Korean Journal of Health Education and Promotion
Vol.3, No.1, 2001, pp. 129-142

Meta-Analysis of AIDS Prevention Programs

Yeon-Hee Kim
Dept. of Health Administration, Nam Seoul University, Cheonan, Korea

CONTENTS

Ⅰ. Introduction

Meta-analysis studies are becoming a central part of research integration in almost all behavioral and medical sciences. For example, Lipsey and Wilson (1993) have recently identified 302 specific meta-analysis studies dealing with the efficacy of a wide array of behavioral interventions. Capitalizing on these advances in research integration and their potential to improve future reviews of empirical research findings, many behavioral science and educational research methods texts (see for example Cooper, 1990; Gall, Borg and Gall, 1996; Hittleman and Simon, 1997) now routinely include a treatment of meta-analysis which is within the grasp of both researchers and practitioners.

Meta-analysis provides a statistical method for combining the quantitative findings from several studies that address the same research problem. McGaw's (1988) explanation suggests that meta-analysis is a strategy for making the integration of the findings of empirical research itself an empirical task rather than an intuitive one. Thus, research integration becomes an "analysis of analysis" or "meta-analysis" in a term coined by Glass(1976).

Since it is a specific method of research integration, a meta-analysis study follows the same basic procedures used research reviews. These procedures are as follows: (1) describing problems to be addressed in the review, (2) collection all relevant research studies, (3) specifying pertinent data in each study, (4) conducting the data analysis and interpretation tasks, and (5) sharing conclusions in a published report.

This study examined the effectiveness of instructional programs designed to improve school-age children's knowledge and understanding of Acquired Immune Deficiency Syndrome (AIDS). The proper way to change or to modify AIDS instructional program is to base such changes on research that addresses content analysis and effectiveness of AIDS prevention programs. Moreover, the most pressing concern among the many issues surrounding AIDS prevention education in school is program effectiveness. Therefore, well-designed summative and formative evaluation studies are required in order to provide information regarding effective AIDS prevention programs. With these research results, more appropriate and effective AIDS prevention programs can be developed and implemented in schools.

Several studies have been conducted and indicated that some AIDS instructional programs were not effective (Slonim-Nevo et al. 1991; Steitz and Munn, 1993). Nevertheless, some of AIDS prevention programs have produced valuable and effective results (Crawford et al. 1990; Lawrence et al. 1995; Sunwoo et al. 1995).

For this study, a meta-analysis is to be conducted on all studies of the effectiveness of AIDS instructional programs published between 1985 and 1997. This time period was selected primarily because few (if any) eligible studies were conducted prior to 1985.

II. Research Hypothesis

The research hypothesis for this meta-analysis study is as follows: Health education programs designed for classroom instruction are an effective means to increase school-age children's knowledge about AIDS.

Problem formulation is completed five specific characteristics that are used to identify the individual studies to be included in the research review. The first four characteristics are based on information given in the common research hypothesis. The fifth characteristic creates a time period. Each characteristic is described below.

The independent variable (treatment) for the meta-analysis study is the AIDS instructional program. The dependent variable (outcome) for the meta-analysis is a criterion measure indicating a student's knowledge about AIDS.

Accordingly, each individual study must use a valid and reliable AIDS knowledge questionnaire. Since a standardized effect size is determined for each individual study, meta-analysis does not require that the same questionnaire be used in all studies to be included in the review.

The target population for the meta-analysis study is school-age children in the United States. Thus, the sample used in each individual study in the review must included only elementary or secondary school students. Following this specification allows the research reviewer to generalize the meta-analysis findings to the population consisting of all school-age children in the United States.

The research design for the meta-analysis study requires a valid test statistics that indicates how the AIDS instructional program has increased a student's knowledge about AIDS. Thus, each study to be included in the research review must have an experimental design that (a) includes both before- instruction and after-instruction measures and (b) a valid statistical test procedure to determine if students participating in the instructional program made significant (meaningful) gains in their knowledge about AIDS.

The specific research designs used to evaluate the program effectiveness for each study to be included in this meta analysis are as following:

Design A.
One-Group Pretest-Posttest Design

This design involves three steps:(1) administration of a before-instruction (pretest) measure to assess knowledge of AIDS, (2) implementation of the AIDS instructional program, and (3) administration of an after- instruction (posttest) measure to assess knowledge about AIDS again. The effects of the AIDS instructional program are determined by comparing the pretest and posttest scores.

Design B.
Posttest-Only Control-Group Design

This design involves three steps: (1) Random assignment of participants or classes to either the instructional program (experimental group) or a control group not receiving this program, (2) Administer the AIDS instructional program to the experimental group only, and (3) administer the after-instructional (posttest) measure to both groups. The effect of the AIDS instructional program is determined by comparing the posttest scores for the two groups.

Design C.
Pretest-Posttest Control-Group Design

This design involves four steps: (1) Random assignment of participants or classes to either the instructional program(experimental group) or a control group not receiving this program, (2) Administration of a before-instruction (pretest)measure to both groups assess knowledge about administration of an after-instruction (posttest)measure to both groups to assess knowledge about AIDS again.

Design C is often the preferred instruction evaluation strategy for three reasons. First, it allows researchers to control interaction between the pretest and the experimental treatment (AIDS instruction program). Specifically, this design helps to eliminate significant effects due to pretest only. Second, when there is no evidence of pretest differences, the data analysis can focus on posttest scores only in the manner indicated for Design B. This data analysis procedure was used for all Design C studies in this meta-analysis. Third, after the experiment is completed, participants in the control group can be given the same experimental treatment. Most important, this procedure which is widely used in health education research helps to maintain equity because it allows all participants to benefit from the experimental treatment. Moreover, parents and guardians are far more likely to sign the required consent forms when they are certain their children will receive the treatment.

Five research studies used Design A. This is a basic within design subjects design that implies that all students provide both a before-instruction (pretest) and an after-instruction (posttest) measure of their knowledge of AIDS. The effect of the AIDS instructional program is determined by comparing pretest and posttest scores. the other five research studies in the meta-analysis sample used Design C. This is a basic research design that uses both an experimental group that receives the AIDS instructional program and a control group that does not receive the program.

Design B was not used in any of the research studies presented in this meta-analysis. However, it was used in the data analysis for all Design C studies. This was done for the following reasons.

First, there was no evidence of pretest differences in the experimental and control groups in the five Design C studies. Thus, each of the five studies began by using two equivalent groups.

Second, since the two groups were equivalent, administration of the AIDS knowledge questionnaire after the instructional program provides two distributions of scores that can be compared to indicate the effect of the program.

Ⅲ. The Effect Sizes

The effect size findings for the ten studies in this meta analysis are summarized in <Table 1>.

<Table 1> The sample sizes and effect sizes for each study

Study	Sample Size	Effect Size
1	482	1.40
2	1550	1.21
3	772	0.98
4	82	0.88
5	34	0.79
6	151	0.36
7	131	0.25
8	770	0.12
9	34	-0.07
10	43	-0.11

Table 1 shows ten studies with sample size and effect size. Inspection of table 1 indicates that the sample sizes used to estimate the effect sizes for the ten meta-analysis studies were very different. The largest sample size is *N* equal to 1,550 in Study Two. The next largest sample size is *N* equal to 772 in Study Three.

The two smallest sample sizes are for Study Five and Study Nine. In each case the *N* value is equal to thirty-four. The range of the sample sizes used to estimate the ten study effect sizes is 1,550 minus 34 which equals 1,516.

Sample size is an important factor in meta-analysis since it is used first to estimate the individual effect size *d* for each study and then to estimate the corresponding confidence interval for each *d*.

A review of the effect size concept is given in Note 1. Those unfamiliar with the effect size concept are encouraged to examine this review.

Three key pieces of information about the effect size variation deserve mention here.

First the ten effect sizes are ranked from the highest d value equal to 1.4 in Study One to the lowest *d* value equal to -0.11 in Study Ten. Thus, the overall range for the ten effect sizes is 1.4 minus (-0.11) which equals 1.51.

Second, there is a positive correlation between effect size magnitude and program effectiveness. Thus, rank ordering the effect sizes from highest to lowest also ranks the ten studies from most effective (Study One) to least effective (Study ten).

Third, the sign of an effect size is also an important indicator. Specifically, when the sign of an effect size is positive, there is a gain in knowledge about AIDS that can be attributed to the instructional program. Accordingly, eight out of the ten studies in the meta-analysis have documented knowledge gains. These gains range from a high *d* value of 1.4 in Study One to a low *d* value of 0.12 in Study Eight.

Similarly, when the sign of an effect size is negative, there is no gain in knowledge that can be attributed to the instructional program. This was the case in Study Nine having a d value equal to -0.07 and in Study Ten having *d* value of -0.11 (see Note 2). Negative *d* values indicate that students on average scored lower in the posttest. These lower scores could be due to one or more factors such as retention (a failure to remember correct facts), ineffective instruction (teaching that leads to confusion rather than clarity), chance scores (guessing more items correct in the pretest), and student attitude (not taking the posttest seriously)

Ⅳ. Confidence intervals

Data provided in an experimental study that uses probability sampling yield both an effect size estimate of the true population effect size and an estimate of the variance due to sampling. These two pieces of information are used to get a 95 % confidence interval for the true value of the population effect size.

These concepts are best explained using an example from Figure 1.

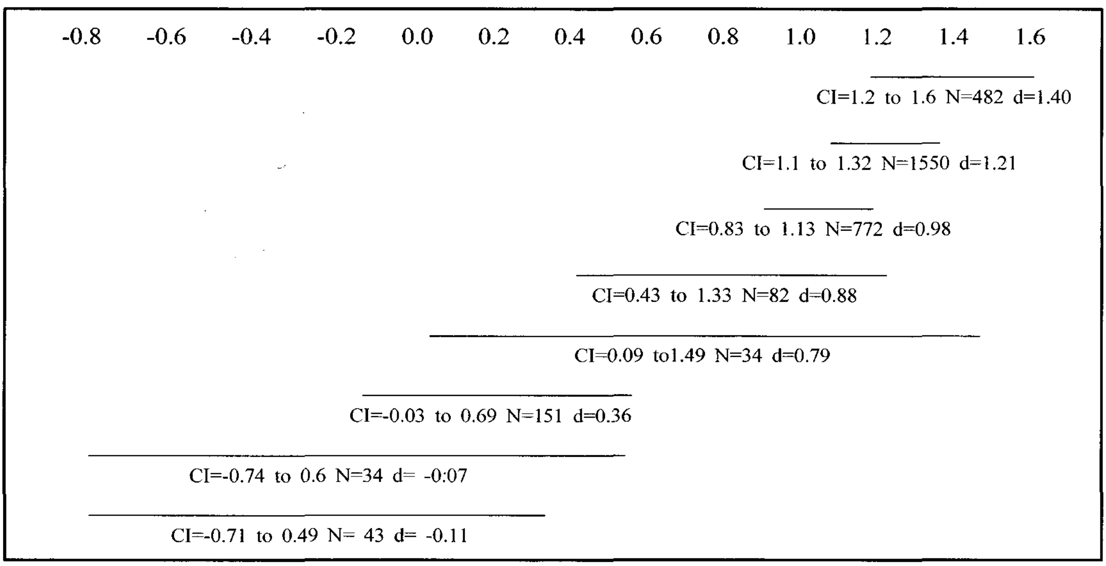
For Study One the effect size estimate is *d* equal to 1.4. This single value, called a point estimate, is the best estimate of the population effect size that would result if this AIDS instructional program were offered to all members of the target population.

Using the estimate of the variance due to sampling in this study, the margin of error for a 95 % confidence interval for the population effect size is 0.20. Thus, the line segment for Study One in Figure 1 has a lower bound of 1.4 minus 0.20 or 1.2 and an upper bound of 1.4 minus 0.20 or 1.6 (see Note 3).

The formal interpretation of this 95 % confidence interval is as follows: Given the sample data for Study One, we are 95 % certain that the population effect size for this AIDS instructional program is between 1.2 and 1.6.

Each of the other line segments in Figure 1 can be interpreted in a similar manner. For example, the Figure 1 information for the second most effective AIDS instructional program has the following interpretation: Given the sample data for Study Two, we are 95 percent certain that the population effect size is between 1.10 and 1.32.

Two important points emerge from a careful study of the effect size information presented Figure 1.



Legend

N is the total size for the study

CI is the range of the confidence interval estimate of an actual effect size.

d is the actual effect size. Note that a positive value favors the outcome of the group receiving AIDS instruction

<Figure 1> Effect Size Scale

Influence of Sample Size. Notice that studies with large sample sizes have smaller line segments depicting their confidence intervals. For example, compare the line segment for Study Two where N equals 1,550 to the line segment for Study Five where N equals 34.

Population Effect Size. It is very unlikely that these ten studies have a common population effect size. This is true because the ten studies in Figure 1 do not have a significant overlap in their line segments.

The obvious lack of overlap can be seen by comparing the line segment for Study One with each of the line segments for Study Six to Study Ten. Since there is no overlap of the Study One line segment with the line segments for any of these five least effective studies. It is relatively safe to suggest that the population effect for Study One is very different from any reasonable common population effect size one might estimate for the five studies used as a comparison group.

V. Estimating the Combined Effect size

One goal of meta-analysis is to combine estimates of effect size to produce an overall average.

The statistically optimal way to average a group of independent estimates is to form a weighted average. The weight for each effect size estimate is the reciprocal of its sampling variance (see Note 4).

The combined estimate D for this meta-analysis is given below:

Step One: Estimation model that assumes effect sizes are homogeneous.

D=0.87 which is a combined estimate using information provided in the ten individual meta-analysis studies elaborated in Figure 1.

R=($0.81 \leq D \leq 0.94$) which is a 95 % confidence interval for the combined estimate D.

Step Two: Estimation model that assumes effect sizes are not homogeneous.

Results for Subgroup One are for the five individual studies in the first homogeneous group

D(1)=1.16 is the combined estimate for all group one studies.

R(1)=($1.08 \leq D \leq 1.24$) is the corresponding 95% confidence interval

Results for Subgroup Two are for the five individual studies in the second group.

D(2)=0.15 is the combined estimate for all group two studies.

R(2)=($0.13 \leq D_2 \leq 0.17$) is corresponding 95 % confidence interval.

The sign of combined effect size estimate is positive; therefore, AIDS instructional programs impact positively on knowledge about AIDS.

The combined estimate D equal to 0.87 is based on the assumption that all studies come from a single population having just one effect size parameter. This assumption is called homogeneity assumption. When it is true the observed variation among the individual effect size estimates is attributed to random sampling fluctuations (sampling error).

Assuming homogeneity, the 95 % confidence interval for the combined estimate 0.87 has the following interpretation: Given this meta-analysis study sample, we are 95 % certain that the population effect size value (parameter) is at least 0.81 and no larger than 0.94.

Inspection of the two combined effect size estimates indicates that there are at least two meaningful subgroups. Specifically, the combined effect size estimate for group one is D(1) equal to 1.16 and the corresponding estimates for group two is D(2) equal to 0.15. The mean difference for these two effect size parameter estimates is 1.16 minus 0.15 which equals 1.01. This estimated difference is very large and clearly significant.

Next, notice the two 95 percent confidence intervals. The lower bound for the 95 percent confidence interval for the 95 % confidence interval for group one is 1.08. The upper bound for the 95 percent confidence interval for group two is 0.17. There is no overlap in these two confidence intervals. Moreover, the distance between the lower bound for the group one interval estimate (1.08) and the upper bound for the group two interval estimate (0.17) is equal to 0.91. This difference is also very large and clearly significant.

Taken collectively, this information for the two subsamples (independent groups) leaves little doubt that at least one moderator variable must exit to explain the difference between the two parameter estimates D(1) and D(2).

The statistical test for homogeneity

The final statistics is used to conduct a homogeneity test. If this statistical test suggests that homogeneity is a fair assumption, the statistical analysis is completed. If the statistical test fails to support a homogeneity assumption, a research must be undertaken to locate other variables to account for the systematic fluctuations (nonrandom variance) in the observed effect sizes.

The graphics information in Figure 1 has already pointed toward the possibility that the homogeneity assumption for the ten effect size estimates in this meta-analysis illustration is unrealistic.

This conclusion about homogeneity was reached using the idea that homogeneity requires some overlap in the line segments representing 95 percent confidence intervals for each d value.

The formal statistical test for homogeneity is carried out using a chi-squared goodness-of-fit test shown below:

$$H(T)= \sum w_i d_i^{\,2} - \frac{(\sum w_i d_i)^2}{\sum w_i} = 206.14$$

The H(T) test statistics value 206.14 exceeded the 95 percent critical value 16.92 of the chi-square distribution with 9 degrees of freedom. Because 206.14 exceeds the critical value, the hypothesis of homogeneity test that all ten studies have the same underlying population effect size was rejected. Therefore, these ten studies did not provide evidence of program effects of the same magnitude.

Applying this chi-squared goodness-of-fit test to the data from this meta-analysis study sample suggests the fit implied by the homogeneity assumption is not appropriate for these ten effect size estimates are heterogeneous (i.e., they have variation in excess of sampling fluctuations), and the meta analyst would be well advised to examine these studies to determine, if possible, why they are deviant.

Ⅵ. Conclusion

The purpose of this study was to analyze and synthesize findings on selected studies related to the effectiveness of instructional programs designed to improve school-age children's knowledge and understanding of Acquired Immune Deficiency Syndrome (AIDS). To achieve this purpose, meta-analysis was applied to analyze and synthesize findings from prior studies.

The first step was used to summarize what was learned from reading each research study in the meta-analysis sample. Summary information was illustrated with the target populations, the AIDS instructional programs, the specific instruments used to measure knowledge about AIDS, the experimental research designs used to assess program effectiveness, and the actual effect size estimates for each study.

The next step was used to apply meta-analysis statistical methods for obtaining a combined effect size parameter estimate, conducting a homogeneity test.

The combined estimate D equal to 0.87 is based on the assumption that all studies come from a single population having just one effect size parameter. However, inspection of the two combined effect size estimates indicates that there are at least two meaningful subgroups. Specifically, the combined effect size estimate for group one is D(1) equal to 1.16 and the corresponding estimates for group two is D(2) equal to 0.15. The mean difference for these two effect size parameter estimates is 1.01. This estimated difference is very large and clearly significant.

The combined effect size for dependent

variable, knowledge about AIDS, was positive, indicating that school-based AIDS instructional program for school age children consistently contribute to increasing a student's knowledge about AIDS.

The final analysis on program characteristics for knowledge about AIDS identified two potential moderator variables to be explored in future research. These potential moderator variables were hours of instruction, parent involvement, program type, and comprehensive concepts. This means that these program potential variables may cause the variation of effects of AIDS instructional programs for the school age children regarding knowledge about AIDS.

The conclusion about homogeneity was reached using the idea that homogeneity requires some overlap in the line segments representing 95 percent confidence intervals for each d value.

The recommendation for future research study and practice is as following:

Policy makers, practitioners, and other consumers of meta-analysis findings were advised to verify that a published meta-analysis of interest is based on a test of the homogeneity assumption. Failure to conduct this test, and failure to perform the required additional analyses when this test suggests effect size estimates are heterogeneous, can easily result in reporting inaccurate and incomplete meta-analysis findings and recommendations.

References

Cooper, H. 1990. Meta-Analysis and the Integrative Research Review in *Research Methods in Personality and Social Psychology* edited by C. Hendrick and M. S. Clark. Newbury Park, CA:Sage.

Crawford, I. et al. 1990. Amultimedia-based approach to increasing communication and the level of AIDS knowledge within families. *Journal of Community Psychology*. 18:361-373

Di Clemente, R,. Boyer, C., and Morales, E. 1988. Minorities and AIDS: Knowledge, Attitudes, and Misconceptions among Black and Latino Adolescents. *American Journal of Public Health*. 78:55-57.

Gall, M. D., Borg, W. R., and Gall, J. P. 1996. *Educational Research*(6th ed). New York. NY: Longman.

Glass, G. V. 1976. Primary, Secondary and Meta-Analysis. *Educational Researcher*. 5(1): 3-8

Hedges, L. V., and Becker, B. J. 1986. Statistical Methods in the Meta-Analysis of Research on Gender Differences in *The Psychology of Gender: Advances through Meta-Analysis* edited by J. S. Hyde and M. C. Linn. Baltimore. MD: Johns Hopkins University Press.

Hedges, L. V., and Olkin, I. 1985. *Statistical Methods for Meta-Analysis*. New York, NY: Academic Press.

Hedges, L. V., Shymansky, J. A., and Woodworth, G. 1989. *Modern Methods of Meta-Analysis: A Practical Guide.* Washington. DC: National Science Teacher Association.

Hittleman, D. R., and Simon, A. J. 1997. *Interpreting Educational Research: An Introduction for Consumers of Research.* Columbus, OH:Merrill

Lipsey, M. W., and Willson, D. B. 1993. The Efficiency of Psychological Educational and Behavioral Treatment: Confirmation from Meta-Analysis. *American Psychologist.* 48(12):1181-1209.

Lawrence, J. S. 1995. Comparison of education versus behavioral skills training interventions in lowering sexual HIV-risk behavior of substance-dependent adolescents. *Journal of Consulting and Clinical Psychology.* 63(1): 154-157

McGaw, B. 1988. Meta-Analysis in *Educational Research Methodology and Measurement: An International Handbook* edited by In J. P. Keeves. New York. NY: Pergamon press.

McNamara, J. F. 1997. Meta-Analysis: A Strategy to Improve Integrative Research Reviews. *International Journal of Educational Reform.* 6(4):490-495.

McNamara, J. F. 1994. *Surveys and Experiments in Education Research.* Lancaster, PA: Technomic Publishing Co., Inc

Slonim-Nevo, V., Ozawa, M., and Auslander, W. 1991. Knowledge, attitudes and behaviors related to AIDS among youth in residential centers. *Journal of Adolescence.* 14: 17-33

Steitz, J. A., and Munn, J. A. 1993. Adolescents and AIDS: Knowledge and Attitudes. *Adolescence.* 28(111):609-619.

Sunwoo, J. et al.1995. School-based AIDS education for adolescents. *Journal of Adolescent Health.* 16: 309-315

Notes

Note1. An example introduced in McNamara (1994) is used in this note.

Assume a school district uses a two independent sample design to compare achievement for a new method (Method1) and an existing method(Method2) used to teach Algebra One. Let the sample size in each of the two student groups be equal to 100.

The two methods are used for an entire school year. Achievement scores at the end of the year for the 100 students in Method One have a mean of ninety. Corresponding scores for the 100students taught using Method Two have a mean of seventy-eight. Thus, the sample mean difference is twelve points.

Let the standard deviation for each of these two sets of scores equal ten. Thus, the common standard deviation (called the pooled standard deviation in the t Test) equals ten.

The effect size (d) is a difference expressed in standardized form. Specifically, the sample mean difference of twelve points on the achievement test is divided by the common standard deviation equal to ten. This yields d equal to 1.2.

Using a sample distribution for the difference of two independent sample means, the effect size d equal to 1.2 has the following interpretations.

Mean Difference Interpretation. Put in strict statistical terms. one can correctly interpret the experimental results in this manner: On average, students taught by Method One had achievement scores that were twelve points higher than the students taught by Method Two.

Standardized Difference Interpretation. Using the standardized difference d equal 1.2, one can make this parallel statement: On average, students taught by Method One had achievement scores that were 1.2 standard deviations higher than the students taught by Method Two.

Policy Interpretation. Sample data in this illustration have direct relevance for making a data-based policy decision about which method should be adopted for future school years.

If only the existing method (Method Two) were used in future school years, the best prediction would be that the end of the year Algebra One score distribution for all students would have a mean score of seventy-eight and a standard deviation of ten.

The new method is obviously more effective for teaching Algebra One. If both methods are within the budget allocated to the Algebra One program, the new method of teaching Algebra One should be the method used to teach all students in future school years.

Note2. All ten effect sizes were calculated using Equation (11) and Equation(12) given below:

$$\mathrm{d} = c_n g = c_n(\overline{Y^F} - \overline{Y^M})/S, \qquad (11)$$

where the values of c_n are given to a very good approximation by

$$c_n = 1 - \frac{3}{4n^M + 4n^F - 9} \qquad (12)$$

Note3. The confidence intervals for all effect sizes were calculated using Equation (13) and Equation(15) in Hedges and Becker.

$$\nu = \frac{n^M + n^F}{n^M n^F} + \frac{d^2}{2(n^M + n^F)} \qquad (13)$$

$$\mathrm{d} - z_\alpha\sqrt{\nu} \leq \delta \leq + z_\alpha\sqrt{\nu} \qquad (15)$$

Note4. The combined estimate D was calculated using Equation (16)and the overall confidence interval was calculated using Equation (19) in Hedges and Becker.

$$d. = \frac{\sum_{i=1}^{k} w_i d_i}{\sum_{i=1}^{k} w_i} \qquad (16)$$

where

$$w_i = 1/\nu = \frac{2(n_i^M + n_i^F)n_i^M n_i^F}{2(n_i^M + n_i^F)^2 + n_i^M n_i^F d_i^2} \quad (17)$$

$$d. - z_\alpha\sqrt{\nu.} < \delta < d. + z_\alpha\sqrt{\nu.}, \quad (19)$$

where z_α is the $100\,\alpha$ percent two-tailed critical value of the standard normal distribution.

ABSTRACT

Meta-analysis methods are becoming a central part of research integration in behavioral and medical science studies. The main goal of the meta-analysis is combining the quantitative findings from several studies which address the same research problem and is sharing conclusions in a published report.

The purpose of this research is to develop meta-analysis approach to evaluate effectiveness of instructional program design to improve school-age children's knowledge and understanding of Acquired Immune Deficiency Syndrome (AIDS). To achieve this purpose, meta-analysis is applied to analyze and synthesize findings for AIDS instructional programs from several prior studies.

Key Words : Meta- Analysis, AIDS (Acquired Immune Deficiency Syndrome)