



HHS Public Access

Author manuscript

Nurs Res. Author manuscript; available in PMC 2017 June 05.

Published in final edited form as:

Nurs Res. 2013 ; 62(1): 45–49. doi:10.1097/NNR.0b013e3182741948.

Composite Variables:

When and How

Mi-Kyung Song, PhD, RN,

Associate Professor, School of Nursing, University of North Carolina at Chapel Hill

Feng-Chang Lin, PhD,

Assistant Professor, School of Public Health, Department of Biostatistics, University of North Carolina at Chapel Hill

Sandra E. Ward, PhD, RN, FAAN, and

Professor Emerita, School of Nursing, University of Wisconsin-Madison

Jason P. Fine, ScD

Professor, School of Public Health, Department of Biostatistics, University of North Carolina at Chapel Hill

Abstract

Background—Use of composite variables is a common practice, but knowledge about what researchers should consider when creating composite variables is lacking.

Objective—The purpose of this paper was to present methods used to create composite variables with attention to advantages and disadvantages.

Methods—Methods of simple averaging, weighted averaging, and meaningful grouping to create composite variables are described briefly, and the context in which one method might be more suitable than the others is discussed. Study examples and comparisons of statistical power among these methods as well as Bonferroni correction are described.

Discussion—Each approach to creating composite variables has advantages and disadvantages that researchers should weigh carefully. With normally distributed data, composite variables provide the greatest increases in power when the original variables (that make up the composite variable) have similar associations with the outside outcome variable.

Keywords

composite variables; statistical power

A composite variable is a variable made up of two or more variables or measures that are highly related to one another conceptually or statistically (Ley, 1972). The individual

Corresponding author: Mi-Kyung Song, PhD, RN, 7460 Carrington Hall, School of Nursing, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 (songm@email.unc.edu).

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.nursingresearchonline.com).

The authors have no conflicts of interest to disclose.

variables making up a composite variable may be scales, single or global ratings, or categorical variables. Using composite variables is a common practice for controlling Type I error rate (e.g., when a sample size is not sufficient for testing multiple comparisons), addressing multicollinearity for regression analysis, or organizing multiple highly correlated variables into more digestible or meaningful information.

The consequences of combining related variables into a composite variable can include alterations of the relationship strength with outside variables (e.g., outcome variables), changes in statistical power, over-reduction or loss of information, and challenges in interpreting the composite variable itself or the relationships with outside variables. However, there has been little discussion about these issues. In this presentation, several commonly used methods to create composite variables are outlined, and actual studies and a numerical illustration are used to demonstrate the importance of these issues in the use of composite variables.

Approaches to Creating a Composite Variable

Creating a composite variable begins with critically analyzing the theoretically intended meaning of a concept so that the variables selected are to create the composite fully and to capture that meaning logically. In general, there are two ways to create composite variables, averaging and meaningful grouping.

Averaging

There are several ways to accomplish averaging. The specific method of averaging can be chosen based on whether the original variables intended to be combined (original variables in short) are continuous or categorical and also depend on the existing knowledge of the original variables. When original variables are continuous, simple averaging or weighted averaging can be considered. Simple averaging is the most commonly used approach to creating a composite variable. In this approach, the composite variable (symbolized as C) is created by summing z scores of the original variables: $C = z_1 + z_2 + \dots + z_p$, where $z = \frac{x - \bar{x}}{SD}$. The z scores have a mean of 0 with a range from negative to positive numbers. Such standardization is necessary when the original variables have different variances so that the association between the composite variable and outside variables will not be unduly affected by any one original variable with a large variance. To enhance interpretation, some researchers transform the composite to a standardized score—for example, a T score:

$T = \bar{X}' + C(SD')$, where \bar{X}' is a new desired mean (e.g., 50) and SD' is a desired standard deviation (e.g., 10; Streiner & Norman, 2003).

While converting the original variable scores to z scores preserves the distribution of the raw scores (z -score transformation does not guarantee a normal distribution), the contributions of the original variables are considered equal, as shown in the algebraic expression C above. When the purpose of creating a composite variable includes controlling Type I error rate for multiple comparisons with outside variables, simple averaging would be appropriate if the original variables are known to have similar relationships (the magnitude) to the outside

variables. The following study example illustrates the use of simple averaging to create a composite variable.

Ward et al. (2009) conducted a randomized controlled trial to test a patient-education intervention designed to reduce cancer pain. The investigators created a composite cancer pain severity score as an endpoint. The composite was made up of five variables that address pain related to cancer and its treatment during the past week: (a) the amount of time the patient had spent in moderate to severe pain (1 = *never* to 5 = *always*), (b) the intensity of usual pain (0 = *none* to 3 = *severe*), and (c) three intensity items (worst pain, least pain, and pain now), each rated 0 = *no pain* to 10 = *pain as bad as I can imagine* (Brief Pain Inventory; Daut, Cleeland, & Flanery, 1983). To create the composite, the investigators used simple averaging by summing *z* scores as described above. To enhance interpretation, they transformed the *z* scores to *T* scores with a range of 0–100; higher scores indicate greater pain severity.

With weighted averaging, weights are assigned to each of the original variables used to construct the composite variable. The weights may be determined from a prior study in which the composite variable has been used. Alternatively, the weights may be constructed using the data from a current study using principal components analysis (PCA). In this case, the weights may be applied to mean-corrected (or centralized scores; without being divided by the standard deviation) or standardized original variable scores to create the composite variable. As shown in the equation below, this technique is appropriate for constructing composite variables that are linear combinations of the original variables (Dillon & Goldstein, 1984; Sharma, 1996): $C = w_1X_1 + w_2X_2 + \dots + w_jX_j$, where w_j is the weight of the j th variable used in the composite variable. The weights, w_j , are chosen to maximize the ratio of the variance of C to the total variation (to account for as much variability in the original variables as possible), with the constraint, the squares of the weights sum to 1. The variability of the composite variable depends on the individual variances and the covariances of the original variables. The resulting composite variable corresponds to the latent dimension in the data that best summarizes the overall structure of the original variables. Other principal components (the technique yields more than one principal component) may be constructed to capture other dimensions in the data.

The weight assigned to a variable is affected by the relative variance of the original variable. Therefore, the choice between mean-corrected scores and standardized scores (for X_j in the equation above) has an effect on the analysis. If there is a compelling reason to believe that the variance of any of the original variables is more influential or important than the others, mean-corrected scores should be used (Sharma, 1996).

There are challenges in using PCA to create a composite variable. There may be different sets of weights obtained using PCA, each leading to a different composite variable. For example, with PCA based on five original variables, there are five possible sets of weights and five composite variables. Because these composite variables are orthogonal to each other, PCA may be used if the goal is to avoid multicollinearity in regression analysis. However, note that some of the composite variables will explain less variability than do the original variables. Also, care is needed if one wishes to conduct tests between each of the

multiple composite variables and the outside variable; to control Type I error rate, a correction for multiple testing would be required.

Principal components analysis is perhaps most useful when the original variables are correlated highly with each other and only a few components are needed to capture their overall structure, leading to a reduction in the number of variables. Nonetheless, the loss of information in discarding some principal components needs to be considered carefully in light of the scientific goals of the study.

An example of the use of PCA is a study by De Pauw et al. (2009). They evaluated relationships among personality traits, self-esteem, and psychopathology in 60 child psychiatric patients. Twelve constructs were assessed in two informants (parent and teacher). The Self-Perception Profile for Children was used to assess self-esteem for children aged between 8 and 12 years. This measure included six subscales (scholastic competence, social acceptance, athletic competence, physical appearance, behavioral conduct, and global self-esteem). Achenbach's System of Empirical-Based Assessment was used to assess adaptive and maladaptive functioning. The authors used two subscales, internalizing and externalizing problem behavior. Finally, the Strengths and Difficulties Questionnaires included four problem scales (emotional symptoms, conduct problems, hyperactivity, and peer problems) and one adaptive scale (prosocial behavior). These 13 scales were entered as variables in PCA to extract the underlying dimension between parent and teacher ratings. The first component (the composite variable) accounted for 72% of the variance. This composite variable was used to assess correlations with outside variables, such as disagreeableness and emotional instability.

Meaningful Grouping

Meaningful grouping is the nonstatistical combination of selected original variables based on the interpretation of the variables' values or scores, guided by the science of the field. Meaningful grouping can be used to create a composite outcome variable from multiple continuous or categorical variables or both. These original variables, when combined into a composite, can indicate an attribute (e.g., high risk for mortality) that is meaningful. A composite variable created by meaningful grouping is often categorical. For example, a composite variable may include categories improved, no change, and worse to indicate the direction of overall change from baseline, to determine whether or not an intervention was efficacious. The key point is that the composite should be meaningful with respect to the context and purpose of the study and should be determined based on the science of the field, with a predefined algorithm.

One of the most common contexts of using meaningful grouping involves clinical trials. Researchers may choose to create a composite variable by meaningful grouping if there is considerable difficulty in deciding which of multiple possible outcomes should be the primary outcome variable, particularly when any one of those possible outcomes is insufficient to represent the treatment effect (Freemantle, Calvert, Wood, Eastaugh, & Griffin, 2003). With respect to power, a major challenge is that the treatment effect represented by a composite outcome variable can be diluted if any one of the original variables is not affected by the treatment or is not sensitive to change. Furthermore, a

composite variable made up of multiple clinical indicators may be meaningful for testing efficacy in a trial, but applicability in clinical practice may be in question if any of the original variables is not readily obtainable in practice.

When power is a main concern, the meaningful grouping used to create a composite variable should be decided before the conduct of a study. The design of the study (e.g., sample size) should be adequate to test hypotheses of interest regarding this composite variable. The use of meaningful grouping after the conduct of a study based on the observed data does not control Type I error rate and is best viewed as exploratory and hypothesis generating for future studies.

In a randomized controlled trial (Song et al., 2009), meaningful grouping was used to create a composite outcome variable to examine the effect of an end-of-life communication intervention on patient–surrogate dyad congruence on goals of care and surrogates’ end-of-life decision-making confidence. Dyad congruence scores ranged from 0 (*incongruent in all three end-of-life scenarios*) to 2 (*congruent in all three scenarios*). Surrogate decision-making confidence was measured using a scale that ranged from 0 (*not confident at all*) to 4 (*very confident*). Because surrogates could be highly confident even if they misunderstood patients’ preferences for end-of-life care, the investigators created a composite variable to group dyads as follows: Dyads were categorized as *improved* if the dyad’s congruence improved from baseline or continued to be congruent and the surrogate had a decision-making confidence score of 3 of 4. Dyads were categorized as *not improved* if the dyad remained incongruent or became incongruent from baseline and the surrogate had a decision-making confidence score of 2. The numbers of improved and not improved dyads in intervention and control groups were compared.

Statistical Power: Simple Averaging, Weighted Averaging Versus Bonferroni Correction

If statistical power is of concern, the consideration of using a composite variable instead of Bonferroni correction may depend on the correlations among the original variables and on the association between the original variables and the outcome variables. This point is illustrated by comparing power under different correlation and association scenarios.

For simplicity, the original variables in X and the outcome variable Y are assumed to follow a multivariate normal distribution with correlation coefficients ρ_{XX} (correlation coefficient between original variables) and ρ_{XY} (correlation coefficient between an original variable and the outcome variable). Since mean and variance have no effect on power, arbitrary values

can be assigned. For a composite variable $C = \sum_{i=1}^p \omega_i Z_i$, where Z_i is the standardized original variables in $X = (x_1, x_2, x_3, \dots, x_j)$, the power can be computed by knowing the sample size n , the true association between C and Y , and Type I error rate α . The true association between C and Y (the effect size), referred to as δ_c , is the key element in the power calculation. The power calculation formula for C is

$1 - \beta_c = \Phi(\delta_c \sqrt{n} - Z_{\alpha/2}) + \Phi(\delta_c \sqrt{n} - Z_{\alpha/2})$, where Φ is the cumulative probability of the standard normal distribution with $Z_{\alpha/2}$ as its $(1 - \alpha/2)$ percentile. It can be seen that the effect size is determined by two kinds of quantities, ρ_{XX} and ρ_{XY} , with n and α , which are fixed.

When X and Y are highly associated, ρ_{XY} is large, the effect size is enhanced, and statistical

power increases. On the other hand, when colinearity exists between the original variables in X (i.e., ρ_{XX} is large), the effect size diminishes and statistical power decreases.

The power for the Bonferroni correction can be derived using the same ideas, but with Type I error α/p , reflecting that p individual tests are done. The combined power of the tests is related to the power of each test as well as their correlation. In general, if one assumes a priori that there is a common effect size δ , then it is sensible to assign equal Type I error probability to each test. The converse would be true if some effect sizes were believed a priori to be larger than others, in which case one might wish to give more Type I error probability to tests with larger effect sizes. The Bonferroni procedure gives equal weight to each test, regardless of the hypothesized effect sizes, and hence may be underpowered relative to a test based on the composite variable, depending on the nature of the associations in ρ_{XX} and ρ_{XY} .

Explicit formulae for composite score method and for Bonferroni correction are provided in Supplemental Digital Content 1, <http://links.lww.com/NRES/A87>.

Figures 1 and 2 illustrate the power comparisons under two kinds of associations between $X = (x_1, x_2, x_3)$ and Y , in which the method to create a composite variable is superior in one scenario whereas Bonferroni correction is favored in the other. In Figure 1, when each original variable (X) has the same association with Y , composite scores based on simple averaging and based on weighted averaging using PCA perform well when the original variables have relatively high correlation, for example, $\rho_{x_2x_2} = \rho_{x_2x_3} = \rho_{x_2x_3} = 0.9, 0.5$, as illustrated in (A) and (B). However, the PCA method performs poorly as shown in (C) when the original variables lack correlations or as shown in (D) when more than two principal components exist in the underlying correlation structure such as $\rho_{x_2x_2} = 0.9$ and $\rho_{x_2x_3} = \rho_{x_2x_3} = 0$. In contrast, when one original variable has a different association with Y (Figure 2), Bonferroni correction is more robust because the effect size in X_1 can be much larger than the other two original variables. Use of composite variables (both simple averaging and PCA scores) is underpowered, since the weight ω_j is not determined by ρ_{xy} . Superior power can be achieved by selecting weights that yield a composite variable with larger effect size for testing association between X and Y . The power formula given previously may be useful in such calculations. In theory, researchers may determine optimal weights that yield largest effect size δ_c .

In summary, with normally distributed data, composite variables provide the greatest increases in power when the original variables have similar associations with the outcome variable (Serlin & Mailloux, 1999). In general, PCA weighted averaging is not recommended when power is a key consideration in selecting a composite variable for the design of a study. However, because it can be used to identify hidden dimensions in the data, PCA may be helpful in exploratory analyses to identify composite variables using the observed data that have associations with the outcome variables. Such composite variables might be used in the design of future studies. With non-normal data, the simple power formulae presented above are not available. Simulation studies may be used to compute power curves such as those in Figures 1 and 2 in order to determine appropriate sample sizes for the study design.

Conclusions

Each approach to creating a composite variable has advantages and disadvantages that researchers should weigh when planning a study. The approach to creating composite variables should be determined during the planning stage and certainly before any testing with outside variables. Once a composite variable has been created and used in analyses, results involving the composite variable should be interpreted at the level of the composite variable, not at the level of the individual original variables. In manuscripts reporting studies that involve composite variables, investigators should describe the rationale for selecting a given approach and the exact method used to create the composite.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Dr. Ronald Serlin, Professor Emeritus, at the Department of Educational Psychology, University of Wisconsin-Madison, for his prior review and comments.

Funding source: NIH, R01NR011464, Song and R01NR013359, Song.

References

- Daut RL, Cleeland CS, Flanery RC. Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases. *Pain*. 1983; 17:197–210. [PubMed: 6646795]
- De Pauw SS, Mervielde I, De Clercq BJ, De Fruyt F, Tremmery S, Deboutte D. Personality symptoms and self-esteem as correlates of psychopathology in child psychiatric patients: Evaluating multiple informant data. *Child Psychiatry and Human Development*. 2009; 40:499–515. [PubMed: 19381800]
- Dillon, WR., Goldstein, M. *Multivariate analysis: Methods and applications*. New York, NY: John Wiley & Sons; 1984.
- Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C. Composite outcomes in randomized trials: Greater precision but with greater uncertainty? *JAMA*. 2003; 289:2554–2559. [PubMed: 12759327]
- Ley, P. *Quantitative aspects of psychological assessment*. London, UK: Gerald Duckworth & Co; 1972.
- Serlin, RC., Mailloux, M. An empirical comparison of three methods for performing univariate analyses with multivariate data. Presented at The American Educational Research Association Annual Meeting; Montreal, Canada. 1999.
- Sharma, S. *Applied multivariate techniques*. New York, NY: John Wiley & Sons; 1996.
- Song MK, Ward SE, Happ MB, Piraino B, Donovan HS, Shields AM, Connolly MC. Randomized controlled trial of SPIRIT: An effective approach to preparing African-American dialysis patients and families for end-of-life. *Research in Nursing & Health*. 2009; 32:260–273. [PubMed: 19205027]
- Streiner, DL., Norman, GR. *Health measurement scales: A practical guide to their development and use*. 3. Oxford, UK: Oxford University Press; 2003.
- Ward SE, Serlin RC, Donovan HS, Ameringer SW, Hughes S, Pe-Romashko K, Wang KK. A randomized trial of a representational intervention for cancer pain: Does targeting the dyad make a difference? *Health Psychology*. 2009; 28:588–597. [PubMed: 19751085]

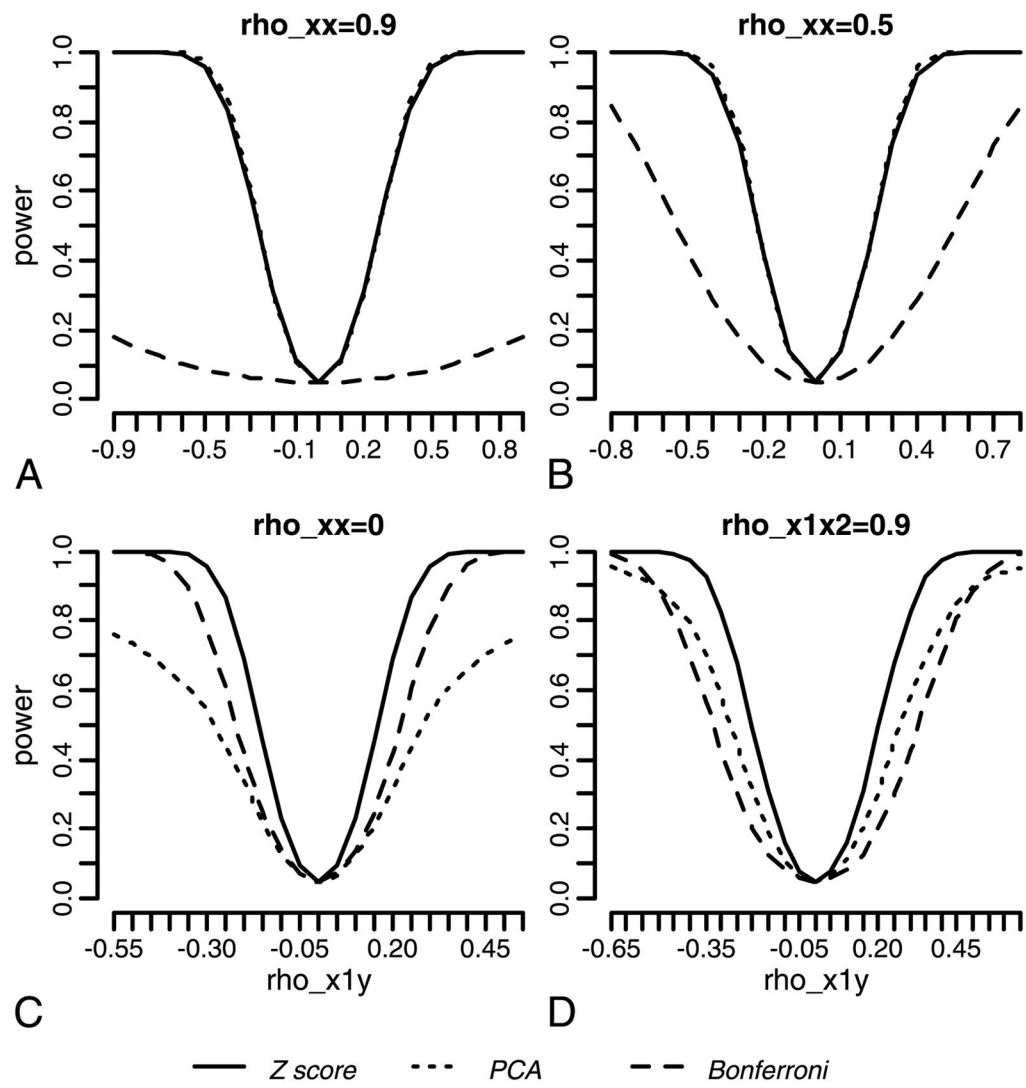


FIGURE 1. Power comparisons when $n = 50$, $a = 0.05$, and $\rho_{x1y} = \rho_{x2y} = \rho_{x3y}$

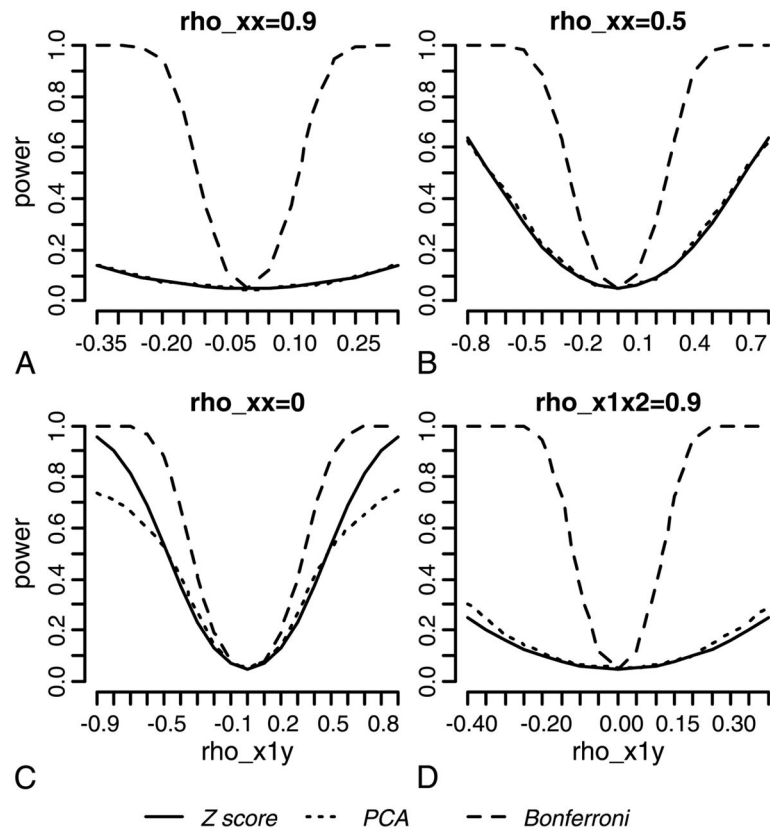


FIGURE 2. Power comparisons when $n = 50$, $\alpha = 0.05$, and $\rho_{x_2y} = \rho_{x_3y} = 0$