Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Estimating Mediation and direct effects of the multiple model with application Kidney

Zain alabdin Aboud Kazem El-Husseini[a] , *Ahmad N. Flaih*[b]

[a] Department of Statistics, University of Al-Qadisiyah, Iraq. Email: zain7728764@gmail.com
[b] *Department of Statistics, University of Al-Qadisiyah, Iraq .Email: :* ahmad.flaih@qu.edu.iq

ARTICLE INFO

Article history:
Received: 16 /07/2019
Rrevised form: 29 /07/2019
Accepted : 18 /08/2019
Available online: 21 /09/2019

Keywords:

Multiple mediator model; Effects Mediation; direct effect ; causal inference .

DOI :10.29304/jqcm.2019.11.4.617

ABSTRACT

This paper presented a study of the analysis of the mediation of a multi-model and the knowledge of the role of these variables in the transfer of the effect of the independent variable to the dependent variable. In this article, the delta method was used to estimate the standard error and because this method violates a number of assumptions, especially with small sampling sizes, the method of shoeboxes was used to overcome such problems. According to the results obtained from the application for this model, it was found that the shoe bass method is better than the Delta method according to numerical results.

MSC.

Research Goal

This search is used to study the simple and multilevel mediation model and estimation of direct and indirect impacts using Ordinary Least Squared (OLS) estimation method & boot- strapping Via a set of medical data specific to each model .

1 . Introduction

Most research in medical, social and economic sciences focuses on the relationship between two variables where the first variable is the independent **X** (causal variable) and the second variable is the **Y** (variable response) , When adding a third variable (**M**) it may be more difficult as the new variable is an intermediary between the

Corresponding author Zain. Al Hussain

Email: zain7728764@gmail.com

Communicated by Qusuay Hatim Egaar

independent variable **X** and the variable **Y**, which is called the mediation variable (**M**). The simplest form of mediation is the addition of the variable M which explains the effect of the variable **X** on the variable of the result **Y**, in which case the independent variable x is the reason for the occurrence of the mediation variable (**M**) which in turn causes the result variable **Y** ie that **X→M→Y** (D. P. MacKinnon, Fairchild, & Fritz, 2007) (Grotta & Bellocco, 2012).

In this article we will study the multiple mediation model with its applications. The simplest forms of multiple mediation consists of two intermediaries working to transfer the effect from the independent variable to the dependent variable and the following figure illustrates the relationship between these variables.

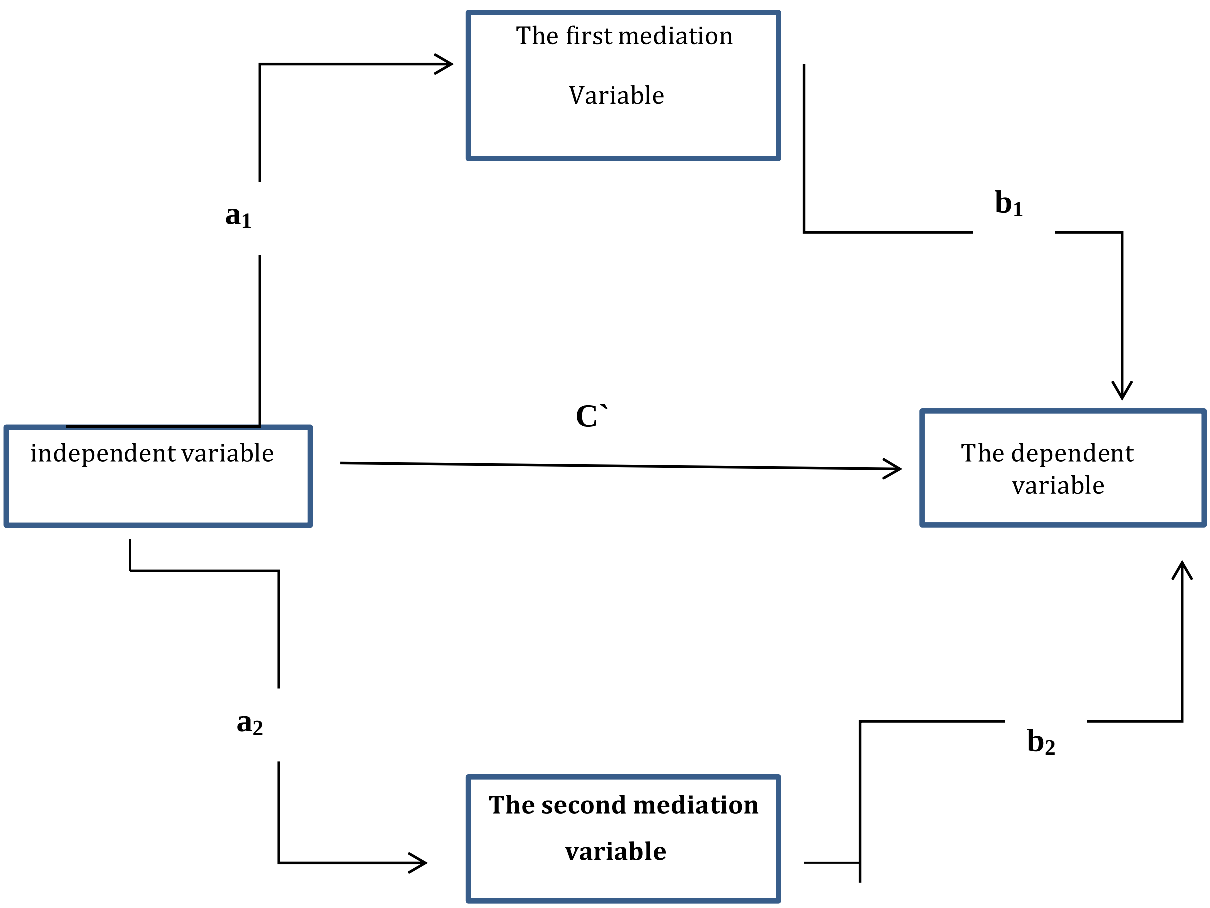


Figure 1: Schematic of a multiple mediation model using two variables.

Most studies are interested in determining or knowing the mechanism in which the independent variable affects the dependent variable, mediation analysis is the most widely used method where it involves the installation of a series of linear regression models in the framework of structural equation modeling (SEM) (Baron & Kenny, 1986). Mediation analysis was used to investigate and study the relationships between a wide range of variables in nutrition research in the American Journal of clinical Practice (Blood Fat Ratio) BMI was selected (BMI) as an intermediary, like (red meat intake, amount calories) on a number of health outcomes (liver cancer, pregnancy complications, risk of arterial disease) (Wu et al., 2003) (Wittenbecher et al., 2015) (Jacobs et al., 2015) (Aleksandrova et al., 2015).

The research group conducted in the literature of nutrition in the field of mediation analysis shows the wide use of this technique in modern research, one of the most cited articles is an article (Baron & Kenny, 1986) . On the impact of mediation and moderation however, many of the references dealt with the basic principles on the analysis of mediation (Krause et al., 2010) (Kenny, 2016) (Imai, Keele, & Tingley, 2010) (Miočević, MacKinnon, & Levy, 2017).

The frequent use of mediation in research and recent articles is evidence of the importance and value of mediation as a tool for understanding important and fundamental processes. The study carried out by the personal and psychosocial bulletin (PSPB) for each experimental article in 2007 and for a period of six months. This study showed that 41% in at least one study is a mediation test (Kashy, Donnellan, Ackerman, & Russell, 2009).

In 2011-2012 bulletin 16% of articles in the psychological Sciences of mediation analysis (Hayes & Scharkow, 2013), Journal of personal and social psychology JPSP included 59% of articles and 65% of PSPB in mediation testing where the study period was from 2005-2009 (Rucker, Preacher, Tormala, & Petty, 2011) (Baron & Kenny, 1986) . As we mentioned, the article (Baron & Kenny 1986) is the most cited article in the field of mediation, where the statistics of web of science that the number of citations reached 20,326 in June of 2013 (Yanai, Okada, Shigemasu, Kano, & Meulman, 2013).

2- Multiple Mediator Model

When studying any process between two variables, it is possible that there is a set of mediation variables in that process. When using a single mediation model, a number of other intermediaries are neglected and these variables may be of great importance to change the results (Briggs, 2006).The researcher uses a multi-mediation model when he has a set of variables that transfer the effect from the independent variable to the dependent variable. multiple mediation model can be uses by few researchers to compare with the single mediation model this is due to the ambiguity of analytical methods as well as the difficulty of multiple mediation models, from the few authors who use the multiple model (Cheung, 2007) (West & Aiken, 1997) (D. P. MacKinnon, 2000).

Examples of the multi-modal model are the study done by (Wen, 2013) ,which is to learn methods of teaching at the student level, the researcher used three mediators (student confidence level, student motivation level, student cooperation level) Assuming that the independent variable is the teacher's experience , and the dependent variable is the student achievement. (Preacher & Hayes, 2008) also developed a number of advantages to define and test a multi-mediation model:

1- Testing the effect of mediation of the independent variable on the dependent variable is similar to the regression analysis with several predictions, with a view to whether there is a comprehensive impact.

2- It is possible to identify the intermediaries that transfer the effect from the independent variable to the result variable.

3-When more number of intermediaries are included in the multi-mediation model this leads to a reduction of bias in the parameters, in the case of a single mediator model, these models may suffer from the problem of the deleted variables, which leads to the estimation of the biased parameters.

4- Put many intermediaries in one model allows the researcher to determine the relative size of the effects of mediation.

Researchers differ in naming variables used in mediation analysis (D. MacKinnon, 2012). describes the mediator as the variable that conveys the effect of causal advances on its descendant. Mediation variables are also called operations because they describe the process that an independent variable effects on a child variable (Judd & Kenny, 1981) , they are called alternative or intermediate ends in medical literature (Prentice, 1989). The purpose of these different labels is to have the names of the variables accurate about the data used and the nature of the study.

In our study we will use a description of the variables as an independent variable, an intermediate variable, and a dependent variable. Mediation analysis is a way to increase the information obtained from research studies when mediation tracks are available. There are three main ways to analyze statistical intermediation (D. P. MacKinnon, 2000):

1- Causal steps

2- Difference in transactions

3- Transaction output

3- Mediation for Interpretation and Design

There are two main uses of the mediating variables in research studies. The first of these uses is the interpretation, when a relationship is created between the independent variable **X** and the dependent variable **Y**, the researchers work to explain why the relationship between the two variables occurs. In this case, the purpose of the mediation analysis is to investigate the processes underlying the relationship between the two variables, examples of this use is the analysis of mediation in psychology as well as in sociology (Lazarsfeld, 1955).

The second method of mediation works to identify the pre-mediated variables related to the dependent variable, rather than explaining the relationship between two variables. In determining these mediation mechanisms, manipulation is designed to change with the selected mediation variables. If the correlation is correct between the mediation variables and the result in this case, manipulation of the mediator leads to a change in the result variable. Recently, the use of mediation in various studies using the design approach has increased because of the usefulness of this approach in applied research.

Mediation in interpretation is more commonly used in basic research to explain the apparent relationship between two variables. Design mediation is of primary use in applied empirical studies (D. MacKinnon, 2012).

4- Evaluating Mediation Utilize Regression Equation

In order to investigate the mediation of a binary mediation model, the following equations should be :

$$\mathbf{Y = i_1 + CX + e_1} \;\ldots\ldots\ldots\ldots\; \mathbf{1}$$

$$\mathbf{Y = i_2 + C`X + b_1M_1 + b_2M_2 + e_2} \;\ldots\ldots\ldots\; \mathbf{2}$$

$$\mathbf{M_1 = i_3 + a_1X + e_3} \text{ ……………3}$$

$$\mathbf{M_2 = i_4 + a_2X + e_4} \text{ ……… 4}$$

Where **Y:** is the dependent variable.

X: is the independent variable.

M1,**M2:** are the mediation variables.

C`: The parameter that links the independent variable to the dependent variable under the influence of the mediation variables.

C: The parameter that links the independent variable with the dependent variable on the absence of the mediation variables.

a1: represents the parameter that links the independent variable with the first mediation variable.

a2: represents the parameter that links the independent variable with the second mediation variable.

b1: The parameter that links the first mediation variable with the dependent variable.

b2: The parameter that links the second mediation variable and the dependent variable.

$\mathbf{e_1, e_2, e_3}$: represents the errors and their resolutions i1, i2, i3, equation 1 represents the estimation of the direct effects of the absence of the intermediation variables. equations 2,3,4 are used to estimate the indirect effect, which determines the mediation model in figure (1) (D. MacKinnon, 2012).

5- The Total Effect

In the multi-mediation model shown in figure (1), that contains two mediation variables, the total effect is the sum of the direct effect of parameter **C`** and the indirect effect of the **M1**, **M2** (Hayes, 2009).

$$\text{Total effect} = \mathbf{C`} + \mathbf{a_1b_1} + \mathbf{a_2b_2}$$

6- Mediation Effects

The overall effect can be divided into direct effects of parameter **C`** and indirect, represented by the effects of the medium $\mathbf{a_1b_1}$, $\mathbf{a_2b_2}$. in the analysis of the mediation, it should be $\mathbf{a_1b_1 + a_2b_2 = C - C`}$. in the case of inequality, the reasons are due to differences in sample sizes between the equations. the parameters of the above models are estimated using different methods, including the smaller squares method (Preacher & Hayes, 2008).

7- Significance Test and Confidence Intervals for Mediation Effects

One of the most widely used methods to test significance of the effect of mediation (**ab**) is to estimate the standard error and compare the resulting **Z-** scores with the critical value of standard normal distribution (D. P. MacKinnon, Lockwood, & Williams, 2004). Also the standard error and the estimated mediation effect can be used to build confidence intervals (C.I) for the mediation effect, its well know that the (C.I) used the standard error in estimation, for this reason, we note that (C.I) used possibly to provide a number of effect values instead of a single value .Confidence intervals are popular tools used in research because they require the researcher to consider the value of the effect as well as its statistical significance (Harlow, n.d.) . In order to test the significance of the indirect effect (median), we need to find the standard error of the sample through the intermediary **(ab)**.The most widely used tests to estimate the standard errors of the indirect effect are as follows:

7-1 Steps to Establish Mediation

As we mentioned in the single mediation model, the multiple mediation model contains a set of steps but more broadly.

Step 1: The independent variable **X** should affect the **Y** result variable through parameter **C`** as shown in equation 1.

Step 2: The independent variable **X** must affect both intermediaries M_1, M_2 by landmarks a_1, a_2 and shown in equations 3 and 4.

Step 3: Mediation variables should affect the dependent variable Y after controlling the independent variable **X** by factors b_1, b_2, as shown in equation 2.

Step 4: The effect of the independent variable **X** on the dependent variable **Y** should be unimportant (direct effect **C`**) to achieve full mediation as in Equation 2, but if there is an independent variable effect on the dependent variable **C`** there will be partial mediation (Wen, 2013).

7-2 Product of Parameter Coefficients Testing

The standard error formula for the internal effect of the multiple mediation model is the same as the one used in the single mediation model where the standard error is given as below :

$$S_{a1b1} = \sqrt{S_{a1}^2 \ b_1^2 + S_{b1}^2 \ a_1^2}$$

Other formulas for standard error can be used for a multiple mediation model which differs from that described in the single mediation model in the mediation variables where :

$$S_{a1b1} + S_{a2b2} = \sqrt{S_{a1}^2 \ b_1^2 + S_{b1}^2 \ a_1^2 + S_{a2}^2 \ b_2^2 + S_{b2}^2 \ a_2^2 + 2a_1 \ a_2 \ S_{b1b2}}$$

The equation above can be rewritten as follows :

$$S_{a1b1} + S_{a2b2} = \sqrt{S_{a1b1}^2 + S_{a2b2}^2 + 2a_1 a_2 \ S_{b1b2}}$$

$2b_1b_2 \ S_{a1\,a2}$ of the equations above should be added when there is a non-zero variation between a_1, a_2 .

Where S_{b1b2} represents the common variation.

In the case of the total effect of the broker, the standard error is given as follow :

$$S_{c-c`} = \sqrt{S_c^2 + S_{c`}^2 - 2r \, S_c S_{c`}}$$

Estimation of the median effect and standard error can be used to construct confidence limits as in equations (2-3-7-2) for similar periods (D. MacKinnon, 2012).

7-3 Bootstrapping

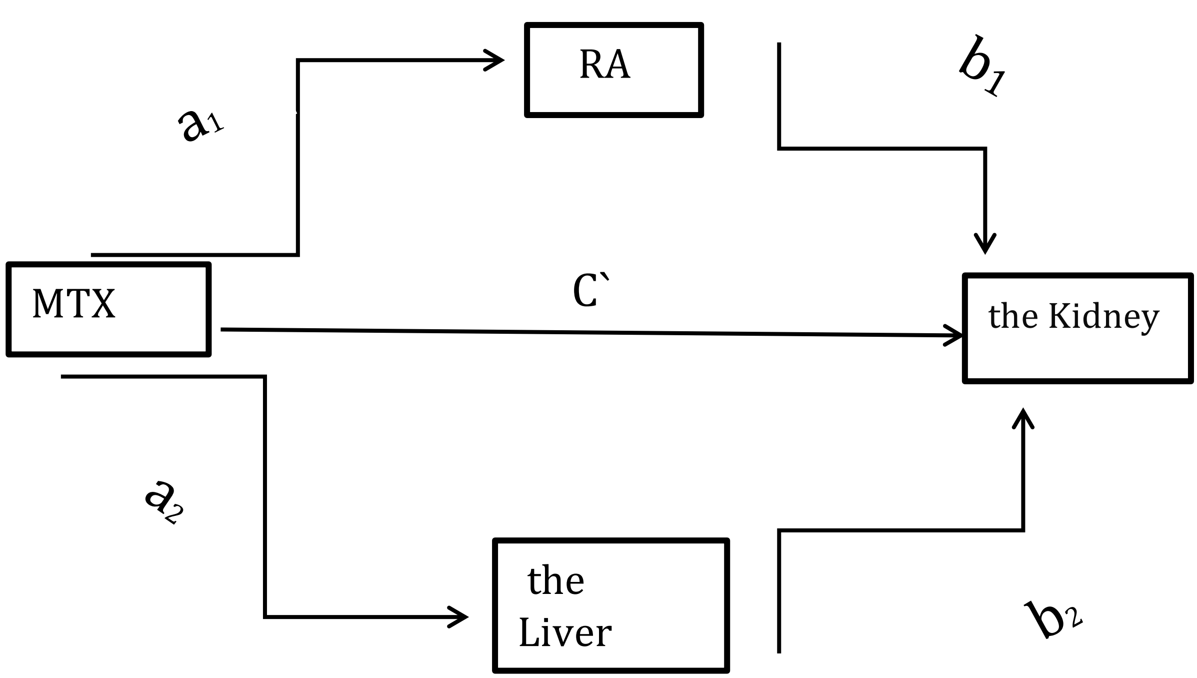
In many situations the sample size is not enough to study, this is one of the problems facing researchers in many studies and to overcome this problem, bolger & shrout suggested in 2002, the use of Bootstrapping method. Where the Delta Method to estimate the standard error does not work well when the sample size is not sufficient large .This limitation leads to use the Bootstrapping resampling method.

The Bootstrapping can be used to find the standard errors of the estimated parameter when the sampling distribution of the estimated parameter is unknown, as well as Bootstrapping can be used to assess the accuracy of the estimated model and to assess its prediction accuracy (Wen, 2013).

The Bootstrapping method is one of the common ways of estimating the indirect effects. Bootstrapping depends on the restructuring with the replacement of a large number of times. Meaning that if we have a sample of size n the bootstrapping process is performed by taking **K** of the samples with the repetition and the replacement process from the original sample size and preferably **K = 1000** at least (Preacher & Hayes, 2008).

Each sample has its own characteristics, such as the medium as well as estimating the indirect effects of each sample. This method is used to conduct sample distributions as a basis for confidence intervals and hypothesis testing (**Keeny,2018).**

In this model data were used for rheumatoid arthritis Where data was collected from the Marjan Educational Hospital in Babylon, Data were obtained through follow-up of monthly examinations of people with the disease who are taking therapy Methotrexate (MTX) treatment, a sample of 40 patients was collected During 2018 records of patients in the hospital. Treatment (MTX) was used as an independent variable (**X**) and the effect of treatment on the disease directly as a variable first mediation (**M1**), The effect of treatment on the liver directly (reduced the ratio of the Album) as a second mediation variable (**M2**), and the effect of treatment on the kidney directly (causing the increase of urea) as a dependent variable (**Y**).



Figure(2): the pathway between the effect of Methotrexate MTX (X) and Rheumatoid Arthritis (RA) (M1) and the Liver function (M2) and the Kidney function (Y) .

8- Mediation Analysis of Rheumatoid Arthritis Data

- A- Causal procedure strategy to Establish mediation We check the correlation between the effect of Methotrexate MTX (X) and Rheumatoid Arthritis RA (M1) and the liver function (M2) and the kidney function (Y) by tracking the relationship path using the correlation matrix as shown in the bellows :

	M1	M2	X	Y
M1	1	-0.30499869	-0.0473036	0.1402038
M2	-0.30499869	1	0.04340307	-0.1206272
X	-0.04730367	0.04340307	1	0.4621566
Y	0.14020385	-0.12062718	0.46215660	1

Table (1) : Correlation matrix among independent variable, dependent variable and mediators.

The correlation matrix shown in table 3.11 shows a correlation between methotrexate and renal function (0.4621), There is also a correlation between methotrexate and liver function (0.0434). There is also a correlation between methotrexate and rheumatoid arthritis (-0.0473). There is also a correlation between Rheumatoid Arthritis (RA) and renal function rate (0.1402), and there is another correlation between renal function and liver function (-0.1206).

-B- Product of Coefficient Approach

When using lm regression method results in R programs ,we have the following estimates of proposed model :

$Y = i_1 + CX + e_1$ ………………….. Equation I

Equation I shows the total (direct) effect of the treatment (MTX) on the kidney.

Parameters	$\hat{C}$

Estimate	4.0748
Std. Error	0.2065
T value	19.73
Pr(>|t|)	<2e-16
C.I	4.479-3.670

Residual standard error	14.13
Multiple R-squared	0.9089
Adjusted R-squared	0.9066
F-statistic	389.3
p-value	< 2.2e-16

Table (2): Estimates total effect (Direct effect) $\hat{C}$.

From table (3.12) we observe the total effect estimate (Direct effect) of the treatment (MTX) on kidney function at a rate of (4.0748) and a standard error at a rate of (0.2065) and there is a significant effect $p<0.05$.

The total error in the model was estimated at (14.13), and with a coefficient of determination of (0.9089), ie that it is possible to determine (0.9089) of the change in the kidney function depending on the treatment (MTX) , while the significance of the model in general is ($p< 2.2e-16$) This indicates a significant model.

From table (3-12) we obtain the following estimations:

Total effect ($\hat{C}_{\text{direct effect}}$) = 4.0748
Std. Error ($S_{\hat{C}}$) = 0.2065

Upper Confidences Limits (UCL) = $\hat{C}_{\text{direct effect}} + Z^*S_{\hat{C}}$

$$= 4.0748 + 1.96\ (0.2065) = 4.479$$

Lower Confidences Limits (LCL) = $\hat{C}_{\text{direct effect}} - Z^*S_{\hat{C}}$

$$= 3.670$$

$$Y = i_2 + C`X + b_1M_1+b_2M_2+ e_2 \quad \text{………………………… Equation (II)}$$

Equation (II) represents the effect of the treatment (MTX) on the kidney (partial effect) with another indirect effect (Mediation effect) represented by rheumatism (M1) and liver (M2).

Parameters	$\hat{C}$`	$\hat{b}1$	$\hat{b}2$
Estimate	1.6287	0.8646	10.1430
Std. Error	0.3161	0.1241	3.4672
T value	5.153	6.967	2.925
Pr(>|t|)	8.79e-06	3.14e-08	0.00585
C.I	2.248-1.0091	1.107-0.621	16.94-3.35

Residual standard error	8.439
Multiple R-squared	0.9692
Adjusted R-squared	0.9667
F-statistic	387.9
p-value	< 2.2e-16

Table (3): Estimates Partial effect and The parameters $\hat{b}_1$ and $\hat{b}_2$.

From table (3.13) we observe the partial effect estimate of the treatment (MTX) on kidney function at a rate of (1.6287) and a the standard error at a rate of (0.3161) and estimate the effect of rheumatism (effect of mediation 1) at a rate of by (0.8646) and a the standard error at a rate of (0.1241) and estimate the effect of the liver (effect of mediation 2) at a rate of by (10.1430) and a the standard error at a rate of (3.4672) , there is a significant effect where $p<0.05$.

The total error in the model was estimated at (8.439), and with a coefficient of determination of (0.9692), ie that it is possible to determine (0.9692) of the change in the kidney function depending on the Rheumatism (RA) and the liver , while the significance of the model in general is ($p< 2.2e-16$) This indicates a significant model.

From table (3-13) we obtain the following estimations:

Partial effect ($\hat{C}`_{\text{partial effect}}$) = 1.6287

Std. Error ($S_{\hat{C}}$) = 0.3161

Upper Confidences Limits (UCL) = $\hat{C}_{\text{partial effect}} + Z^* S_{\hat{C}}$

= 1.6287+ 1.96 (0.3161) = 2.248

Lower confidences limits (LCL) = $\hat{C}_{\text{partial effect}} - Z^* S_{\hat{C}}$

= 1.0091

$\hat{b}_{1\text{ indirect effect}}$ = 0.8646

Std. Error ($S_{\hat{b}1}$) = 0.1241

Upper Confidences Limits (UCL) = $\hat{b}_{1\text{ indirect effect}} + Z^* S_{\hat{b}1}$

= 0.8646 + 1.96 (0.1241)

= 1.107

Lower Confidences Limits (LCL) = $\hat{b}_{1\text{ indirect effect}} - Z^* S_{\hat{b}1}$

= 0.621

$\hat{b}_{2\text{ indirect effect}}$ = 10.1430

Std. Error ($S_{\hat{b}2}$) = 3.4672

Upper Confidences Limits (UCL) = $\hat{b}_{2\text{ indirect effect}} + Z^* S_{\hat{b}2}$

= 10.1430 + 1.96 (3.4672)

=16.938

Lower Confidences Limits (LCL) = $\hat{b}_{2\text{ indirect effect}} - Z^* S_{\hat{b}2}$

= 3.347

$M_1 = i_3 + a_1 X + e_3$ ……………………………. Equation (III)

Equation (III) represents the effect of treatment (MTX) on Rheumatism (RA) in a manner direct .

Parameters	$\hat{a}_1$
Estimate	2.1244

Std. Error	0.1647
T value	12.9
Pr(>|t|)	1.19e-15
C.I	2.447-1.801

Residual standard error	11.27
Multiple R-squared	0.8101
Adjusted R-squared	0.8053
F-statistic	166.4
p-value	<1.193e-15

Table (4): Estimates parameter $\hat{a}_1$.

From table (3.14) we observe the total effect estimate of treatment (MTX) In a manner direct on Rheumatism (RA) at a rate of (2.1244) and a the standard error at a rate of (0.1647) , there is a significant effect where $p<0.05$.

The total error in the model was estimated at (11.27), and with a coefficient of determination of (0.8101), Ie that it is possible to determine (0.8101) of the change in the Rheumatism (RA) depending on the treatment (MTX) , while the significance of the model in general is ($p< 1.193e-15$) this indicates a significant model.

From table (3-14) we obtain the following estimations:

$\hat{a}_{1\ \text{direct effect}} = 2.1244$

Std. Error ($S_{\widehat{a1}}$) = 0.1647

Upper Confidences Limits (UCL) = $\hat{a}_{1\ \text{direct effect}} + Z^*S_{\widehat{a1}}$

$= 2.1244 + 1.96\ (0.1647)$

$= 2.447$

Lower Confidences Limits (LCL) = $\hat{a}_{1\ \text{direct effect}} - Z^*S_{\widehat{a1}} = 1.801$

$M_2 = i_4 + a_2X + e_4$ …………………….. Equation (IV)

Equation (IV) represents the effect of treatment (MTX) on the kidney function in a manner direct .

Parameters	$\hat{a}_2$
Estimate	0.060080
Std. Error	0.005894
T value	10.19
Pr(>|t|)	1.48e-12
C.I	0.0713-0.049

Residual standard error	0.4033
Multiple R-squared	0.7271
Adjusted R-squared	0.7201
F-statistic	103.9
p-value	< 1.484e-12

Table (5): Estimates parameter $\hat{a}_2$.

From table (3.14) we observe the total effect estimate of treatment (MTX) In a manner direct on the kidney function at a rate of (0.060080) and a the standard error at a rate of (0.005894) there is a significant effect where $p<0.05$.

The total error in the model was estimated at (0.4033), and with a coefficient of determination of (0.7271), ie that it is possible to determine (0.7271) of the change in the kidney function depending on the treatment (MTX) , while the significance of the model in general is ($p<1.484e-12$) this indicates a significant model.

From table (3-14) we obtain the following estimations:

$\hat{a}_{2\ \text{indirect effect}} = 0.060080$

Std. Error ($S_{\widehat{a2}}$) = 0.005894

Upper Confidences Limits (UCL) = $\hat{a}_{2\ \text{direct effect}} + Z^*S_{\widehat{a2}}$

$= 0.060 + 1.96\ (0.0058) = 0.0713$

Lower Confidences Limits (LCL) = $\hat{a}_{2\ \text{direct effect}} - Z^*S_{\widehat{a2}}$

$= 0.0486$

Standard error can also be calculated:

$\hat{a}_{1\text{indirect effect}} * \hat{b}_{1 \text{ indirect effect}} = 2.1244 * 0.8646$

$= 1.836756$

$\hat{a}_{2 \text{ indirect effect}} * \hat{b}_{2 \text{ indirect effect}} = 0.060080 * 10.1430$

$= 0.6093914$

It was found that MTX (**X**) treatment significantly affected renal function (**Y**) ($\hat{C}_{\text{direct effect}} = 4.0748$, Std. Error ($S_{\hat{C}}$) = 0.2065, $t_{\hat{c}}$ = 19.73), providing evidence of a statistically significant intervention effect at 4.0748 units. The effect of MTX (**X**) was statistically significant for both RA ($\hat{a}_{1 \text{ indirect effect}}$ = 2.1244, Std. Error ($S_{\widehat{a1}}$) = 0.1647, and $t_{\widehat{a1}}$ = 12.9). As well as on liver function (**M2**)

($t_{\widehat{a2}}$ = 10.19, $\hat{a}_{2 \text{ indirect effect}}$ = 0.060080, Std. Error ($S_{\widehat{a2}}$) = 0.005894).

The effect of rheumatoid arthritis was also statistically significant

($\hat{b}_{1 \text{ indirect effect}}$ = 0.8646, Std. Error ($S_{\widehat{b1}}$) = 0.1241, $t_{\widehat{b1}}$ = 6.967), The results of liver function also showed a statistically significant effect ($\hat{b}_{2 \text{ indirect effect}}$ = 10.1430, Std. Error ($S_{\widehat{b2}}$) = 3.4672, $t_{\widehat{b2}}$ = 2.925).

The treatment of MTX has had a difference in the proportion of rheumatoid arthritis as well as a change in the proportion of liver function, leading to a change in renal function. Where the effect was statistically significant

(Partial effect ($\hat{C}`_{\text{partial effect}}$) = 1.6287, Std. Error ($S_{\hat{C}`}$) = 0.3161, $t_{\hat{C}`}$) = 5.153).

As well The average effect equal:

$\hat{a}_{1 \text{ indirect effect}} * \hat{b}_{1 \text{ indirect effect}} + \hat{a}_{2 \text{ indirect effect}} * \hat{b}_{2 \text{ indirect effect}} = 1.836756 + 0.6093914 = 2.4461$

$\hat{C}_{\text{direct effect}} - \hat{C}`_{\text{partial effect}} = 4.0748 - 1.6287 = 2.4461$

So

$\hat{a}_{1 \text{ indirect effect}} * \hat{b}_{1 \text{ indirect effect}} + \hat{a}_{2 \text{ indirect effect}} * \hat{b}_{2 \text{ indirect effect}} =$

$\hat{C}_{\text{direct effect}} - \hat{C}`_{\text{partial effect}}$

The median effect of MTX therapy through changes in both rheumatoid arthritis and liver function was 2.4461 renal function change units.

Standard errors can be calculated using the standard error equation to estimate the mean effect as shown in the following:

$$S_{\hat{a}1\widehat{b1}} = \sqrt{s^2_{\widehat{a1}}\hat{b}^2_1 + \hat{a}^2_1\, s^2_{\widehat{b1}}}$$

$$= \sqrt{(0.1647)^2\,(0.8646)^2 + (0.1241)^2\,(2.1244)^2}$$

$= 0.2996376$

$$S_{\hat{a}2\widehat{b2}} = \sqrt{s^2_{\widehat{a2}}\hat{b}^2_2 + \hat{a}^2_2\, s^2_{\widehat{b2}}}$$

$$\sqrt{(0.005894)^2\,(10.1430)^2 + (3.4672)^2\,(0.060080)^2} =$$

$= 0.2167182$

$$S_{\hat{a}1\widehat{b1}} + S_{\hat{a}2\widehat{b2}} = \sqrt{s^2_{\hat{a}1\hat{b}1} + s^2_{\hat{a}2\hat{b}2} + 2a_1\, a_2\, S_{\hat{b}1\widehat{b2}}}$$

$$= \sqrt{0.2996376 + 0.2167182 + 2(2.1244)(0.060080)(0.0845)}$$

$= 0.5379259$

The confidence limits were as follows:

Upper Confidences Limits (UCL) = Mediated effect + $Z^* S_{\hat{a}1\widehat{b1}}$

$= 2.4461 + 1.96\,(0.2996376)$

$= 3.03339$

Lower Confidences Limits (LCL) = Mediated effect – $Z^* S_{\hat{a}1\widehat{b1}}$

$= 2.4461 - 1.96\,(0.2996376)$

$= 1.85881$

Upper Confidences Limits (UCL) = Mediated effect + $Z^* S_{\hat{a}2\widehat{b2}}$

$= 2.4461 + 1.96\,(0.2167182)$

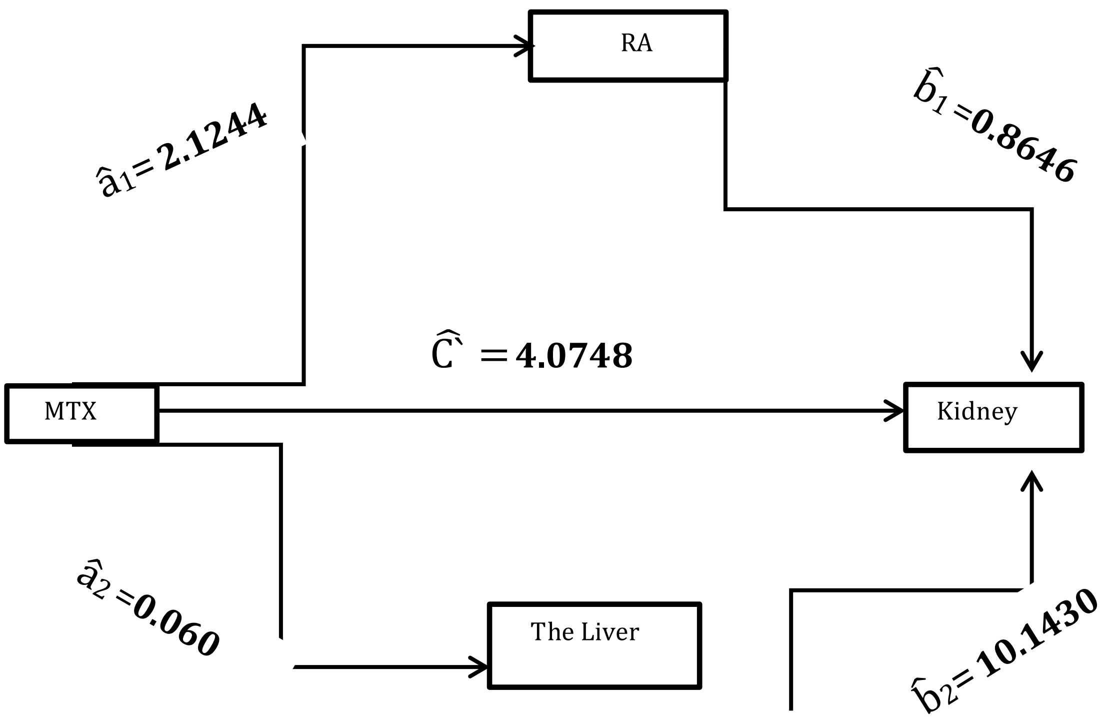
$= 2.870868$

Lower Confidences Limits (LCL) = Mediated effect – $Z^* S_{\hat{a}2\widehat{b2}}$

$$= 2.4461 - 1.96\ (0.2167182)$$

$$= 2.021332$$

All results show that there is statistical significance and significance of the effect of the variables (MTX, rheumatoid arthritis, liver function, renal function)



Figure(3): Shows estimates of variables on the chart .

- C - Bootstrapping Estimation

Either when using the bootstrapping method be estimate the intervals of confidence as follows :

In order to obtain more accuracy, in bootstrapping, I chose to take 1500 samples of 5,000 cases with replacement from the original sample and calculated each mediated effect. Appendix includes the R code with sample command of Bootstrapping.

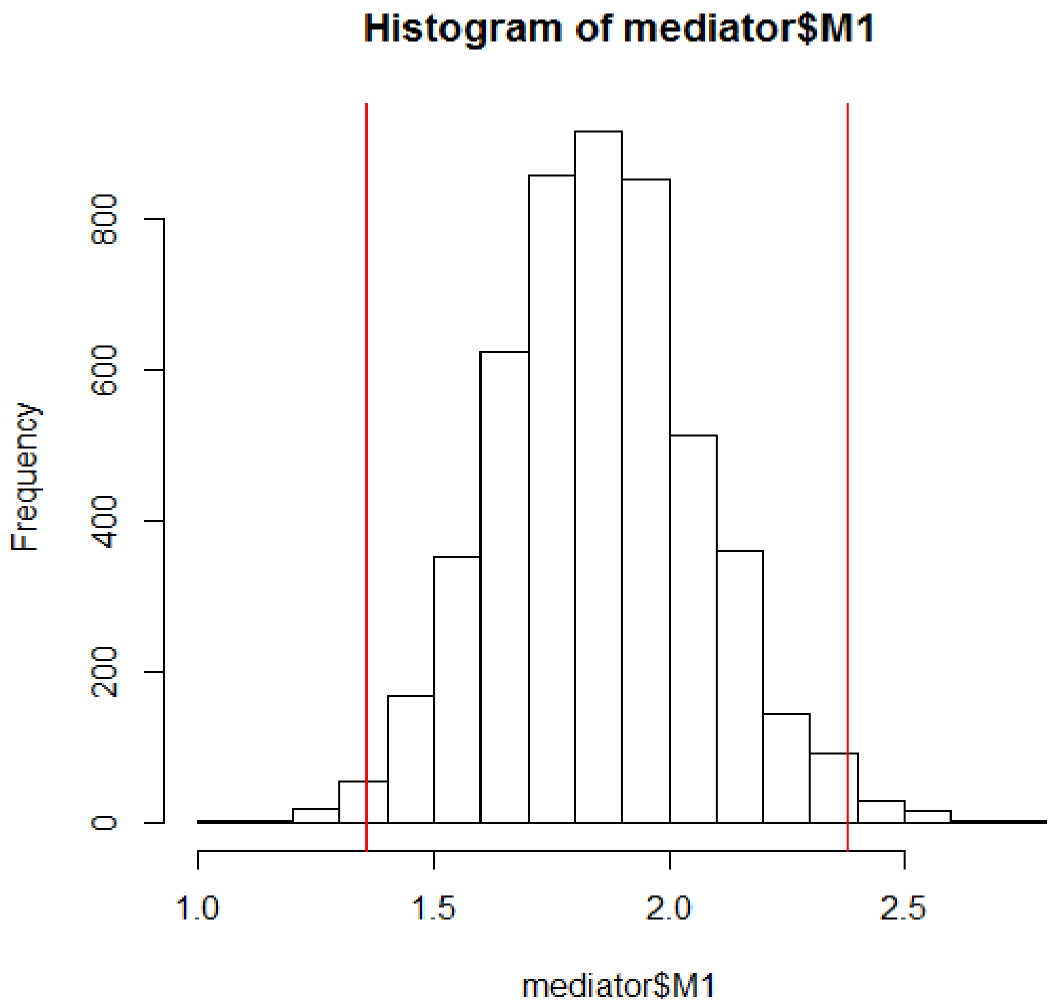
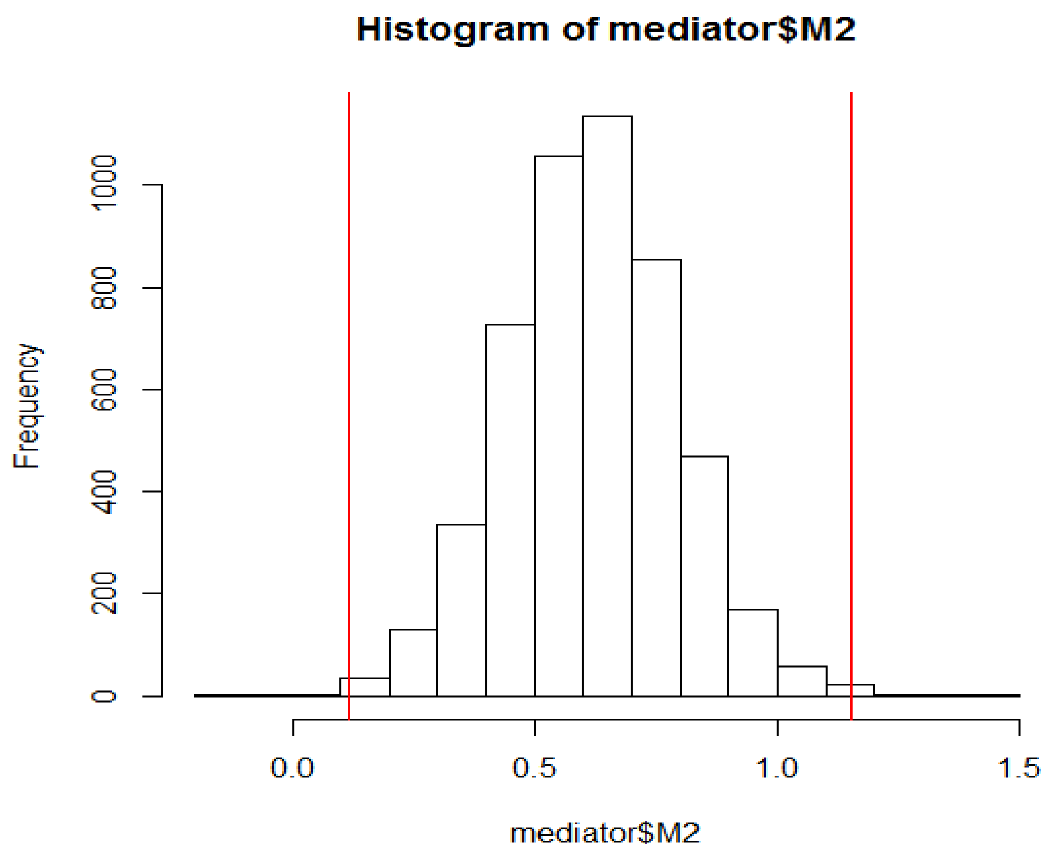


Figure (4): The diagram of the effect is shown by the kidneys. The red lines represent the minimum and the highest for every 95% of the confidence interval.

The above diagram shows the confidence interval for the first mediation variable (Rheumatoid Arthritis) with a confidence interval of 95%. Since the period is limited between (1.3 - 2.4), ie, zero does not belong to this period, it can be said that the effect of the mediation variable is different from zero and therefore there is an effect for this variable.



Figure(5): The diagram of the effect is shown by the liver. The red lines represent the minimum and the highest for every 95% of the confidence interval.

The above diagram shows the confidence interval for the second mediation variable of (The Liver) with a confidence interval of 95%. Since the period is limited between (0.2 - 1.1), ie, zero does not belong to this period, it can be said that the effect of the mediation variable is different from zero and therefore there is an effect for this variable.

9- Conclusion

We can conclude from the results obtained from analyzing the data of the multiple mediation model of rheumatoid arthritis in the previous chapter on the following:

1- The effect of the independent variable (MTX) on the dependent variable (renal function) (total effect) is significant and at a rate of (4.0748).

2- The effect of the independent variable (MTX) on the dependent variable (renal function) (partial effect) with mediation variable a presence is significant and at a rate of (4.0748).

3 - Effect of the first mediation variable (rheumatoid arthritis) on the dependent variable (renal function) is significant and by at a rate of (0.8646).

4- The effect of the second mediating variable (liver function) on the dependent variable (renal function) is significant and by at a rate of (10.1430).

By observing the correlation between the studied variables and the estimation of the parameters, as well as the estimated trustworthiness, it turns out that the method of shoeboxes is better than striking the transactions.

References

Aleksandrova, K., Bamia, C., Drogan, D., Lagiou, P., Trichopoulou, A., Jenab, M., … Pischon, T. (2015). The association of coffee intake with liver cancer risk is mediated by biomarkers of inflammation and hepatocellular injury: data from the European Prospective Investigation into Cancer and Nutrition–. *The American Journal of Clinical Nutrition*, *102*(6), 1498–1508.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173.

Briggs, N. E. (2006). Estimation of the standard error and confidence interval of the indirect effect in multiple mediator models. The Ohio State University.

Cheung, M. W. L. (2007). Comparison of approaches to constructing confidence intervals for mediating effects using structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(2), 227–246.

Grotta, A., & Bellocco, R. (2012). Causal mediation analysis on survival data: an application on the National March Cohort. PhD thesis, Univ. Milano-Bicocca, Milan.

Harlow, L. L. (n.d.). Mulaik. SA, and Steiger, JH, Eds., 1997. *What If There Were No Significance Tests*.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*(4), 408–420.

Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter? *Psychological Science*, *24*(10), 1918–1927.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309.

Jacobs, S., Schiller, K., Jansen, E. H. J. M., Boeing, H., Schulze, M. B., & Kröger, J. (2015). Evaluation of various biomarkers as potential mediators of the association between Δ5 desaturase, Δ6 desaturase, and stearoyl-CoA desaturase activity and incident type 2 diabetes in the European Prospective Investigation into Cancer and Nutrition–Potsdam . *The American Journal of Clinical Nutrition*, *102*(1), 155–164.

Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, *5*(5), 602–619.

Kashy, D. A., Donnellan, M. B., Ackerman, R. A., & Russell, D. W. (2009). Reporting and interpreting research in PSPB: Practices, principles, and pragmatics. *Personality and Social Psychology Bulletin*, *35*(9), 1131–1142.

Kenny, D. A. (2016). Power analsis app MedPower. Learn how you can do a mediation analysis and output a text description of your results: Go to mediational analysis using DataToText using SPSS or R.

Krause, M. R., Serlin, R. C., Ward, S. E., Rony, R. Y. Z., Ezenwa, M. O., & Naab, F. (2010). Testing mediation in nursing research: Beyond Baron and Kenny. *Nursing Research*, *59*(4), 288.

Lazarsfeld, P. F. (1955). Interpretation of statistical relations as a research operation. *The Language of Social Research: A Reader in the Methodology of Social Research*, 115–125.

MacKinnon, D. (2012). *Introduction to statistical mediation analysis*. Routledge.

MacKinnon, D. P. (2000). Contrasts in multiple mediator models. *Multivariate Applications in Substance Use Research: New Methods for New Questions*, 141–160.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annu. Rev. Psychol.*, *58*, 593–614.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, *39*(1), 99–128.

Miočević, M., MacKinnon, D. P., & Levy, R. (2017). Power in Bayesian mediation analysis for small sample research. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(5), 666–683.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891.

Prentice, R. L. (1989). Surrogate endpoints in clinical trials: definition and operational criteria. *Statistics in Medicine*, *8*(4), 431–440.

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, *5*(6), 359–371.

Wen, S. (2013). Estimation of multiple mediator model.

West, S. G., & Aiken, L. S. (1997). Toward understanding individual effects in multicomponent prevention programs: Design and analysis strategies.

Wittenbecher, C., Mühlenbruch, K., Kröger, J., Jacobs, S., Kuxhaus, O., Floegel, A., … Adamski, J. (2015). Amino acids, lipid metabolites, and ferritin as potential mediators linking red meat consumption to type 2 diabetes–. *The American Journal of Clinical Nutrition*, *101*(6), 1241–1250.

Wu, H., Dwyer, K. M., Fan, Z., Shircore, A., Fan, J., & Dwyer, J. H. (2003). Dietary fiber and progression of atherosclerosis: the Los Angeles Atherosclerosis Study. *The American Journal of Clinical Nutrition*, *78*(6), 1085–1091.

Yanai, H., Okada, A., Shigemasu, K., Kano, Y., & Meulman, J. J. (2013). *New Developments in Psychometrics: Proceedings of the International Meeting of the Psychometric Society IMPS2001. Osaka, Japan, July 15–19, 2001*. Springer Science & Business Media.