subtracted from Goodman standard error formula. Both formulaic methods generate smaller SEs than the empirical method, which is consistent with the findings from Tein and MacKinnon’s study [9].

Our study also examined the characteristics of the two mediated effect methods of \( ab \) and \( t-t' \) under different scenarios of hazard, mortality and the amount of censoring in the model. Our findings have found point estimates of \( t-t' \) have relatively larger variability than that of \( ab \) within the same specifications. In most of the scenarios, the point estimate of \( t-t' \) is larger than \( ab \), although the difference between two methods regarding both variability and estimates decreases as sample size increases. In the special case of no censoring data is present, our findings are consistent with the existing literature [9] that under the condition of no censored case, the equivalence of \( ab \) and \( t-t' \) applies to log-survival time model. Within each of the mediated methods of \( ab \) and \( t-t' \), the increasing hazard rate is associated with the smallest variability of point estimates, followed by the constant rate having moderate variability. The decreasing hazard rate has relatively the largest variability of point estimates. Different characteristics of \( ab \) and \( t-t' \) have been found when the mortality rate varies. For example, estimates from \( t-t' \) show a relatively larger variability than the point estimates of \( ab \) with mortality rates of 30% and 50%. However, the variability in the two methods is close at a mortality rate of 80%. Variability of \( ab \) shows a better symmetric pattern than that of \( t-t' \) within the same sample size when the mortality rate is smaller than 80%. The percentage of censoring data has an impact on the variability of the point estimates. More censored data is associated with higher variability on both methods.
Our study extended the existing method for examining mediation effects in survival analysis by including the censored data. We also examined how the two methods would act differently under different hazard, mortality and amount of censoring, respectively, in the model. It is proved theoretically and also in the simulation that under the circumstance of no censored data, $ab$ and $t-t'$ are theoretically equivalent to each other, in both OLS regression and log-survival time models. However, differences in the point estimates have been found between two methods with or without censoring. The comparison between two methods shows that the point estimate of $t-t'$ always has larger variability than that of $ab$. Both $ab$ and $t-t'$ behave differently when the hazard rate or mortality rates or the censoring amount varies individually or together. Some recommendations regarding the choice between two methods under different scenarios might be made. For example, at the condition of when there is an increasing or decreasing hazard rate, $t-t'$ does not perform as well as $ab$ unless the sample size $n$ reaches 1000. However, when there is a constant hazard rate present, two methods start to perform similarly to each other when the sample size is 200. So $ab$ is recommended for use when no constant hazard is present with moderate sample size. When a constant hazard rate is present, the choice of $ab$ and $t-t'$ can become flexible even with moderate sample size. When the mortality rate is less than or equal to 50%, $ab$ and $t-t'$ perform similarly regarding the estimates and variability even when the sample size is as large as 1000. However, when the mortality rate increases to 80%, both methods show a similar pattern of variability and symmetry of point estimates. As a result, $ab$ is recommended to be used if there is at most half of the mortality rate in the model. Again, no significant difference is detected between two methods when the mortality rate is over 50%. Both
methods show a larger variability of estimates when censoring is present. When there are 50\% or less censored data in the model, \( ab \) always shows a smaller variability than \( t-t' \), even when sample size increases to 1000. However, when more than half of the data are censored (percent censored=70\% in current study), both \( ab \) and \( t-t' \) show a very similar pattern regarding the estimates and variability when the sample size \( n \) is 1000. Therefore, when less than half of the data are censored, \( ab \) is recommended to be used to estimate mediation effects, especially with a moderate sample size. As over 50\% of data are censored in the model, no difference between two methods are detected should large sample size is allowed. In summary, \( ab \) performs better than \( t-t' \) in most of the scenarios of hazard and mortality rates and percent censored; however, the two methods become less distinguishable when the sample size increases to 1000.

There are limitations of the current study. One limitation is that it does not compare the two methods in the log-hazard procedure. The reason is that results from no censored cases in log-survival time model show a stable scenario from the previous study. Further research might want to examine the properties of the two methods in a log-hazard model where more unstable parameters need to be defined and the true values of point estimate need to be carefully derived. Another limitation is that this study conducts only 1000 replications in each of the simulations. Our findings have shown some obvious association between SE or MSE and hazard rate, mortality rate and amount of censoring; however, no such pattern is apparent with respect to the calculation of bias. We expect the results could show a better clear pattern regarding the differences of the two methods with more replications of simulations. Another limitation is that this study does not discuss time-varying mediating variables. We would expect to see the same or even more
issues with more complicated models. In addition, the model had a simpler scenario as
the true mediation was assumed to be known and there was no interaction term. It is often
the case that the model is unknown in the reality; therefore, it may be more difficult to
detect the true relations between X, M, and Y.
References


Figure 1. Models showing total effect and mediated effect of X on Y
Figure 2. Distribution of bias of $ab$ from true value at different hazard rates
Figure 3. Distribution of standard errors from point estimate of $ab$ at different hazard rates
Figure 4. Distribution of mean square errors from point estimate of $ab$ at different hazard rates.
Figure 5. Distribution of bias of $ab$ from true value at different mortality rates
Figure 6. Distribution of standard errors from point estimate of $ab$ at different mortality rates
Figure 7. Distribution of mean square errors from point estimate of $ab$ at different hazard rates
Figure 8. Distribution of bias of $ab$ from true value at different censoring percents
Figure 9. Distribution of standard errors from point estimate of $ab$ at different censoring percents
Figure 10. Distribution of mean square errors from point estimate of $ab$ at different censoring percents
Figure 11. Comparisons of standard errors between Sobel's formula, Goodman's formula and empirical data.
Figure 12. Boxplots of the distribution of $ab$ and $t-t'$ at different hazard rates ($n=1000$ replications, percent censored=70%, mortality rate=30%)
Figure 13. Boxplots of the distribution of $ab$ and $t$-$t'$ at different mortality rates ($n=1000$ replications, percent censored=70%, increasing hazard rate)
Figure 14. Boxplots of the distribution of $ab$ and $t-t'$ at different censoring percents (n=1000 replications, mortality rate=30%, increasing hazard rate)
1 Introduction

Unintentional weight loss (UWL) in older adults is a common but serious problem and is associated with increased mortality, significant adverse clinical outcomes and progressive morbidity [1]. It is well-known that both intentional and unintentional weight loss without exercise could result in both muscle and fat loss. It also reduces physical function especially for older adults as people lose muscle naturally when they age [2].

After examining the literature published from 1990 to 2003 addressing the prevalence, determinants, and consequences of undernutrition among 65 years or older adults, Payette and his colleagues [3] have summarized that a high prevalence of unintentional weight loss as well as very low energy intakes have been reported in this population. In clinical practice, it is reported that up to 27% of frail, community-dwelling older adults over 65 years have experienced unintentional weight loss of at least 5 kg over the last 2 years [4]. The presence of underlying health problems or poor nutritional status and reduced energy intake due to medical illness are frequently implied as the causes of unintentional weight loss among the elderly [1].

Previous studies have reported numerous factors associated with unintentional weight loss in older people. They include older age [5-6], disability [7], cognitive impairment [8], low education level [6], poorer health status [6], previous admission to
Numerous studies have shown the associations between unintentional weight loss and mortality in older adults. Locher and colleagues [11] have observed in a longitudinal study that in older adults the odds of experiencing mortality with unintentional weight loss were 1.67 times as those who didn’t report weight loss. In a 5-year study in seniors, it is reported that unintentional weight loss resulting in a BMI change of at least 2.0 units is significantly predictive of mortality [12]. A fair amount of unintentional weight loss within 5 to 10 years is associated with increased risk of mortality or overall health decline [13-17]. The risk of complications and likelihood of admission to an institution [4,18] associated with weight loss in older adults is even higher for those who have been in hospitalized [19].

Some risk factors predicting weight loss are also associated with mortality. However, we are unaware of any research that has evaluated the causal effect of risk factors on mortality through the mediation effects of unintentional weight loss for older adults in the context of survival analysis. Mediation happens when a third variable M is added to this X (risk factor) → Y (mortality) relation, whereby X causes changes in M, and M has a causal influence on Y, so X → M → Y. In other words, part of the causal relationship between X and Y can be explained by the mediator should mediation effect exist.

Since this study involves time-to-event binary dependent variable, Equations (1)-(3) were used to estimate the single-mediator in the context of proportional hazard model:

\[ M = i_3 + a_1x_1 + \ldots + a_kx_k \]  

(1)
\[ Y = \log h(t) = i_1 + t_1 x_1 + \ldots + t_k x_k \quad (2) \]

\[ Y = \log h(t) = i_2 + t_1 x_1 + \ldots + t_k x_k + bM \quad (3) \]

In these equations, \( i_1 \) and \( i_2 \) and \( i_3 \) are intercepts; \( Y \) is a time-to-event variable, with 1 indicating mortality in community-dwelling older adults; the \( X \) variables are the risk factor or predictor variables, such as gender, race, age, and nutritional variables; and the \( t \) weights are the log odds of mortality that one expects to see to increase for each unit increase in these risk factor variables. \( M \) is the change in weight observed over some period of time in kilograms, and \( b \) is the increase in log odds for mortality for each kilogram of weight change. Therefore, \( ab \) log odds of mortality are changed through weight loss resulting from the effects of a risk factor variable. Here effect \( t \) in Equation (2) is called the total effect; \( t' \) in Equation (3) is called the direct effect.

There are two ways to calculate the mediated effect, as either \( ab \) or \( t-t' \) [20]. In this paper, the mediation effect of weight loss was calculated and compared using two methods. The assessment of the mediation effect of unintentional weight loss can help researchers have a better understanding of the causes of mortality, as well as improve the chance to detect what is good for older adults’ health status and their remaining life.

Literature has shown that the 24-hour dietary recalls are widely used to obtain a participant’s mean daily caloric intake to assess the undernutrition of a subject. In a recent published paper examining the reliability and predictive validity of caloric intake measures from 24-hour dietary recalls [21], caloric intake deficiencies relative to estimated energy requirements was not found to predict weight loss in homebound older adults. Despite the popularity of using 24-hour dietary recalls, other screening instruments, for example, the serum albumin, have been used to identify undernutrition.
and its associated risk [22-23]. Albumin is a single protein that is synthesized in the liver during periods of adequate nutrition and not synthesized during inadequate nutrition. Although albumin synthesis does not decrease with aging, albumin levels decline by 0.8 g/L per decade in persons over 60 years of age [24]. Centenarians therefore have a significantly lower serum albumin level [25] than young persons. Limitations in using the serum albumin level as a marker for undernutrition are its long half-life of about 18 days [26], and therefore, no storage capacity for excess production of albumin in the body. As a result, the measurement errors usually occur in the presence of uremia [27] and stem from a lack of specificity [28]. Although these limitations impair the diagnostic values of albumin, some studies still use it as a prognostic tool, and it has been shown that a low serum albumin level is highly predictive of in-hospital mortality [29] and of mortality in the general population [30]. There is also evidence that it is associated with increased risk of in-hospital complications, readmissions, and prolonged duration of hospital stay [31]. Some literature also suggests that a low serum albumin concentration is associated with morbidity and increased mortality in older adults [22-23,32]. Therefore, in this paper, we still would like to use albumin as an indicator to assess the relationship between undernutrition and mortality. In particular, a cut point of less than 3.5g/dl, which has been widely used as a maker of undernutrition [33-39], is used in our current study.

The purpose of this study was to evaluate the multivariate determinants of observed weight loss over a 4-year period in a sample of community-dwelling older adults. The same risk factors and the observed weight loss were then evaluated as to whether they predicted mortality over the next 4.5 years of the study. These data and analyses allowed us to examine the relationship between key risk factors and mortality,
and to estimate the mediating effect of weight loss in accounting for these mortality risk factors. In this study, undernutrition was assessed by albumin level rather than the usual mean daily caloric intake. The ultimate goal was to answer the proposed question in our research, whether weight loss mediates the undernutrition-mortality relationship.

2 Methods

2.1 Sample and Design

The UAB Study of Aging (SOA) originally consisted of 1000 community-dwelling older adults living in the State of Alabama. It is a population-based, longitudinal observational study of mobility among older African Americans and Whites. Recruitment was based on a sample of aged 65 years or older residing in five central Alabama counties. Race, gender, and urban/rural residence were used to stratify the sampling. The sample consisted of 50% African American, 50% female, and 51% rural. The study protocol was reviewed and approved by the UAB Institutional Review Board. Participants were administered a questionnaire consisting of mobility, life space assessment, and overall health status [40] in their homes using a standard interview format.

One thousand participants underwent a baseline in-home assessment, followed by follow-up in-home evaluations every 6 months. The study is still ongoing and at the time of this paper, assessments of up to 102 months (8 years and a half) have been collected for available participants. The actual weight of the participants was measured at baseline and at a 48-month in-home assessment. As this paper focuses on the mediating effect of weight loss on the relationship between undernutrition (assessed by serum albumin) and
mortality, a sub-sample of 350 participants who had measured weight at both baseline and 48 months and albumin measurement at 48 months was drawn for the analyses. There are five major reasons for the remaining 650 subjects to be excluded from the current analyses. These reasons included the following: 1) 219 participants died before the 4-year assessment; 2) 48 participants had been lost to follow up for other reasons before the 4-year assessment; 3) 109 subjects continued to participate in telephone interviews but refused to participate in the 4-year home visit; 4) 85 subjects did both baseline and 4-year home visits, but didn’t have weight measured at one or both visits because many of them were in wheelchairs or bed bound and couldn’t be weighed; and 5) 189 subjects had measured weight at both home visits, but did not provide valid blood samples for albumin measurement.

2.2 Procedure and Measures

Weight Loss. Weight change was derived by subtracting a participant’s actual baseline weight from their 48-month weight. Weight change was analyzed as a continuous variable in kilogram.

Albumin. An overnight fasting blood sample was collected and analyzed for albumin measurement. Depending upon the analyses, serum albumin was analyzed as a continuous variable or was coded as a binary variable with a cut-off point of 3.5g/dl.

Life-space. The UAB Study of Aging Life-Space Assessment (LSA) was used to assess mobility. During the interview by the LSA, a person reported the moving distance
and frequency within one month preceding the assessment. The assessment of mobility can widely range from a person’s bedroom to independent travel outside the person’s town [41-42]. The LSA was scored by a set of 5-level questions and then summing the 5 level scores [40]. A score range of 0-120 was used for LSA, with higher scores indicating higher mobility. An intraclass correlation coefficient of 0.96 implied that LSA has a high test-retest reliability [41].

*Mortality.* Since our analysis sample only included participants who were weighed and provided albumin samples at 48 months, the mortality in the analyses consisted of participant deaths that occurred over a 4.5-year period between the 48-month and the 102-month assessments. A total of 63 death events are modeled in our proportional hazards models of mortality.

*Predictor Variables.* We controlled for baseline body mass index (BMI), count of comorbidity, age, gender, and race in our analyses. BMI was obtained by collecting height and weight and categorized according to NHLBI Clinical Guidelines [43]. A comorbidity index [44-45] was created, based on the diagnoses of medical conditions but no consideration of severity, giving one point for each disease category of the Charlson comorbidity index. A diagnosis needed to be verified and the analyses only contained the verified diagnoses [44], so this variable counted the number of comorbidities for each of the participants. Age, gender, and race were assessed by self-report. There were two categories for race (White and African American). Age was included as a continuous variable.
2.3 Statistical Methods

Data analyses were performed using SAS version 9.1 programming language. Descriptive statistics were completed first to compare the sub-sample of 350, the excluded sample of 650, and total sample of 1000 participants.

Generalized linear model (GLM) was examined to assess the multivariate predictors including albumin-assessed undereating on weight loss. The dichotomous albumin variable using the cutoff point of 3.5g/l was used in the GLM procedure.

Three mediation effect models including GLM and 2 Cox proportional hazard models, as shown in Equations (1), (2) and (3), were examined to determine if the effects of the multivariate determinants, especially serum albumin on mortality, could be explained by the mediation effect of weight loss.

Mediation effect was calculated in two ways: The product of coefficients method ab and the difference of coefficients method t-t'.

Under the assumption that the dependent variable is continuous and the residual variance follows normal distribution, the equivalence of \( ab \) and \( t-t' \) was shown by Mackinnon et al. using ordinary least squares regression [46]. However, for mediation analyses with binary dependent variables, the equivalence of \( ab \) and \( t-t' \) does not always hold [20]. In a more recent paper, MacKinnon and colleagues [47] conducted a simulation study in the model with binary outcomes. Their findings indicated that the asymptotically unbiased estimates can be directly produced by the \( ab \) method while the \( t-t' \) method must be scaled properly before it is used to estimate mediated effect.

In order to have a more accurate test of the mediation effect, MacKinnon et al. [48] in their simulation study incorporated a set of different values of \( a \) and \( b \) to estimate
the empirical distribution of $ab$ and determined critical values rather than the values from the standard normal test. MacKinnon et al. derived a program called PRODCLIN from a previously existing Fortran program called FNPROD and have improved the product of coefficients method on it by computing the critical values [49-50]. In this paper, mediation effect calculated using product of coefficients method was examined using the PRODCLIN method [48,50]. A total of 63 (18% of 350) death events over a 4.5-year period are modeled in our proportional hazard model.

3 Results

3.1 Descriptive Statistics

Table 1 presents baseline characteristics for two subsamples and whole sample of participants. The two subsamples were similar in the proportion of gender, BMI category, BMI values, and baseline weight, but there were significant differences between the groups in age, race, life space score, number of comorbidity, serum albumin, and proportion of death. In particular of our interest, the group of 350 had a significantly younger age, more white subjects, higher life space score, fewer number of comorbidity, and lower percentage of death than the remaining 650 subjects. The group of 350 also had a significantly higher serum albumin level than the remaining sample; however, it should be pointed out that there were only 51 subjects with albumin level in the group of 650. Although some researchers might question the reliability of the comparison based on the very different sample sizes, this result actually indicates that subjects with measured weight at both visits had relatively more stable health status and, therefore, better nutrient biochemical parameters.
3.2 Mediation Analysis Models

Tables 2 through 4 show the details of mediation effect models as displayed in Equations (1)-(3), respectively. The analyses for all three tables were conducted only on the subgroup of 350 participants. The relationship between comorbidity count and weight loss/mortality was examined (data is not shown), and including it in the models would impair the effect of other risk factors. To main the efficiency of the models, comorbidity count was only displayed in the descriptive statistics basis; it was not included into any of the mediation models.

Table 2 presents the multivariate effect of predictors on weight loss (mediator) for the subgroup of 350 participants. It shows that factors such as being African American and higher BMI category significantly predicted observed weight loss. Table also shows older age and lower life space score were associated with weight loss at borderer significance after controlling for gender and albumin. The effects in this model are equivalent to the \(a\) path in Equation (1).

Table 3 presents the multivariate effect of the same set of risk predictors on mortality (dependent variable). Among the 350 subjects in the analysis, 88.6% of them had reported albumin\(\geq 3.5\text{g/dl}\). The result shows that the odds of experiencing mortality for participants who had albumin less than 3.5g/dl were 2.3 times as those who had albumin\(\geq 3.5\text{g/dl}\) after controlling the other variables. Those effects are called total effects (\(t\) path in Equation (2)) in the model for each of the predictors respectively.

Table 4 presents the multivariate effect of the same risk predictors and the observed weight loss (mediator) on mortality. It shows weight loss, albumin less than 3.5g/dl and male subjects were significantly associated with mortality. Here the effect of
weight loss is equivalent to the $b$ path in Equation (3), and the effects of the other risk factors in the model are called direct effects ($t'$ path in Equation (2)).

Table 5 presents the mediation effects of weight loss on each of the risk factors calculated using both $ab$ and $t-t'$ methods. Table 5 also shows the 95% asymmetric confidence limits for $ab$ using the PRODCLIN method [47-48]. The data were generated based on Equations (1) and (3) for each of the predictors and observed weight loss, respectively, where $X$ (predictors) and $M$ (weight loss) followed normal distributions and were generated using the RANNOR function with zero as the seed. The values of all the parameters specified in the model came from the estimates through Table 2 and Table 4. The parameter estimates ($a$, $b$) and standard errors calculated from both GLM and Cox proportional model of each of the predictors were put in the PRODCLIN program. A 95% asymmetric confidence limits were calculated for the mediation effect $ab$. Table 5 shows that 95% confidence limits do not include zero for both race and BMI category comparison between underweight and obesity group. The result indicates that weight loss had significant mediation effects on race and one of the BMI category comparisons using the product of coefficients. In Table 5, a column shows the proportion of the mediation effect over the total effect calculated by two methods, respectively. Notice that, for both race and BMI, the absolutely value of proportion is greater than 1 for both methods.

This is the situation in which the mediated effect and direct effect have different signs in the model [51-53]; i.e. when the inconsistent mediation occurs. When this happens, knowing the significance of the relation between predictors $X$ and dependent variable $Y$ is not enough to interpret the results, as there are cases where an overall $X$ to $Y$ relation is not significant, but the mediation exists [47]. We have a very similar
scenario in the study. Let’s look at the mediation effect calculated by $ab$ and $t-t’$. For most of the variables in Table 5, the effects from both methods are not comparable, and effects calculated from $t-t’$ are consistently larger than $ab$. This is consistent with the findings we have shown in Chapter 3 of this dissertation. For survival analysis without censored data, the equivalence of $ab$ and $t-t’$ does not hold anymore. After introducing censored data into the model, $t-t’$ behaves even worse than $ab$, especially with small sample sizes. This is why we prefer to use $ab$ rather than $t-t’$ with censored data in the model. Although the difference between $ab$ and $t-t’$ decreases with fewer censored data and large sample size, the difference cannot be ignored. We must also discuss the inconsistent mediation in our model. For example, BMI has a nonsignificant direct effect of 0.467 on mortality (From Table 4), and has a mediation effect of $ab=-0.25$, which has the opposite sign from direct effect. The model in Table 2 indicates that obesity subjects had a higher likelihood of losing weight, but the models in Tables 3 and 4 show that obesity subjects were more likely to survive. The overall relation between BMI and mortality may seem to be very small due to the existence of two opposite processes, which explains why the absolute proportion of -1.15 is greater than 1. The sign of the proportion also indicates that the inconsistent mediation has happened. In addition, the 95% asymmetric confidence limit shows the significance of this mediation effect. Table 2 shows that obesity BMI is significantly associated with weight loss; therefore, here we can interpret this as meaning that almost all the BMI effect on mortality was explained by the inconsistent mediation of weight loss. It is the same with the Race variable; however, there is a large difference in the proportion of the mediation effect between two methods. Notice both effects on $b$ path and $t’$ path are very small, which are close to zero. Albumin
was found to have significant total and direct effect on mortality, but was not found to be significantly associated with weight loss; therefore, the 95% asymmetric confidence limit did not show its effect on mortality mediated through observed weight loss. Notice that the proportion of mediation effect is only 0.06, which would have been interpreted as 6% of the albumin effect on mortality being explained by weight loss. The relationships among risk factors, weight loss, and mortality are shown in Figure 1.

4 Discussion

Our findings indicate that race and BMI are risk factors of weight loss and, as a result, weight loss significantly predicts mortality. These relationships were true for participants who had measured weight at both baseline and 48 months and albumin measurement at 48 months. The BMI category in our analysis was baseline BMI where higher BMI was associated with weight loss. The relationship between 48-months BMI and weight loss was also examined (data was not shown) and showed that lower BMI was associated with weight loss. It could be explained that after 48 months, the higher BMI participants had lost weight and had reduced their BMI; therefore, at 48 months assessment, current lower BMI actually predicted weight loss. These findings are consistent with previous reports in the literature regarding the risk factors of weight loss and mortality [5-7,11]. In this paper undernutrition was assessed by albumin level, an alternative to 24-hour dietary recall criteria. Our multivariate model does not show that albumin is the risk factor of weight loss. As mentioned earlier, a recent published paper by the same author examined the predictive validity of caloric intake in the homebound older adults [21], and caloric intake assessed undernutrition was not found to predict
weight loss. The two studies examined the same problem by using different criteria to assess undernutrition; neither of the results detected a significant association between undernutrition and weight loss. It should be pointed out that both studies have found that undernutrition significantly predicts mortality. Literature has shown that under-reporting of caloric intake is an issue of 24-hour dietary recall that might result in its non-significant relationship between weight loss. Our analysis with albumin avoided the possibility of under-reporting; however, it did not produce any more promising results. That might be due to the limitations mentioned earlier; i.e., that albumin level declines for people over 60 years of age and might not be a very good indicator for undernutrition especially for older adults [24]. Our models have successfully detected the undernutrition-mortality relationship and other risk factors which are consistent with the findings in the literature; therefore, it is reasonable to say that our models with current number of predictors and given sample size are reliable and stable to detect the significance should it exists. Although the limitations of albumin as an indicator of undernutrition might be the reason that we failed to detect the significance, undernutrition itself may not be a risk factor as strong as the other factors of observed weight loss. Future research might want to take efficient action to improve the collection of nutrition information.

When it comes to the calculation of mediation effects, both product of coefficients and the difference of coefficients methods were introduced. For most of the predictors in the model, the effects from both methods are not comparable, and effects calculated from $t-t'$ are consistently larger than $ab$. Those findings are consistent with the simulation results shown in Chapter 3, where after introducing censored data into the model, $t-t'$
behaves even worse than $ab$ especially with study duration as short as less than 5 years, which is similar to our study period of 4.5 years and small sample sizes. Five years is a reasonable length for most of the studies that last. We would like to suggest using the $ab$ method rather than $t-t'$ with censored data in the model for this scenario.

The 95% asymmetric confidence limits calculated from the PRODCLIN method [47-48] show that the causal relationship between some risk factors and mortality are mediated by weight loss, especially when the factors are also significantly associated with weight loss itself.

Our study has shown an extended application of mediation analysis by describing a situation where an inconsistent mediation effect exists in terms of the BMI variable from the analytic model. Taking into account inconsistent mediation effects can be conducted to better describe more complicated causal relations between variables, especially when more than one mediator exist and the mediated effects have different signs. It can also explain why a relation which is supposed to be present does not show a strong magnitude.

This study has contributed novel significant input to the research of causal effect of risk factors on mortality through the mediation effect of weight loss in the context of survival analysis, and the results are encouraging. There are some limitations in our study. In order to maintain the efficiency and feasibility of the model under the restrictions of sample size and number of events, we only looked at a limited number of risk factors of weight loss and mortality. Albumin is the variable we are especially interested in to assess undernutrition regardless of its potential limitation as an undernutrition indication for older adults, and a dichotomous measurement with cut-off
point of 3.5g/dl has been found to have significant association with mortality. The analysis has shown that the significance of path $a$ effect is very important for us to detect a significant mediation effect. Therefore, without a significant association between albumin and weight loss, we were not able to detect the mediation effects of weight loss on albumin-mortality relationship. Future studies might want to investigate the mediation effect of weight loss on the undernutrition-mortality relationship using different criteria to assess undernutrition, and also investigate further the mediation effect in the context of survival analysis with time-dependent variables.
References


Table 1. Descriptive statistics for two subgroups and whole sample of participants

<table>
<thead>
<tr>
<th>Variables</th>
<th>Subgroup (Mean (SD) or N (%))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=350  a</td>
</tr>
<tr>
<td>Age **</td>
<td>73.6 (5.8)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>172 (49.1)</td>
</tr>
<tr>
<td>Male</td>
<td>178 (50.9)</td>
</tr>
<tr>
<td>Ethnicity **</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>142 (40.6)</td>
</tr>
<tr>
<td>White</td>
<td>208 (59.4)</td>
</tr>
<tr>
<td>Weight at Baseline (kg)</td>
<td>79.2 (16.7)</td>
</tr>
<tr>
<td>Weight Change from Baseline to 48 Months</td>
<td></td>
</tr>
<tr>
<td>Mean (Sd)</td>
<td>-1.91 (6.3)</td>
</tr>
<tr>
<td>Range</td>
<td>(-33.60,23.61)</td>
</tr>
<tr>
<td>Serum Albumin **</td>
<td>3.8 (0.32)</td>
</tr>
<tr>
<td>Body Mass Index at Baseline</td>
<td></td>
</tr>
<tr>
<td>Mean (Sd)</td>
<td>27.7 (5.5)</td>
</tr>
<tr>
<td>Range</td>
<td>(16.8, 49.0)</td>
</tr>
<tr>
<td>Underweight (BMI&lt;18.5)</td>
<td>23 (6.6)</td>
</tr>
<tr>
<td>Normal weight (18.5 ≤ BMI&lt;25)</td>
<td>84 (24.0)</td>
</tr>
<tr>
<td>Overweight (25 ≤ BMI&lt;30)</td>
<td>138 (39.4)</td>
</tr>
<tr>
<td>Obese (BMI ≥ 30)</td>
<td>105 (40.0)</td>
</tr>
<tr>
<td>Life Space at Baseline **</td>
<td>74.3 (21.1)</td>
</tr>
<tr>
<td>Comorbidity Count at Baseline **</td>
<td>1.9 (1.4)</td>
</tr>
<tr>
<td>Death at Any Time **</td>
<td>63 (16.0%)</td>
</tr>
</tbody>
</table>

\[a\] Participants who had actual measured weight at both baseline and 48-month and albumin measurement  
\[b\] Participants who were excluded from the analyses  
\[c\] Whole sample of participants  
\[d\] Sample size which is different from subgroup N is displayed  
** \(p<0.01\)
Table 2. Generalized linear model for multivariate determinants of weight loss \(^a\) for the subgroup of 350 participants

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin (≥ 3.5 g/dl vs &lt;3.5g/dl)</td>
<td>1.160</td>
<td>0.971</td>
<td>0.233</td>
</tr>
<tr>
<td>Gender (Female vs. Male)</td>
<td>-0.534</td>
<td>0.706</td>
<td>0.450</td>
</tr>
<tr>
<td>Race (African American vs. White)</td>
<td>-1.820</td>
<td>0.700</td>
<td>0.0098**</td>
</tr>
<tr>
<td>Age</td>
<td>-0.111</td>
<td>0.060</td>
<td>0.065</td>
</tr>
<tr>
<td>Baseline Life Space</td>
<td>0.031</td>
<td>0.018</td>
<td>0.085</td>
</tr>
<tr>
<td>BMI Category</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;18.5 vs. ≥ 30</td>
<td>5.511</td>
<td>1.407</td>
<td>0.0001**</td>
</tr>
<tr>
<td>18.5-25 vs. ≥ 30</td>
<td>0.984</td>
<td>0.935</td>
<td>0.294</td>
</tr>
<tr>
<td>25-30 vs. ≥ 30</td>
<td>0.941</td>
<td>0.805</td>
<td>0.243</td>
</tr>
</tbody>
</table>

\(^a\) Weight loss is calculated by subtracting weight at baseline from weight at 48 month.

\(^*\) p<0.05

\(^**\) p<0.01
Table 3. Cox proportional hazard model for multivariate effects of predictors on mortality \(^a\) for the subgroup of 350 participants

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate</th>
<th>Hazard-ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin ((\geq 3.5) g/dl vs &lt;3.5g/dl)</td>
<td>-0.852</td>
<td>0.426</td>
<td>0.007**</td>
</tr>
<tr>
<td>Gender (Female vs. Male)</td>
<td>-0.548</td>
<td>0.578</td>
<td>0.055</td>
</tr>
<tr>
<td>Race (African American vs. White)</td>
<td>-0.004</td>
<td>0.996</td>
<td>0.989</td>
</tr>
<tr>
<td>Age</td>
<td>0.029</td>
<td>1.029</td>
<td>0.194</td>
</tr>
<tr>
<td>Baseline Life Space</td>
<td>-0.01</td>
<td>0.990</td>
<td>0.159</td>
</tr>
<tr>
<td>BMI Category &lt;18.5 vs. (\geq 30)</td>
<td>0.186</td>
<td>1.204</td>
<td>0.721</td>
</tr>
<tr>
<td>BMI Category 18.5-25 vs. (\geq 30)</td>
<td>-0.168</td>
<td>0.845</td>
<td>0.656</td>
</tr>
<tr>
<td>BMI Category 25-30 vs. (\geq 30)</td>
<td>-0.256</td>
<td>0.774</td>
<td>0.418</td>
</tr>
</tbody>
</table>

\(^a\)Mortality includes death from 48 months to 102 months

\(^\text{**}\)p<0.01
Table 4. Cox proportional hazard model for multivariate effects of predictors including weight change on mortality\textsuperscript{a} for the subgroup of 350 participants

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate</th>
<th>Hazard-ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight Change</td>
<td>-0.045</td>
<td>0.956</td>
<td>0.037*</td>
</tr>
<tr>
<td>Albumin ((\geq 3.5) g/dl vs &lt;3.5g/dl)</td>
<td>-0.824</td>
<td>0.439</td>
<td>0.0095**</td>
</tr>
<tr>
<td>Gender (Female vs. Male)</td>
<td>-0.597</td>
<td>0.550</td>
<td>0.0396*</td>
</tr>
<tr>
<td>Race (African American vs. White)</td>
<td>-0.093</td>
<td>0.911</td>
<td>0.745</td>
</tr>
<tr>
<td>Age</td>
<td>0.026</td>
<td>1.026</td>
<td>0.247</td>
</tr>
<tr>
<td>Baseline Life Space</td>
<td>-0.008</td>
<td>0.992</td>
<td>0.256</td>
</tr>
<tr>
<td>BMI Category</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;18.5 vs.(\geq 30)</td>
<td>0.467</td>
<td>1.595</td>
<td>0.386</td>
</tr>
<tr>
<td>18.5-25 vs. (\geq 30)</td>
<td>-0.106</td>
<td>0.899</td>
<td>0.779</td>
</tr>
<tr>
<td>25-30 vs. (\geq 30)</td>
<td>-0.201</td>
<td>0.818</td>
<td>0.528</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Mortality includes death from 48 months to 102 months

*\(p<0.05\)

**\(p<0.01\)
Table 5. Mediation effects of weight loss on risk factors-mortality relationship

<table>
<thead>
<tr>
<th>Mediation Effect</th>
<th>t-t'</th>
<th>ab</th>
<th>(t-t')/t</th>
<th>ab/ (ab+t')</th>
<th>95% Asymmetric confidence limits for ab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin (≥3.5 g/dl vs &lt;3.5 g/dl)</td>
<td>-0.028</td>
<td>-0.053</td>
<td>0.033</td>
<td>0.060</td>
<td>(-0.171, 0.028)</td>
</tr>
<tr>
<td>Gender (Female vs. Male)</td>
<td>0.050</td>
<td>0.024</td>
<td>-0.091</td>
<td>-0.042</td>
<td>(-0.036, 0.102)</td>
</tr>
<tr>
<td>Race (African American vs. white)</td>
<td>0.089</td>
<td>0.083</td>
<td>-23.041</td>
<td>-7.671</td>
<td>(0.004, 0.199)</td>
</tr>
<tr>
<td>Age</td>
<td>0.003</td>
<td>0.005</td>
<td>0.107</td>
<td>0.162</td>
<td>(-0.0003, 0.014)</td>
</tr>
<tr>
<td>Baseline Life Space</td>
<td>-0.002</td>
<td>-0.001</td>
<td>0.190</td>
<td>0.150</td>
<td>(-0.004, 0.0001)</td>
</tr>
<tr>
<td>BMI Category</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;18.5 vs. ≥30</td>
<td>-0.282</td>
<td>-0.250</td>
<td>-1.517</td>
<td>-1.150</td>
<td>(-0.547, -0.017)</td>
</tr>
<tr>
<td>18.5-25 vs. ≥30</td>
<td>-0.062</td>
<td>-0.045</td>
<td>0.369</td>
<td>0.296</td>
<td>(-0.155, 0.033)</td>
</tr>
<tr>
<td>25-30 vs. ≥30</td>
<td>-0.055</td>
<td>-0.043</td>
<td>0.215</td>
<td>0.175</td>
<td>(-0.140, 0.024)</td>
</tr>
</tbody>
</table>

*p<0.05*
Figure 1. Path models showing relationship among risk factors, unintentional weight loss and mortality

* $p<0.05$

** $p<0.01$

Solid lines indicates the multivariate effect on weight loss, which can be referred to be $a$ effect in Equation (1)

Dash lines between unintentional weight loss to mortality can be referred to be $b$ effect in Equation (3)

Dash lines between predictors to mortality can be referred to be $t'$ effect (also called direct effect) in Equation (3)
SUMMARY AND FUTURE RESEARCH

1 Summary

This research focused on the examination of mediation effects in the context of survival analysis when there were censored data. There are two ways to calculate the mediated effect in the single-mediator model, as either product of coefficients methods \( ab \) or the difference in coefficients method \( t-t' \). The equivalence of both methods was shown under ordinary least squares regression [20]. It also has been shown that under the condition of no censored data and no time-variant mediators, the differences of two methods are negligible using the LIFEREG procedure, although there are consistent differences between \( ab \) and \( t-t' \) for the log-hazard model [21].

It is appealing to examine whether the conclusion of Tein and Mackinnon’s study applies to data that have censored cases and if there are new features can be explored. As part of this research, we have improved the characteristic of the mediation model by including the censored data, and have found more appealing features and comparisons between the methods which can lead a better real-world decision. Furthermore, as a practical application, with a longitudinal observation study to assess the relationship between under-eating, unintentional weight loss, and mortality in homebound or community-dwelling older adults, the effect of weight loss was examined to evaluate whether weight loss mediated the under-eating-mortality relationship. Below we describe in detail each of the aspects:
1. Methodological Extension: We extended Tein and MacKinnon’s study in a more complicated scenario where the censored data were included. Censored survival data are very common and censoring requires special treatment. In particular, right censoring is far more common than left censoring in the practical research because observation is terminated before the target event occurs [22]. We examined the property of \( ab \) in terms of bias, SE, and MSE under different scenarios of hazard rates, mortality rates, or percentage of censoring. Our findings did not show any clear pattern in bias under different specifications, although both SE and MSE have been found to perform differently. For example, point estimates with decreasing hazard rate have the largest SE and MSE, followed by the constant hazard rate, and the estimates with the increasing hazard rate have the smallest values of SE and MSE. We further compared the point estimates from \( ab \) and \( t-t' \) to the true value of \( ab \) and examined how the characteristics of two methods change over sample sizes after manipulating the hazard rate, mortality rate, and percentage of censoring. Some features have been found to hold regardless the censored restriction. However, there are some significant differences in the results with introduction of censored data. Unlike the linear model situation, the difference in coefficients and product of coefficients estimators can lead to substantially different estimates and inferential conclusions. Generally speaking, under most scenarios of specifications, the point estimate of \( t-t' \) has larger variability than that of \( ab \), although the difference between two methods regarding both variability and estimates decreases as sample size increases. Both \( ab \) and \( t-t' \) behave differently when hazard rate, mortality rate, or censoring amount vary individually or together. Some recommendations regarding the choice between two methods under different scenarios might be made. For
example, in the case of no constant hazard rate, $ab$ is recommended for use with moderate sample sizes. When a constant hazard rate is present, the choice of $ab$ and $t-t'$ can become flexible even with moderate sample sizes. When there is less than half of the mortality rate in the model, $ab$ is recommended, although the difference between the two methods becomes negligible when the mortality rate is over 50%. Although both methods show a larger variability of estimates when censoring is present, they behave differently under specific cases. For instance, $ab$ performs better and thus is recommended to be used when less than half of the data are censored and only moderate sample size is available. As the percentage of censoring is over 50% and a large sample size is allowed, the difference between two methods becomes negligible. The choice between $ab$ and $t-t'$ depends on a lot other elements involved in the research. Two methods can always be looked at in pair to examine the difference. When there are many censoring data in the model, neither method gives a perfect point estimate of the mediation effect. Our findings have shown that $ab$ performs better than $t-t'$ in most of scenarios of hazard rate, mortality rate, and percent censored, especially with moderate sample size, although two methods become less distinguishable when sample size increases to be 1000. Therefore, in order to increase the accuracy and specificity which are essential to most designs, we therefore recommend using the product of coefficient methods in a LIFEREG procedure when only a moderate sample size is available.

2. Practical application on longitudinal study: We applied this improved method to a population of older adults to investigate the relationship between under-eating and mortality through the mediating effect of unintentional weight loss. First of all, the reliability and validity of self-report caloric intake deficiency measures from 24-hour
dietary recalls were examined for a homebound older adult population. Specifically, we examined the predictive validity of 24-hour dietary recalls by evaluating the effect of caloric intake deficiencies at baseline on observed weight loss over the subsequent 6 months. This validation aspect of the study also consists of the $a$ path of our undereating $\rightarrow$ UWL $\rightarrow$ mortality mediation model. In the study, we hypothesized that a relationship would exist between caloric intake discrepancy (caloric intake minus EER) and observed weight loss in homebound older adults. Our findings, however, revealed a low correlation between caloric intake discrepancy and actual weight loss, even in a subsample we believe to be more reliable with respect to health status. Further, there was no association between caloric intake discrepancy and weight loss in either the logistic regression or GLM analyses for this subsample.

We also applied this improved method to a population of community-dwelling older adults to further examine the mediation effect of weight loss on the undereating-mortality relationship. Paths $a$, $b$, and $t'$ were all examined in the study. In this study, $a$ path of our undereating $\rightarrow$ UWL $\rightarrow$ mortality mediation model was examined again by using an alternative method to assess undernutrition. Albumin level, which was used as the alternative criterion to 24-hour dietary recall to evaluate undereating, was not found to have significant association with unintentional weight loss. Both studies have detected the significance of UWL $\rightarrow$ mortality, the $b$ path of mediation model. Both $ab$ and $t-t'$ methods were used to calculate the mediation effect. For most of the predictors in the model, the effects from both methods are not comparable, and the effects calculated from $t-t'$ are consistently larger than $ab$. Those findings are consistent with the simulation results shown in the simulation study, where after introducing censored data into the
model, \( t-t' \) is farther away from the true value of mediation effect than \( ab \) with a study duration of 4 years regardless of the sample sizes, hazard rate and mortality rate. Four years is a reasonable average length since many studies usually have a 3-year trial with 2 years follow up. We would like to suggest using the \( ab \) method rather than \( t-t' \) with censored data in the model for this scenario. The 95% asymmetric confidence limits show that the causal relationship between some risk factors and mortality are mediated by weight loss, especially when the factors are also significantly associated with weight loss itself.

3. Novel application: Our study has contributed a novel input to the research regarding the causal effect of risk factors on mortality through the mediation effect of weight loss in the context of survival analysis. Our simulation was conducted to improve the existing method by examining the mediation effect based on different combinations of sample sizes, amount of censored data, varieties of hazard rate, mortality rate and effect sizes. Then we have applied the extended method to the specific studies to examine each path of the mediation model we have proposed. Our findings have shown that unintentional weight loss mediated some risk factors and mortality relationship, especially when the risk factors were associated with weight loss itself (significant \( a \) path). However, our findings did not show the mediation effect of weight loss on the undereating-mortality relationship. When it comes to the calculation of the mediation effect with censoring data present, our research has shown that neither \( ab \) nor \( t-t' \) performs perfectly. However, after carefully comparing the two methods under different scenarios of hazard and mortality rates and the amount of censoring, \( ab \) is found to be a better point estimate than \( t-t' \) in most of time.
2 Future Research

We have examined different perspectives on research regarding older adults, and here are several directions which can be pursued as part of our future research.

1. Under-reporting of 24-hour dietary recalls might be a factor that impair that fails to allow us to detect the true relationship between under-eating and unintentional weight loss. Future research might want to take efficient action to improve the collection of nutrition information.

2. Future studies might want to investigate the mediation effect of weight loss on the undernutrition-mortality relationship using different criteria to assess undernutrition.

3. A limitation of the current study is that it does not compare the two methods in the log-hazard procedure. The reason is that results from non-censored cases in a log-survival time model show a stable scenario from the previous study. The extension of the research by introducing the censored data is necessary in order to add a further complexity to the model. However, one must be careful not to involve too many unstable parameters that could compromise the results. Further research might want to examine the properties of two methods in the log-hazard model where more unstable parameters need to be defined and the true values of point estimate need to be carefully derived. Another limitation is that this study conducts only 1000 replications in each of the simulations. A better and clear pattern with respect to bias, standard error, and mean square error, as well as the comparison between the two methods, is expected to include more replications of simulations. Another limitation is that it does not discuss time-varying mediating variables. We would expect to see the same or even more issues with more complicated models. In addition, the model had a simpler scenario, as the true
mediation was assumed to be known and there was no interaction term. Future research might want to investigate more features of $ab$ and $t-t'$ with respect to point estimates, standard errors, and other factors with time-dependent variable and interaction terms.
LIST OF GENERAL REFERENCES


45. Institute of Medicine, Committee on Nutritional Services for Medicare Beneficiaries. The Role of Nutrition in Maintaining Health in the Nation’s Elderly: Interventions and Assessments Can Help Beneficiaries. Washington, DC 2001: National Academy Press.


APPENDIX B

WEIGHT LOSS AND MORTALITY IN THE COMMUNITY-DWELLING OLDER ADULTS (MORBILITY AMONG OLDER AFRICAN AMERICANS AND WHITES)