Paper SD64

Using SAS to Examine Mediator, Direct and Indirect Effects of Isolation and Fear on Social Support Using Baron& Kenny Combined with Bootstrapping Methods

Abbas S. Tavakoli, DrPH, MPH, ME; Sue P. Heiney PhD, RN

Abstract

This study presentation examines mediator, direct and indirect effects of isolation and fear on social support by using two methods: Baron & Kenny, and Bootstrapping. This paper used a cross-sectional data from the longitudinal study randomized trial design in which 185 participants were assigned to the therapeutic group (n=93) who received by teleconference with participants interacting in real time with each other and control group (n=92) who received usual psychosocial care (any support used by the patient in the course of cancer treatment. Baron and Kenny (1986) steps and Hayes (2004) were used to examine for direct and indirect effects. Results of Baron indicated that the relationship between fear and social support was significant (c =-1.151 (total effect) (p=.0001)) and that there was significant relationship between isolation and fear (**α** =1.22 (p=.0001)). Also, previously significant relationship between fear and social support was not significant (c' =-.40 (direct effect) (p=.1876) when both fear and isolation were in the model. The indirect effect was -1.11 and Sobel test was significant (P=.0001). The results of bootstrapping methods indicated the direct effect wares -.41 (95% CI: -.42, -.40 for normal theory and -.41 (95% CI: -.99, .14 for percentile) and indirect effect was -1.06 (95% CI: -1.09, -1.08 for normal theory and -1.09, -1.55 for percentile). The result showed both methods had significant indirect effect.

Keywords: SAS, Mediator, Direct, Indirect.

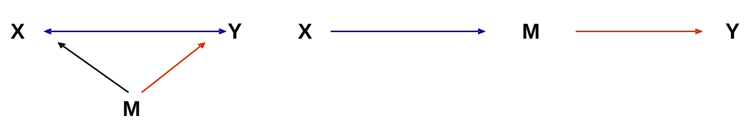
University of South Carolina, College of Nursing.

The project described was supported by Award Number R01CA107305 from the National Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

At the time a portion of this research was conducted Heiney was employed as Manager, Psychosocial Oncology at Palmetto Health Cancer Centers (PHCC) in Columbia, SC.

Introduction

Mediation is common in the health science. There is often problem to explain why the relationship between explanatory variable (X) and criterion variable (Y) become non- significant when we introduce the third variable in the model. This could be the third variable functions as a mediator. Baron & Kenny (1986) define a mediator as a variable to the extent that it accounts for the relation between the independent variable and the outcome variable. Diagrammed below showed two possibilities of mediator effect:



Partial Full

The general model is described in terms of mediated effects. We assume multivariate normal distributions and normally distributed error terms throughout. Two ways can be calculated the effect of adding a third variable based on either the difference between two regression parameters (c - c') (indirect effect) or the multiplication of two regression parameters (**α** β). **In the first method**, the following two regression equations are estimated.

Model 1: $Y = i_1 + cX + \varepsilon_1$
Model 2: $Y = i_2 + c' X + \beta M + \varepsilon_2$
Model 3: $M = i_3 + \alpha X + \varepsilon_3$

Where Y is the outcome or dependent variable, X is independent variable, M is the mediator, c (total effect) codes the relationship between the independent to the outcome in the first equation, c' (direct effect) is the coefficient relating the independent to the outcome adjusted for the effects of the mediator, $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ code unexplained variability, and the intercepts are a_1 and a_2. **A second method** used model 2 and 3 to estimate indirect effect. The indirect t effect or mediator effect is calculated by multiply β coefficient from model 2 and **α** coefficient from model 3.
Baron and Kenny (1986) have proposed four steps in establishing mediation:

Step 1: Show that the independent variable is correlated with the outcome (**Model Y =X).**
Step 2: Show that the independent variable is correlated with the mediator **(Model M = X)**.
Step 3: Show that the mediator affects the relationship of independent on outcome variable **(Model Y = M X)**.
Step 4: To establish that M completely mediates the X-Y relationship, the effect of X (IV) on Y (DV) controlling for M should be zero (estimate and test path c'). The effects in both Steps 3 and 4 are estimated in the same regression equation.
Variable M **completely mediates** the X-Y relationship when all four of these steps are met, and if the first three steps only are met, then **partial mediation** is indicated.
To determine the extent to which a mediator contributed to the total effect on the outcome variable, Sobel's (1982) test is performed. Sobel test assume normality and no measurement error. Different standard error to test indirect effect was proposed by Aroian (1994) and Goodman (1960). Hayes (2010) method estimates and tests the indirect effect of independent variable (X) on dependent viable (Y) through mediation (M) by using bootstrapping.

Purpose

The purpose of this paper is using SAS to examine mediator, direct, and indirect effect of isolation in fear on social support by two methods: Baron & Kenny and bootstrapping.

Background

This paper used the cross-sectional data from research grant (R01) collected in the first of three assessments of a longitudinal study designed to test and compare the effects of a therapeutic group using teleconference for African American women with breast cancer on social disconnection, a sense of being cut off from partners, family and friends due to side effects of treatment and fatalistic beliefs about cancer. A therapeutic group by teleconference may assist African American women with breast cancer to feel connected to women in a similar situation, to learn ways to talk about cancer and to decrease fatalistic beliefs. A randomized trial design was used in which 185 participants were assigned to the therapeutic group (n=93) and control group (n=92). The therapeutic group intervention was led by two social workers experienced in working with oncology patients and leading support groups. The intervention was delivered by teleconference with participants interacting in real time with each other. Control group was defined as any support used by the patient in the course of cancer treatment. The randomization was stratified by treatment type. Data were collected at baseline, the end of the intervention (8 weeks from baseline), and 16 weeks from baseline.

Data Analyses

All data analyses were performed using **SAS/STAT** ® statistical software, version 9.4 (SAS, 2008). Proc MEAN used to describe the data. PROC CORR and REG were used to analyze this study. Bootstrapping used to estimate direct and indirect effects through the repeated sampling of data (n=5000). Also, two different methods (normal distribution theory and percentile) used to calculate confidence interval for direct and indirect effects. Pearson correlation and regression procedures were used to examine the interrelationships among the study variables. *P*-values less than or equal to .05 were considered significant.

Results

Table 1 showed descriptive statistics for social support, isolation, and fear.

Table 1. Measure of center and dispersion for variables

Variable	N	Mean	Std Dev	Minimum	Maximum
SSQ	184	100.77	13.58	44.00	120.00
Isolat	185	28.37	8.10	20.00	59.00
fear	185	2.95	3.15	0.00	16.00

*. SSQ (Social Support), isolate (Isolation), & fear (Fear).

Table 2 revealed descriptive statistics for social support, isolation, and fear using bootstrap.

Table 2. Measure of center and dispersion for variables using bootstrap

Variable	N	Mean	Std Dev	Minimum	Maximum
mssq	5000	100.79	1.00	96.88	104.36
misolat	5000	28.37	0.60	26.25	30.54
mfear	5000	2.95	0.23	2.14	3.84

*. MSSQ (Social Support), Misolate (Isolation), & Mfear (Fear).

The result showed the means from original sample and bootstrapping are identical. However, the standard deviation in bootstrapping is smaller than original sample.

Table 3 indicated pairwise Pearson correlation.

Table 3. Pairwise Pearson correlation

Pearson Correlation Coefficients Prob > |r| under H0: Rho=0 Number of Observations			
	SSQ	Isolat	fear
SSQ	1.00000	-0.56682 <.0001	-0.35118 <.0001
Isolat		1.00000 185	0.47256 <.0001 185
fear			1.00000 185

*. SSQ (Social Support), isolate (Isolation), & fear (Fear).

The results indicated negative linear relationship between social support and isolation (r=-.57) and linear negative relationship between social support and fear (r=-.35). Also, there is linear positive relationship between isolation and fear (r=.47).

Three regression equations were run to test for mediation (baron & Kenny method), First, the outcome (social support) was regressed on the predictor variable (fear).This relationship was significant (c =-1.51 (p=.0001)). There was a negative linear relationship between fear and social support. Since the first model was significant second and third regression models were run. In the second model, the mediator (isolation) was regressed on the predictor variable (fear). The result indicated that there was significant relationship between mediator and predictor variable (**α** =1.21 (p=.0001)). The third model included regressing equation both independent variable (fear) and mediator variable (isolation) on outcome (social support). The result indicated that the previously significant relationship between predictor (fear) and the outcome (social support) became non-significant (c' =-.40 (p=.1876). Therefore, there is an evidence of mediator effect for loneliness in the relationship between fear and social support (see Table 4 & Figure 1).

Table 4. Regression results Baron & Kenny method

Model : SSQ = fear (Y=X)

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	4164.61571	4164.61571	25.60	<.0001
Error	182	29604	162.66118		
Corrected Total	183	33769			

Root MSE	12.75387	R-Square	0.1233
Dependent Mean	100.76630	Adj R-Sq	0.1185
Coeff Var	12.65688		

Parameter Estimates								
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate	Squared Semi-partial Corr Type II	Squared Partial Corr Type II
Intercept	1	105.25087	1.29210	81.46	<.0001	0	.	.
fear	1	-1.51406	0.29922	-5.06	<.0001	-0.35118	0.12333	0.12333

Model: Isolation = fear (M=X)

Parameter Estimates								
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate	Squared Semi-partial Corr Type II	Squared Partial Corr Type II
Intercept	1	24.77629	0.72253	34.29	<.0001	0	.	.
fear	1	1.21683	0.16775	7.25	<.0001	0.47256	0.22332	0.22332

Model: SSQ = Isolation fear (Y = X M)

Parameter Estimates								
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate	Squared Semi-partial Corr Type II	Squared Partial Corr Type II
Intercept	1	127.18984	3.16709	40.16	<.0001	0	.	.
Isolat	1	-0.89389	0.12048	-7.42	<.0001	-0.52076	0.20445	0.23321
fear	1	-0.40027	0.30260	-1.32	0.1876	-0.09284	0.00650	0.00957

*. SSQ (Social Support), isolate (Isolation), & fear (Fear)

<u>Figure 1 Mediator Model:</u> Isolation (Isolat) as mediator of fear to social support (SSQ)

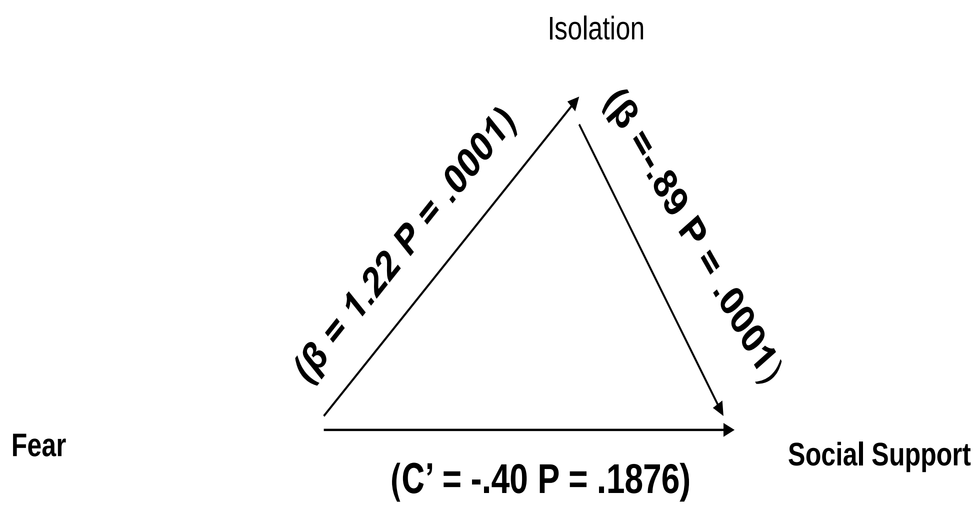
<u>Step. 1</u>

β =-1.51 (p=.0001)

Fear ------------------- --------→ Social Support (SSQ)

<u>Step 2 and 3.</u>

Figure1: Isolation (Isolat) as Mediator of fear to Social Support (SSQ).



Indirect Effect= c – c' = -1.51 – (-.40)=-1.11

Three different tests (Sobel, Goodman, and Aroian) used to examine the indirect effect for original sample. Table 5 indicated the result of the three tests. All tests were statistically significant (p less .0001) which indicated there was indirect effect.

Table 5. Indirect test results

a: SSQ(X)Coeff Model M=x	sa: Standard error (X) Model M=x	b: isolat (M) Coeff Model Y=X M	sb: SE (M) Model y = X M	Total Effect
1.24599	0.16165	-0.89389	0.12048	-1.51406

Standard error total effect	ratio indirect effect over total effect	ratio indirect effect over direct effect	Sobel test	P value Sobel test
0.29922	0.73563	2.78259	5.34541	9.0211E-8

Goodman test	P value Goodman test	Aroian test	P value Aroian test
5.32221	.000000103	5.36892	7.921E-8

Table 6 showed result of indirect effect using bootstrap method. The results of bootstrapping method indicated the direct effect was -.41 with 95% CI: -.42 to -40 for normal theory and -.99 to.14 for percentile. Also, indirect effect was -1.08 with 95% CI: -1.09 to -1.08 for normal theory and -1.09 with 95% CI -1.55 to -.66 for percentile.

Table 6. Results of Bootstrap (n=5000)

Method I: Normal Distribution CI

Means fear Bootstrap direct effect	Standard error fear	lower CI fear	upper CI fear	Means indirect effect Bootstrap	Standard error indirect effect	lower CI indirect effect
-0.40971	.004134275	-0.41652	-0.40291	-1.08582	.003183895	-1.09106

upper CI indirect effect	Means total effect Bootstrap	Standard error total effect Bootstrap	lower CI total effect	upper CI total effect
-1.08058	-1.49553	.00485855	-1.50353	-1.48754

Method 2: Percentile CI

Direct effect fear Original	Direct effect Means fear Bootstrap	direct effect lower CI fear	direct effect upper CI fear	isolat Original	Means isolat Bootstrap	lower CI isolat	upper CI isolat	indirect effect Original
-0.40027	-0.40971	-0.99101	0.14278	-0.89389	-0.89282	-1.12266	-0.67526	-1.08772

Means indirect effect Bootstrap	lower CI indirect effect	upper CI indirect effect	total effect Original	Means total effect Bootstrap	lower CI total effect	upper CI total effect
-1.08582	-1.55154	-0.66300	-1.48798	-1.49553	-2.19077	-0.84297

Conclusion

This paper examined the influence of isolation in the relationship between fear and social support, whether the relationship was influenced by a mediator effect. The result revealed that there was mediator effect for isolation in the relationship between fear and social support. The result indicated that there was significant indirect effect with both Sobel, Goodman, and Aroian tests and bootstrapping. Those tests for indirect effect are valid when the assumption of normality of the sampling distribution can be met. Bootstrapping is powerful technique to calculate confidence interval for indirect effect without any assumptions about sampling distribution. Therefore, our recommendation is to use bootstrapping to examine indirect effect.

References

1. Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical consideration. Journal of Personality & Social Psychology, 51, 1173-1182.
2. Goodman, L. A. (1960). On the exact variance of products. *Journal of the American Statistical Association*, **55**, 708-713.
3. Hayes, A. F., Preacher, K.J. (2010). Estimating and testing indirect effects in simple mediation models when the constituent paths are nonlinear. Multivariate Behavioral Research, 45, 627-660.
4. MacKinnon, D.P. & Dwyer, J.H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review*, 17, 144-158
5. Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. Sociological Methodology **13**: 290–312

Contact Information

Abbas S. Tavakoli, DrPH, MPH, ME
College of Nursing
University of South Carolina
1601 Greene Street
Columbia, SC 29208-4001
Fax: (803) 777-5561
E-mail: abbas.tavakoli@sc.edu

Attachment

SAS Syntax

```
**Let us first get a new file with 5000 ;
%let rep=5000;
 proc surveyselect data=med out=outboot
seed=3292012   method=urs  samprate=1
outhits rep=&rep; run;
**** Descriptive and correlation ***;
ods rtf; ods listing close;
proc means data=med maxdec=2;
    var    ssq isolat fear;
    title ' means '; run;

proc corr data=med;
var    ssq isolat fear;
```

```
    title ' Correlation '; run;
proc means data=outboot noprint ;
    var  var    ssq isolat fear;
 by replicate;
  output out=outall mean= mssq misolat mfear;   run;

proc means data=outall maxdec=2;
var mssq misolat mfear; run;
ods rtf close; ods listing; quit; run;

ods rtf; ods listing close;
**** Baron & Kenny method ***;
%macro regb  (d,i,t);
proc reg data=med  ;
      model &d = &i / stb pcorr2 scorr2;
      title ' Regression model/ ' &t;
       title3 ' Baron Kenny method';
     %mend regb;
%regb (ssq,fear , social support Step 1: Y=x );
%regb (isolat,fear , social support Step 2: m=x );
%regb (ssq,isolat fear , social support Step 3: Y=x m );run;
ods rtf close; ods listing; quit; run;

*** Soble , Goodman, Aroian tests ****;
data test;
            set med;            x= fear ; m= isolat ; y= ssq ;  run;

proc reg data=test noprint;
  model y=x;   model m=x;    model y=x m;
   ods output ParameterEstimates=regout;    run; quit;

* Here we select only the statistics required to compute the tests *;
 data stest;
            set regout;
            if model = 'MODEL2' and variable = 'x' then a  = Estimate;
            if model = 'MODEL2' and variable = 'x' then sa = StdErr;
            if model = 'MODEL3' and variable = 'm' then b = Estimate;
            if model = 'MODEL3' and variable = 'm' then sb = StdErr;
            if model = 'MODEL1' and variable = 'x' then te = Estimate;
    if model = 'MODEL1' and variable = 'x' then se = StdErr; run;
proc summary data=stest nway;
var a sa b sb te se; output out=stesto max= ; run;

proc format;
value Testf  1 = 'Sobel'  2 = 'Goodman'  3 = 'Aroian';

data stest2; set stesto;
zvs = (a*b)/sqrt(((b*b)*(sa*sa))+((a*a)*(sb*sb)));
abssobel = abs(zvs);
ps = 2*(1-CDF('NORMAL',abssobel));
toteff = (a*b)/((a*b)+(te-(a*b)));
```

```
ratio = (a*b)/((te-(a*b)));
Test=1;
format test testf.; run;

******************************;
* Goodman test *;
******************************;
data stest3; set stesto;
zvgi = (a*b)/sqrt(((b*b)*(sa*sa))+((a*a)*(sb*sb))+((sa*sa)*(sb*sb)));
absgood = abs(zvgi);
pgi = 2*(1-CDF('NORMAL',absgood));
Test=2;
format test        testf.; run;

******************************;
* Aroian test *;
******************************;
data stest4; set stesto;
zvgii = (a*b)/sqrt(((b*b)*(sa*sa))+((a*a)*(sb*sb))-((sa*sa)*(sb*sb)));
absgood2 = abs(zvgii);
pgii = 2*(1-CDF('NORMAL',absgood2));
Test=3;
format test testf.; run;

data stest5 (drop = _type_ _freq_ test);
Merge stest2 stest3 stest4; run;

ods rtf; ods listing close;
proc print split = '*' data=stest5 ;
  var a sa b sb te se  toteff ratio abssobel ps absgood pgi absgood2 pgii;
label
   a  ='a: fear(X)Coeff*Model M=x'
   sa = 'sa: Standard error (X)* Model M=x'
    b = 'b: isolation (M) Coeff* Model  Y=X M'
    sb = 'sb: SE (M)* Model y = X M'
    te = 'Total Effect'
    se = ' Standard error* total effect'
    toteff = ' ratio indirect effect over * total effect'
    ratio  = ' ratio indirect effect over * direct effect'
    abssobel = ' Sobel test'
    ps     = ' P value * Sobel test'
    absgood        = ' Goodman test'
    pgi  = 'P value * Goodman test'
    absgood2 = ' Aroian test'
    pgii = ' P value * Aroian test' ;
    Title ' Indriect tests'; run;
ods rtf close; ods listing;  quit; run;

*** Bootstarpping method  ***;
%macro regcf  (d,i,t);
```

```
proc reg data=outboot outest=fst3 noprint;
 by  replicate;
      model &d = &i / stb pcorr2 scorr2;
      title ' Regression model/ ' &t;

%mend regcf;
 %regcf (ssq,fear isolat, fear and isolation on social support);
run; quit;

%macro regcs  (d,i,t);
proc reg data=outboot outest=sst3 noprint;
 by  replicate;
      model &d = &i / stb pcorr2 scorr2;
      title ' Regression model/ ' &t;
        %mend regcs;
%regcs (isolat,fear , social support  m=x );
run; quit;

data fst3 (drop= fear isolat ssq);   set fst3;
ffearf = fear;
fisolat = isolat; run;

data sst3 (drop= fear isolat);
  set sst3;
sfearf = fear;  run;

data ast3;    merge fst3 sst3; run;

data st3 (drop =  ssq );    set ast3;
inef = sfearf*fisolat;
teff = ffearf + inef;run;

proc means data= st3 maxdec=2 noprint ;
     var ffearf fisolat inef teff;
     output out=st3out      mean = mfearf misolat minef mteff
                    stderr=sfearf sisolat sinef steff
                    n = nfearf nisolat ninef nteff
                    lclm = lfearf lisolat linef lteff
                    uclm = ufearf uisolat uinef uteff;   run;
data st3out (drop = _type_ _freq_);     set st3out; run;

data all ;
   merge st st3out ;
bfearf = fearfo - mfearf;
bisolat = isolato - misolat;
binef = inefo - minef; Run;

*** Method I to calculate % cI for effects  ***;
lfearf = mfearf - (tinv (.95, nfearf-1)*sfearf);
ufearf = mfearf + (tinv (.95, nfearf-1)*sfearf);
lisolat = misolat - (tinv (.95, nisolat-1)*sisolat);
```

```
uisolat = misolat + (tinv (.95, nisolat-1)*sisolat);
linef = minef - (tinv (.95, ninef-1)*sinef);
uinef = minef + (tinv (.95, ninef-1)*sinef);
lteff = mteff  - (tinv (.95, nteff -1)*steff );
uteff = mteff  + (tinv (.95, nteff -1)*steff);run;

ods rtf; ods listing close;
proc print  split = '*' data = all ;
   var  mfearf sfearf lfearf ufearf
        minef sinef linef uinef
        mteff steff lteff uteff;
label
       mfearf = 'Means*fear*Bootstrap*direct effect'

       sfearf = 'Standard error*fear'
       lfearf = 'lower CI*fear'
       ufearf = 'upper CI* fear'

       minef = 'Means*indirect effect*Bootstrap'
       sinef = 'Standard error*indirect effect'
       linef = 'lower CI*indirect effect'
       uinef = 'upper CI*indirect effect'

       mteff = 'Means*total effect*Bootstrap'
       steff = 'Standard error*total effect*Bootstrap'
       lteff = 'lower CI*total effect'
       uteff = 'upper CI*total effect';
      title 'printing result  ';
       title3 ' 95% CI / Method I  ';  run;
ods rtf close; ods listing; quit;  run;
*** MEthod II to calculate % cI for effects  ***;
%LET ALPHA=.05;
%let a1 = %sysevalf (&alpha/2*100);
%let a2 = %sysevalf ((1-&alpha/2)*100);

proc univariate data =st3 alpha=.05;
      var ffearf fisolat inef teff;

output out=pmeth mean = mfearf misolat minef mteff
                 pctlpts =&a1 &a2
                 pctlpre = lfearf lisolat linef lteff
                 pctlname= p025 p975;  run;
            title 'univariate  result  ';
            title3 ' 95% CI / Method  II '; run;

data allu ;     merge st pmeth ;
lfearfu = lfearfp025;
lisolatu = lisolatp025;
linefu  = linefp025;
lteffu  = lteffp025;
ufearfu = lfearfp975;
```

```
uisolatu = lisolatp975;
uinefu  = linefp975;
uteffu  = lteffp975;

bfearfu = fearfo - mfearf;
bisolatu = isolato - misolat;
binefu = inefo - minef;
bteffu = teffo - mteff;;  run;
ods rtf; ods listing close;

proc print  split = '*' data = allu ;
   var fearfo mfearf    lfearfu ufearfu        ;
   var    isolato misolat   lisolatu uisolatu;
    var inefo minef   linefu uinefu ;
    var teffo mteff   lteffu uteffu;
label
      fearfo = 'Direct effect*fear*Original'
       mfearf = 'Direct effect *Means*fear*Bootstrap'
      lfearfu = 'direct effect*lower CI*fear'
      ufearfu = 'direct effect *upper CI* fear'
      bfearfu = 'Direct effect *Bias*fear'
      isolato = 'isolat*Orginal'
      misolat = 'Means*isolat*Bootstrap'
      lisolatu = 'lower CI*isolat'
      uisolatu = 'upper CI*isolat'
      bisolatu = 'Bias*isolat'
      inefo = 'indirect effect*Original'
      minef = 'Means*indirect effect*Bootstrap'
      linefu = 'lower CI*indirect effect'
      uinefu = 'upper CI*indirect effect'
      binefu = 'Bias*indirect effect'
      teffo = 'total effect*Orginal'
      mteff = 'Means*total effect*Bootstrap'
      lteffu = 'lower CI*total effect'
      uteffu = 'upper CI*total effect '
      bteffu = 'Bias*total effect';
      title3 ' 95% CI / Method  II ';run;
ods rtf close; ods listing;  quit; run;
```