

Review

Publishing Nutrition Research: A Review of Multivariate Techniques—Part 2: Analysis of Variance

JEFFREY E. HARRIS, DrPH, RD, LDN; PATRICIA M. SHEEAN, PhD, RD; PHILIP M. GLEASON, PhD; BARBARA BRUEMMER, PhD, RD; CAROL BOUSHEY, PhD, MPH, RD

ABSTRACT

This article is the eighth in a series exploring the importance of research design, statistical analysis, and epidemiology in nutrition and dietetics research, and the second in a series focused on multivariate statistical analytical techniques. The purpose of this review is to examine the statistical technique, analysis of variance (ANOVA), from its simplest to multivariate applications. Many dietetics practitioners are familiar with basic ANOVA, but less informed of the multivariate applications such as multiway ANOVA, repeated-measures ANOVA, analysis of covariance, multiple ANOVA, and multiple analysis of covariance. The article addresses all these applications and includes hypothetical and real examples from the field of dietetics.

J Acad Nutr Diet. 2012;112:90-98.

This article is the eighth in a series exploring the importance of research design, statistical analysis, and epidemiology in nutrition and dietetics research. The purpose of this series is to provide dietetics practitioners with tools to enhance general understanding of key concepts inherent in high-quality nutrition research by providing relevant examples and additional resources. These articles are intended for the seasoned researcher

J. E. Harris is a professor and didactic program director, Department of Health, West Chester University, West Chester, PA. P. M. Sheean is an assistant professor of preventive medicine, Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL. P. M. Gleason is a senior fellow, Mathematica Policy Research, Geneva, NY. B. Bruemmer is a senior lecturer emeritus, Graduate Program in Nutritional Sciences, University of Washington, Seattle. C. Boushey is an associate professor and director, Coordinated Program in Dietetics, Purdue University, West Lafayette, IN.

Address correspondence to: Jeffrey E. Harris, DrPH, RD, LDN, Department of Health, 213 Sturzebecker Health Sciences Center, West Chester University, West Chester, PA 19383. E-mail: jharris@ucupa.edu

Manuscript accepted: September 16, 2011.

Copyright © 2012 by the Academy of Nutrition and Dietetics.

2212-2672/\$36.00

doi: 10.1016/j.jada.2011.09.037

as well as the nutrition research novice. In addition, this article is the second in a series of three on multivariate statistical techniques.

In this article we describe the various uses of a statistical technique called analysis of variance (ANOVA), including multivariate applications. The following applications of ANOVA are addressed: one-way ANOVA, repeated-measures ANOVA, multiway ANOVA, multiple analysis of variance (MANOVA), analysis of covariance (ANCOVA), and multiple analysis of covariance (MANCOVA). The descriptions will be enhanced through the use of hypothetical dietetics-related examples. Finally, examples from the *Journal* provide real dietetics-related illustrations of how the techniques have been used in practice. **Figure 1** provides definitions of terms used throughout this article.

WHEN TO USE ANOVA

The *t* test

Suppose 40 people are randomly selected from a large population. Twenty are assigned to take a serum low-density lipoprotein (LDL)-lowering drug and the other 20 a placebo. The mean LDL levels are 90 mg/dL (2.34 mmol/L) and 120 mg/dL (3.12 mmol/L), respectively, following treatment. In this situation an independent *t* test is used to determine whether there is a statistically significant difference in LDL levels between the experimental group and control group. What if a third group of 20 subjects was assigned to a Mediterranean diet and their LDL was 100 mg/dL (2.6 mmol/L)? What statistical test can be used to determine whether there is a statistically significant difference between three groups rather than two? The *t* test is limited to comparing the means of two groups. At its most basic, ANOVA is used to compare the means of two or more groups for statistically significant differences.

The Basics of ANOVA

ANOVA is a series of statistical tests that can be used to compare the means of two or more independent groups. ANOVA determines which factors contribute to the overall variation in a dataset. With LDL values from 100 people, why might the values range from 70 mg/dL (1.82 mmol/L) to 190 mg/dL (4.94 mmol/L)? What specific genetic, lifestyle, and demographic factors might account for the variation? ANOVA allows us to determine what factors can explain the variation.

Figure 1. Definitions of terms used in statistical analysis. Based on Rosner B. *Fundamentals of Biostatistics*. 6th ed. Belmont, CA: Thomson Higher Education; 2006.

Dataset 1		
Group 1	Group 2	Group 3
90	130	180
95	133	181
100	135	176
95	133	185
105	140	181
89	135	180
98	137	183
mean=96	mean=135	mean=181

Within

Between

Figure 2. Illustration of within-group variance vs between-group variance.

The common calculated value for expressing variation is the variance. Because all factors accounting for variation in data can rarely be determined there is explained variance and unexplained variance. ANOVA allows the determination of how much of the variance can be explained and which factors explain a significant amount of the variance. For the above example, does the treatment using diet and drugs significantly explain the variance in means among all subjects?

The *F* Ratio

The test statistic for ANOVA is the *F* ratio rather than the *t* value that is used to compare the means of two groups using the *t* test. The *F* ratio is the explained variance divided by the unexplained in a data set. The explained variance is also called the between-group variance and the unexplained variance is called the within-group variance. Figure 2 presents a dataset that illustrates the concepts of within- and between-group variance. As presented in Figure 2, the between-group variance is that which exists between group means and within-group variance is that which exists within each separate group in a data set. If there is no difference between group means then the between-group and within-group variances are equal, yielding $F=1$. If there are statistically significant differences between groups, then the between-group variance is greater than within, yielding an $F > 1$. This appears to be the case in Figure 2. The variation in group means appears to be significantly greater than the variation of values within the groups.

When the *F* ratio is calculated it is compared to a critical value based on a set level of significance, usually 0.05 (a 5% probability that an observed difference between group means is by chance, and not a real difference). If the *F* ratio exceeds the critical value then there is a statistically significant difference between the means of the groups. These concepts will be operationalized with an example in the section on one-way ANOVA. The *F* ratio is the test statistic for all types of ANOVA.

Although the above description captures the basic idea behind ANOVA, based on the number and type of variables being measured, different versions of ANOVA are used. In different situations, one-way ANOVA, repeated-measures ANOVA, multiway ANOVA, ANCOVA, MANOVA, and MANCOVA may be used. Each of these versions is discussed below.

ONE-WAY ANOVA

One-way ANOVA is used in a situation in which a researcher wishes to examine whether there is a significant relationship between one categorical variable with two or more categories and one quantitative variable. One-way ANOVA is an extension of the *t* test for independent samples. The *t* test for independent samples tests if there is a statistically significant difference between means of two independent samples for a specific quantitative variable. When the means of more than two samples are to be compared one-way ANOVA is applied.

Suppose four equal-sized random samples are chosen from a population; one sample takes 20 mg of an LDL-lowering medication, another eats a Mediterranean-style diet, the third group eats the standard American diet, and the final group eats a vegan diet. One can think of the four different samples reflecting four categories of a single categorical variable capturing the individuals' diet or use of medication (the categorical variable could be called "treatment"). A researcher may want to examine whether this categorical variable is related to a quantitative variable such as an individuals' LDL level. One-way ANOVA is the appropriate statistical test to determine whether the LDL levels of these groups are significantly different. One-way ANOVA is appropriate to determine whether the treatment specifically explains a significant amount of the variance in the LDL level compared to that amount of the variance that is unexplained. As mentioned previously, the test statistic resulting from ANOVA is the *F* ratio. A high *F* ratio indicates the greater likelihood of a statistically significant difference between groups.

The One-Way ANOVA Calculation

The above concepts can be illustrated with an example calculation. The mean LDL levels for the LDL-lowering drug, Mediterranean diet, vegan diet, and American diet groups are 87, 107, 100, and 133 mg/dL (2.26, 2.78, 2.6, and 3.46 mmol/L), respectively. Are these LDL levels statistically significantly different from one another?

By calculating the between-group and within-group variances the *F* ratio can be determined so true statistical differences can be determined. Table 1 presents the calculations of the *F* ratio. In Table 1, the between-group mean square is the between-group variance and the within-groups mean square is the within-group variance. When the former is divided by the latter it yields $F=69$ —a value much greater than 1—indicating that the between-group variance is substantially higher than the within-group variance. Therefore, the variation in treatments explains a significant amount of variance in LDL levels. The *F* ratios that must be exceeded for levels of significance of 0.05 and 0.01 for the statistically significant differences in LDL level based on treatment are 2.87

Table 1. One-way analysis of variance calculations^a for comparing the hypothetical serum low-density lipoprotein cholesterol level means of four treatment groups (ie, lipid-lowering drug, Mediterranean diet, vegan diet, or American diet)

Source of variation	Sum of squares	df	Mean square	F	P value
Between-group variance	11,248	3	3,749	69	<0.001
Within-group variance	1,971	36	55		
Total	13,219	39			

^aRefer to Figure 1 for definitions of terms used in this Table.

and 4.38, respectively. These two values are selected from an *F* ratio table based on the between- and within-group degrees of freedom (Table 1) (1). These values are called critical values, *F* ratios that must be exceeded for confidence that there is either <5% or 1% probability that the differences in means between the groups are by chance, rather than real differences. Sixty-nine significantly exceeds these two values indicating that the mean LDL levels of the four treatment groups are significantly different.

Interpreting the Results

The result of a one-way ANOVA indicates if differences in mean values of the quantitative variable are statistically significant, indicating that there is some relationship between the categorical variable and quantitative variable. However, ANOVA does not indicate the exact way that the groups defined by the categorical variable are different from one another. In the above case, for instance, even with a significant ANOVA *F* ratio, the researcher would not know for sure if mean LDL values for the group taking the LDL-lowering drug differ significantly from the group following the Mediterranean diet. Follow-up tests must be done to determine the groups that are actually different from one another. These tests are called post hoc tests with names such as least square difference, Scheffé test, Tukey-Kramer test, Duncan multiple range test, Fisher exact test, Newman-Keuls test, and Dunnett test. In all variations of ANOVA, post hoc tests must be done to ferret out the specific categories in which there are actual differences. These post hoc tests allow multiple pairwise comparisons of means among the groups being compared. The mean LDL values for the four groups mentioned above are 87 mg/dL (2.26 mmol/L) (LDL-lowering drug), 107 mg/dL (2.78 mmol/L) (Mediterranean diet), 133 mg/dL (3.46 mmol/L) (American diet), and 100 mg/dL (2.6 mmol/L) (vegan diet). A post hoc test allows the determination of which of these means are statistically different from one another. When the Tukey-Kramer test is applied to this situation it reveals statistically significant differences between all the means but the Mediterranean and vegan diets. The LDL-lowering drug is superior to all treatments in producing a lower LDL level with the Mediterranean and vegan diets intermediate, and the American diet least effective. With one-way ANOVA it is important to know not only that the groups are statistically different but also which groups

are different from one another. Investigators would want to know which approach led to the lowest LDL levels. For an example of the application of one-way ANOVA refer to the study by Smith and colleagues (2) of heritage retention, bean intake stage of change, and fiber intake among Hispanic mothers.

REPEATED-MEASURES ANOVA

With one-way ANOVA each subject was measured just once for a given variable. In the LDL example above subjects in each of the four intervention groups were measured only once for serum LDL level. In contrast, repeated-measures ANOVA involves repeated measurement of a variable under different conditions or at different times on the same subjects. Repeated-measures ANOVA is used in a situation in which a group of subjects has the same variable measured several times over a specified time course or after exposure to two or more conditions.

Application to Repeated Measurements Taken Over Time

Suppose a sample of people with high LDL levels participates in an LDL-reduction educational program. LDL levels are measured before and after the program, so each subject has two measurements. Pre-test/post-test differences are of interest to the investigator. The investigator is interested in the effects of the program on LDL levels. The appropriate statistical test in this case is the paired *t* test (also called *t* test for dependent samples).

Perhaps the investigator is interested in looking at the trends in LDL level over time and the sustainability of the effects after the program. Specifically she is interested in monthly changes in LDL level during the 4-month program and for 4 months following the program. Each subject would be measured nine times (including a pre-program measurement) across an 8-month period. In this case the appropriate statistical test is the repeated-measures ANOVA. This type of ANOVA involves a quantitative variable, in this case LDL levels that are measured a repeated number of times on each of the subjects in the sample. In this example, this test would determine whether there are statistically significant differences in LDL levels over time when there are greater than two time-related measurements. Each subject has his or her LDL level measured at a number of time points in the future. The *F* ratio is the test statistic.

If the calculated *F* ratio indicates statistical significance a post hoc test must be done to determine at which points in the time sequence of LDL levels there are statistically significant differences. Suppose the mean values are 130 mg/dL (3.38 mmol/L) pre-program, and 120 mg/dL (3.12 mmol/L), 110 mg/dL (2.86 mmol/L), 100 mg/dL (2.6 mmol/L), 90 mg/dL (2.34 mmol/L), 91 mg/dL (2.36 mmol/L), 95 mg/dL (2.63 mmol/L), 95 mg/dL (2.63 mmol/L), and 92 mg/dL (2.39 mmol/L) at the consecutive 1-month follow-up periods. As in the case of the one-way ANOVA, conducting a post hoc test such as the Tukey-Kramer test will enable the investigator to determine which of the means are significantly different. For example, is there a statistically significant difference between means of 130 mg/dL (3.38 mmol/L) and 120 mg/dL (3.12

Table 2. Taste ratings^a of subjects exposed to three different plant-based burgers illustrating a situation in which repeated measures analysis of variance would be used

Subject	Nut-based burger	Soy-based burger	Grain-based burger
1	7	5	3
2	8	6	2
3	9	6	3

^aRating scale: 1=terrible to 10=fantastic.

mmol/L)? How about between the means of 90 mg/dL (2.34 mmol/L) and 91 mg/dL (2.36 mmol/L)? The post hoc tests indicate which means are statistically different. Maybe the most significant changes occurred during the 4-month program, but stabilized thereafter?

A study by Bourque and colleagues (3) is an outstanding example of the use of repeated-measures ANOVA to examine changes in outcome variables over time. They examined the effects of 12 weeks of weight-bearing and non-weight-bearing endurance exercise on iron status in inactive women. They measured outcomes at 2, 4, 8, and 12 weeks.

Application to Subjects Receiving Multiple Treatments

An alternative application of repeated-measures ANOVA involves situations in which each subject is exposed to more than one experimental condition and is measured on a quantitative variable after exposure to each condition. For example, each subject in a sample is asked to rate the taste of three types of plant-based burgers on a scale from 1 to 10, 1 being terrible and 10 being fantastic. Each subject eats one type of burger and rates it, then the next, and finally the third. Mean ratings are calculated for each burger. Table 2 illustrates the type of data that would be collected in this situation. Because each subject is measured multiple times, in this case three times, it is a repeated-measures situation. Mean taste ratings are compared for the three burgers and a repeated-measures ANOVA is used for statistical analysis. If the *F* ratio is statistically significant then a post hoc test would be done to determine which burger had the best rating.

An example of repeated-measures ANOVA being applied to a situation where each participant is exposed to multiple conditions and measured on a quantitative variable with each exposure is the study by Hertzler and Clancy (4), which aimed to determine whether kefir (a fermented milk drink) improves lactose tolerance in adults with lactose intolerance. Each subject in the study was exposed to four different potential lactose-containing foods (including kefir) and measured on indicators of lactose intolerance.

MULTIWAY ANOVA

In a previous example presented here, one-way ANOVA was used to determine whether there was a statistically significant difference in LDL levels between four different interventions, the categorical variable. The four interven-

Table 3. Changes in serum 25-hydroxycholecalciferol level related to 5,000 IU vitamin D-3/d supplementation vs a placebo and 30 min/d sunlight exposure or not during 1 mo; multiway analysis of variance is an appropriate choice to analyze these data

Condition	Placebo	Vitamin D-3 supplementation
	←————— ng/mL —————→	
No sunlight	−5	15
Sun exposure	20	65

tions were LDL-lowering medication, Mediterranean-style diet, standard American diet, and vegan diet. The quantitative variable was LDL levels. Suppose there is interest in not just looking at the effect of treatment approaches on LDL level, but also sex. Each group could be composed of an equal number of men and women. Is there a different effect among men and women? Which statistical test would be used in this situation because clearly one-way ANOVA cannot be used? In situations in which the investigator desires to examine relationships between more than one categorical variable (such as sex and treatment) and one quantitative variable (LDL level) multiway ANOVA (sometimes referred to as multifactor or factorial ANOVA) is used. It is apparent that more than one factor can affect a parameter such as LDL level. Multiway ANOVA is a tool used to determine whether there is a relationship between more than one independent categorical variable and one dependent quantitative variable.

Suppose there is a study with subjects randomized to four groups. One group gets 30 minutes of sunlight per day and takes a placebo. A second avoids sunlight and takes a placebo. A third takes 5,000 IU vitamin D-3 and avoids sunlight. A fourth takes 5,000 IU vitamin D-3 and gets 30 minutes of sunlight per day. Each group has serum levels of 25-hydroxycholecalciferol (the main laboratory measurement to determine vitamin D-3 status) measured before and after a month long intervention. Change in serum 25-hydroxycholecalciferol is the outcome variable. In this case there are two categorical independent variables, sunlight exposure and vitamin D-3 supplementation. Each has two categories, exposure or no exposure, and there is one quantitative outcome variable, change in serum 25-hydroxycholecalciferol. In this case the data are analyzed using multiway ANOVA.

The mean changes in 25-hydroxycholecalciferol for each group can be found in Table 3. The multiway ANOVA can isolate the influence of each categorical variable, commonly referred to as main effects. Main effects are the independent effects of each categorical variable on the quantitative variable. In this case it would be the independent effects of supplements and sun exposure. Each main effect has an associated *F* ratio that enables the investigator to determine the magnitude of each effect.

Also, multiway ANOVA can be used to look at the combined effects of the categorical variables, also known as interactions. An interaction is present when the effect of one categorical variable on a dependent quantitative variable is dependent on one or more categorical variables.

Table 4. An example of a multiway analysis of variance table related to examining the relationship between vitamin D-3 supplementation and exposure to sunlight, and differences in serum 25-hydroxycholecalciferol levels after treatment^a

Source of Variation	Sum of squares	df	Mean square	F	P value
Supplement	5,281.25	1	5,281.25	281.67	<0.0005
Sunlight	7,031.25	1	7,031.25	375.00	<0.0005
Supplement × sunlight ^b	781.25	1	781.25	41.67	<0.0005
Error	300.00	16	18.75		
Total	13,393.75	19			

^aRefer to Figure 1 for definitions of terms used in this Table.

^bInteraction between the supplement and sunlight variables.

Sunlight and D-3 Supplementation and Serum 25-OH D-3

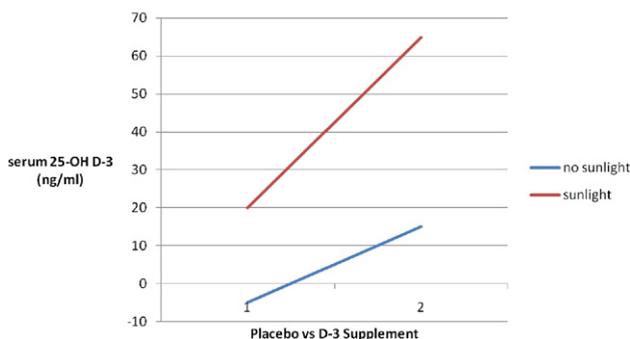


Figure 3. Graphic representation of the interaction between serum 25-hydroxycholecalciferol (25-OH D3) and sunlight exposure and vitamin D-3 supplementation for hypothetical multiway analysis of variance example.

For example, an interaction would be present if the combined effect of aerobic exercise and diet on LDL levels is greater than each by themselves. In other words, the effect of one factor is reliant on the other factor and the effects of the factors are not independent. Interactions have a separate *F* ratio. As noted previously, the magnitude of the *F* ratio is an indicator of the statistical significance of the interaction.

In the study discussed here, the investigators desired to examine the main effects of vitamin D-3 supplementation and sunlight exposure, and the interaction. So there would be *F* ratios calculated for the main effects and the interactions (see Table 4 for an example of an ANOVA statistical table). In the vitamin D-3 example, it appears from the means (Table 3) that there is a disproportionate effect of the combined treatments compared to any alone; hence, an interaction (Figure 3). So the answer to, “What is the effect of vitamin D-3 supplementation,” the answer is it depends on whether the person gets sunlight or not. In other words, the effect of the sunlight is dependent on whether or not subject received supplements. In this case there is a synergistic effect of sunlight and supplementation, not just main effects. If there was no interaction,

supplementation and sunlight exposure could have independent effects, but no joint effects.

So with multiway ANOVA it is important to look for statistically significant *F* ratios for the main effects and interactions. If the *F* ratio for the interaction is significant then it is irrelevant to report any main effects, because the variables are not independent and so cannot have independent main effects.

A study by Mendoza and colleagues (5) examining the relationship between the independent variables study year and socioeconomic status, and the dependent variable reduction of energy density in Texas middle school students illustrates the use of multiway ANOVA. The data analysis was done after collecting data on changes in energy density in the diets of the middle school students after changes in public school policy.

ANCOVA

ANCOVA, an extension of ANOVA, is a tool that explores the relationship between one or more categorical variables and a quantitative dependent variable while adjusting the relationship for one or more quantitative or categorical variables. These variables for which adjustment is made are called covariates. A covariate is a variable that is possibly related to an outcome of interest that can be a primary variable of interest or a confounding one. For instance, if the effect of caffeine on blood pressure is being explored, body weight is a variable that would need to be controlled. For a given dosage caffeine has a greater effect on those with lower vs higher body weight. Therefore, body weight, which can confound the potential relationship between caffeine and blood pressure, is a covariate that must be controlled in this study of caffeine. With ANCOVA, the covariate is often a confounding variable that is controlled by the use of the ANCOVA procedure. ANCOVA is used primarily in two contexts. The first is when participants in a study differ in pre-test measurements and the differences in these baseline covariates could affect the post-test measurements. By utilizing ANCOVA investigators may adjust the post-test measurements for differences in pre-test ones. Second, ANCOVA is used when one or more quantitative or categorical variables are related to an outcome of interest in a study and could affect the influence of the primary categorical variables of interest on the outcome. ANCOVA can adjust the study results for these factors.

Adjusting Post-Test Measurements for Pre-Test Differences

Suppose that at a particular factory an occupational nurse has identified a high prevalence of hypertension among the workers. The nurse wants to know whether it might be effective to offer a behavior-oriented lifestyle program to attempt to combat hypertension. He contracts with a registered dietitian (RD) to design a study to determine the feasibility and possible effectiveness of initiating this program. The RD draws two small random samples from among the factory workers. One group will receive the hypertension control program and the other an educational program on investments.

Subjects’ diastolic blood pressure is measured before and after the program in both groups and the pre-test

Table 5. Mean diastolic heart rates (mm Hg) of experiment and control group members before and after a hypertension control program; these data illustrate a situation in which analysis of covariance would be used

Group	Pretest diastolic heart rate	Post-test diastolic heart rate
	←——— mm Hg ———→	
Hypertension control program	90	70
Control	110	110

mean diastolic blood pressures in the two groups are found to be significantly different. This pre-test difference will potentially confound the post-test results. ANCOVA is used to adjust results for this potentially confounding difference between groups. In this example above, suppose the pre-test mean diastolic pressure is 110 mm Hg in the control group and 90 mm Hg in the experimental group. The post-test mean values are 110 mm Hg in the control group and 70 mm Hg in the experimental group (Table 5). ANCOVA adjusts these post-test blood pressures to neutralize the pre-test differences. When the adjusted post-test values are compared using ANCOVA, an *F* ratio will be generated as the test statistic and can be evaluated for statistical significance. ANCOVA can adjust outcome variables for differences between groups on quantitative variables. If this *F* value is statistically significant, then ANCOVA will have told us that there is a statistically significant difference in posttest blood pressure between the experimental and control groups after adjusting for the potentially confounding effect of the pre-test blood pressure.

A study by Cottone and Byrd-Bredbenner (6) illustrates the use of ANCOVA in the adjustment of post-test variables for pre-test differences in an experimental and control group to determine the effect of the film *Super Size Me* on fast-food knowledge, psychosocial measures, and awareness-raising effectiveness scores among young adults.

Adjusting the Relationship between an Independent and Dependent Variable for Another Confounding Independent Variable

To aid in further clarification of the ANCOVA concept, suppose a dietetics instructor compares two basic nutrition courses on nutrition knowledge. In one course she used lecture-discussion and in another problem-based learning. Style is the independent variable. Nutrition knowledge is the dependent variable. In looking at the grade point averages (GPAs) of students, one class has a mean GPA of 3.25 and the other 2.50. Because of the differences, GPA is the confounding variable. At the end of the course nutrition knowledge scores can be adjusted for the GPA differences. Students with higher GPAs might just naturally have a greater propensity to do better in the course. Therefore, using ANCOVA allows the investigator to make a fairer comparison of the knowledge scores adjusting for GPA.

The study by Ventura and colleagues (7) of Latino children aged 10 to 17 years examining the relationship between metabolic syndrome markers and dietary intake is an excellent illustration of the use of ANCOVA adjusting data for potentially confounding covariates. The investigators compared the dietary intake of those children with more than three or zero metabolic syndrome markers controlling for sex, age, and total energy intake.

MANOVA

With one-way and multiway ANOVA, the relationship between one or more categorical independent variables and one quantitative dependent variable is explored. What tool might be used if an investigator wants to explore more than one outcome/dependent variable? MANOVA is used in the case of looking at the relationship between one or more categorical independent variables and more than one quantitative dependent variable. In certain situations an investigator is interested in examining the effect of categorical variables on a combination of several somewhat correlated and theoretically related dependent variables.

Reasons for Using MANOVA

Why not just run separate one-way or multiway ANOVAs for each of the dependent variables? Doing so multiplies what is called family-wise error. Family-wise error is defined as increasing the probability of seeing a statistically significant difference when there is none by doing multiple statistical tests when you are examining a potential relationship between a given set of independent and dependent variables. In a situation with one independent variable and five dependent variables for a given sample with a level of significance for the individual one-way ANOVAs of 0.05, the chance of getting a statistically significant difference by chance is $(0.95)^5 = 0.23$. By running five separate one-way ANOVAs, there is a 23% chance that the tests will yield differences by chance rather than be real differences. This is too great of a chance. In statistics it is accepted practice to avoid running multiple tests on similar concepts with the same sample or samples. Running these relationships all at once reduces the likelihood that differences will arise by chance.

Another reason for using MANOVA is that sometimes a combination of dependent variables better represents a phenomenon than does a single dependent variable. For instance, in measuring overall physical fitness measures of aerobic fitness, flexibility, agility, and strength would better represent the concept of “total physical fitness” than any one of the variables alone. MANOVA would be used to look at how certain categorical variables might be related to overall physical fitness.

Application of MANOVA

Suppose an investigator chooses three random samples from a population, and one group will be fed a Dietary Approaches to Stop Hypertension (DASH) diet (this diet is an evidence-based approach to controlling hypertension). A second group is fed a high-fat, highly processed, animal-protein-based diet (American diet). Finally, a

Statistical test	Dependent variable	Independent variable	Covariate	Situation
One-way ANOVA	One quantitative variable	One categorical variable with ≥ 2 categories	No	Four treatments (LDL ^a -lowering drug, Mediterranean diet, vegan diet, or American diet) and serum LDL level.
Repeated Measures ANOVA	One quantitative variable measured two or more times under different conditions or at different times	One or more categorical variables	No	Sample of people measured for preprogram LDL level and after education program measured once a month over 8 mo to look for trends; each person in a sample eats three different plant-based burgers and rates them for taste.
Multiway ANOVA	One quantitative variable	Two or more categorical variables	No	Effect of sunlight (sunlight or no) and vitamin D-3 supplements (supplements or placebo) on changes in serum 25-hydroxycholecalciferol
Analysis of covariance	One quantitative variable	One or more categorical variables	One or more covariates	Pre-test and post-test diastolic blood pressure measured on one group who received hypertension control program and another that does not, differences in post-test diastolic blood pressure adjusted for differences in pre-test. Adjusting nutrition knowledge scores for grade point average after one group receives a nutrition course using problem-based learning and another, lecture/discussion.
Multiple ANOVA	Two or more quantitative variables	One or more categorical variables	No	One group Dietary Approaches to Stop Hypertension diet, another American diet, a third fed a vegan diet measured on multiple coronary risk-related physical parameters.
Multiple of covariance	Two or more quantitative variables	One or more categorical variables	One or more covariates	Same example as multiple ANOVA but three groups adjusted for differences in waist circumference.

Figure 4. Applications of different types of analysis of variance (ANOVA) to varying data situations. ^aLDL=low-density lipoprotein.

third group is fed a vegan diet. The groups will stay on a metabolic ward and be fed these diets in a test kitchen for 6 weeks. Systolic and diastolic blood pressure, serum LDL, serum high-density lipoprotein cholesterol, serum triglycerides, serum C-reactive protein, and blood clotting time will be measured after 6 weeks for all groups. This group of variables represents coronary heart disease risk. Rather than run separate one-way ANOVAs for each outcome/dependent variable a MANOVA can be run to examine this concept of coronary heart disease risk. Differences between groups on all outcome variables are tested simultaneously as a mathematical combination. In conjunction, correlations can be calculated between the different dependent variables to explore potential relationships.

The test statistics used for MANOVA include the Wilks λ , Hotelling-Lawley trace, and Roy largest root statistic. The most commonly used test statistic is Wilks λ . Wilks λ is the ratio of the unexplained variance in the combination of dependent variables to the sum of the explained and unexplained combination variable variance. Wilks λ ranges between 0 and 1, with a lower value representing greater likelihood of statistical significance. Because unexplained and explained variance is involved in the calculation of the test statistic, the Wilks λ value can be

mathematically converted to an F ratio if the user is more comfortable with this more common test statistic.

If the test statistic is significant then the difference between groups for the combination of dependent variables is statistically significant. In a post hoc manner individual one-way ANOVAs are run for each dependent variable to see what variables were affected most by the independent variable. To reduce the probability that differences will arise by chance due to multiple comparisons, a Bonferroni correction is used. The level of significance, which is typically 0.05, is divided by the number of dependent variables being evaluated with separate ANOVAs. If five dependent variables are in the combination variable then $0.05/5=0.01$ (0.01 is the significance level). This new conservative level of significance is used as the criteria for statistical significance of the F ratios for the individual ANOVAs.

For the example given above regarding diets and multiple dependent variables related to coronary risk, if the test statistic is significant, indicating that diet had a statistically significant effect on the combination of dependent variables (Wilk's λ closer to 0 than 1), procedures are then used examining individual ANOVAs for each dependent variable to see which was affected most. In other words, there might be a statistically significant

difference in serum LDL and serum C-reactive protein between the groups, but not for the other dependent variables. Because there are seven dependent variables, 0.05 is divided by seven for a level of significance of 0.007 for the individual ANOVAs. Because more than two groups are involved, then pairwise post hoc tests must be done to see which groups were different from one another on a dependent variable after the individual ANOVAs are run. So the DASH diet group is compared to the vegan group, the DASH group is compared to the American diet group, and so on for each dependent variable.

Brann and Skinner (8) used MANOVA to explore potential relationships between two independent variables, parental status (mother or father) and son's body mass index category (average vs high), and a combination variable of a variety of dependent variables, including scores representing parents' perceptions of their son's weight, child feeding practice scores, and parenting styles scores. This is a great example illustrating the steps in the MANOVA analysis procedures.

MANCOVA

In the diet and coronary risk example used above for the MANOVA section, suppose the DASH, American diet, and vegan diet groups were different in waist circumference, indicating that the groups have differing amounts of visceral fat. High visceral fat levels can affect most of the coronary risk variables listed previously. MANCOVA, an extension of ANCOVA and MANOVA, can adjust the mathematical combination of the dependent variables for differences in waist circumferences between groups. MANCOVA is used to test differences between groups for a mathematical combination of more than one outcome/dependent variable while adjusting this combination for the effects of quantitative or categorical variables (covariates). As with MANOVA test statistics such as Wilks λ are calculated. The test statistic is calculated once the mathematical combination of dependent variables is adjusted for the covariates (in our example, waist circumference). Then post hoc ANOVAs are examined for each adjusted dependent variable similarly to MANOVA. Again, if more than two groups are compared on the dependent variables subsequent pairwise post hoc tests will be applied to examine the nature of the differences. The levels of significance are handled the same way, with Bonferroni corrections.

Podar and colleagues (9) used MANCOVA to investigate changes in anthropometric measures in college students with lower vs higher intakes of low-fat dairy product intake during a school year. This study illustrates the introduction of covariates to adjust a variable representing a combination of dependent variables (anthropometric measures).

CONCLUSIONS

In scientific studies, investigators attempt to control as many variables as possible that could affect the relationship between their primary independent variables and dependent variables. Statistical tests are one way that control can be achieved. Many RDs are familiar with multiple regression as a way to control for the influential effects of other variables. Contributions of multiple independent variables to a dependent variable can be elucidated in multiple regression. However, ANOVA is also a useful statistical method that can help control for the effects of important and potentially influential study factors. At the most basic level, ANOVA can be used to examine the relationship between one independent variable and a dependent variable. In addition, ANOVA can be expanded to evaluate the relationship between multiple independent variables and multiple dependent variables, including adjustment for covariates (see Figure 4). ANOVA is an important tool in the repertoire of any RD conducting or interpreting research, as was demonstrated in examples of the application of ANOVA in dietetics-related contexts.

STATEMENT OF POTENTIAL CONFLICT OF INTEREST:
No potential conflict of interest was reported by the authors.

References

1. Table of critical values for the F distribution (for use with ANOVA). University of Sussex Web site. <http://www.sussex.ac.uk/Users/grahamb/RM1web/F-ratio%20table%202005.pdf>. Accessed June 21, 2011.
2. Smith WE, Day RS, Brown LB. Heritage retention and bean intake correlates to dietary fiber intakes in Hispanic mothers—*que sabrosa vida*. *J Am Diet Assoc*. 2005;105(3):404-411.
3. Bourque SP, Pate RR, Branch JD. Twelve weeks of endurance exercise training does not affect iron status measures in women. *J Am Diet Assoc*. 1997;97(10):1116-1121.
4. Hertzler SR, Clancy SM. Kefir improves lactose digestion and tolerance in adults with lactose maldigestion. *J Am Diet Assoc*. 2003;103(5):582-587.
5. Mendoza JA, Watson K, Cullen KW. Change in dietary energy density after implementation of the Texas Public School Nutrition Policy. *J Am Diet Assoc*. 2010;110(3):434-440.
6. Cottone E, Byrd-Bredbenner C. Knowledge and psychosocial effects of the film *Super Size Me* on young adults. *J Am Diet Assoc*. 2007;107(7):1197-1203.
7. Ventura EE, Davis JN, Alexander KE, et al. Dietary intake and the metabolic syndrome in overweight Latino children. *J Am Diet Assoc*. 2008;108(8):1355-1359.
8. Brann LS, Skinner JD. More controlling child-feeding practices are found among parents of boys with an average body mass index compared with parents of boys with a high body mass index. *J Am Diet Assoc*. 2005;105(9):1411-1416.
9. Podar KH, Hosig KW, Nickols-Richardson SM, Anderson ES, Herbert WG, Duncan SE. Low-fat dairy intake and body weight and composition changes in college students. *J Am Diet Assoc*. 2009;109(8):1433-1438.