Published in final edited form as:
Struct Equ Modeling. 2014 April 3; 21(2): 318–324. doi:10.1080/10705511.2014.882696.

Examining Measure Correlations with Incomplete Data Sets

Tenko Raykov,
Michigan State University

Brooke C. Schneider,
VA Health Care System, West Los Angeles

George A. Marcoulides, and
University of California at Riverside

Peter A. Lichtenberg
Wayne State University

Abstract

A two-stage procedure for estimation and testing of observed measure correlations in the presence of missing data is discussed. The approach uses maximum likelihood for estimation and the false discovery rate concept for correlation testing. The method can be utilized in initial exploration oriented empirical studies with missing data, where it is of interest to estimate manifest variable interrelationship indexes and test hypotheses about their population values. The procedure is applicable also with violations of the underlying missing at random assumption, via inclusion of auxiliary variables. The outlined approach is illustrated with data from an aging research study.

Keywords

auxiliary variable; correlation; false discovery rate; incomplete data; missing at random

Examining Measure Correlations with Incomplete Data Sets

Missing data may be unavoidable in most current empirical studies in the behavioral, social, educational, and biomedical sciences as well as cognate disciplines. Frequently in these areas, scholars are interested in examining observed measure correlations. For instance, this may be the case in initial exploration oriented phases of their research, where it could be of concern to estimate these parameters and possibly test hypotheses about their population values, in particular for significance. A traditional approach has relied thereby on list-wise deletion and frequently on unadjusted significance-level multiple testing using the resulting correlation estimates.

That conventional procedure, which seems to be still often utilized in empirical research in these sciences, can be criticized on at least two counts. One, via list-wise deletion an investigator arrives at a subsample from an originally available one, which may not be

Correspondence on this article may be addressed to Tenko Raykov, Measurement and Quantitative Methods, Michigan State University, 443A Erickson Hall, East Lansing, MI 48824.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

representative of the studied population (unless the missing completely at random assumption is fulfilled, which one could expect rarely to be the case in applications unless induced by design; e.g., Allison, 2001; Graham, 2003). Two, subsequent multiple testing of correlation coefficients with unprotected significance level does not control the family-wise error rate (FWER), and has therefore attracted a great deal of criticism. Even if the widely used Bonferroni-procedure is employed then, the resulting testing approach can be overly conservative and thus miss potentially important findings about variable interrelationships (e.g., Johnson & Wichern, 2002).

This article discusses an alternative, two-stage procedure for examining observed measure correlations with incomplete data sets. The outlined method is based on the missing at random assumption, and its plausibility is enhanced by inclusion of auxiliary variables (e.g., Enders, 2010). When testing each measure correlation for significance (or equality to a non-zero pre-specified value) in the second step, the false discovery rate concept is utilized within an application of the Benjamini-Hochberg (BH) multiple testing approach. The BH method controls the FWER and is in general more powerful than conventionally employed multiple testing procedures (Wasserman, 2004).

The paper is structured as follows. The next section deals with the background, notation and assumptions underlying the remaining discussion. A factor analytic model is then adopted that can be straight-forwardly employed to estimate correlations in incomplete data sets. A following subsection is concerned with an application of the BH method for testing correlations. The data illustration section demonstrates subsequently the application of the outlined approach in an empirical study. The conclusion section deals with limitations and extensions of the outlined procedure.

Background, Notation, and Assumptions

Suppose k continuous measures are collected on a sample of n persons in a study of a population of interest ($k, n > 1$). Denote these observed variables by $y_1, \ldots, y_k$ and assume that a researcher is interested in evaluating their indexes of (linear) interrelationships. Let y $= (y_1, \ldots, y_k)'$ be the vector of measures (with priming denoting transposition in the rest of the article), and designate by $\Sigma = Corr(\mathrm{y}) = [\sigma_{ij}]$ their population correlation matrix ($i, j = 1, \ldots, k$). In behavioral, social, educational and biomedical studies, some values may well be missing in the resulting data set, that is, the pertinent $n \times k$ data matrix, denoted M, can be incomplete. In this setting, suppose a concern of an investigator is with (i) estimating each of the $k(k-1)/2$ elements of the matrix Σ and (ii) examining for each of its entries the hypothesis that its population counterpart equals 0 (or a given non-zero value; see below). These concerns can arise in early stages of empirical social and behavioral research, in particular when initial exploration of patterns of observed variable interrelationships is of interest (e.g., Tabachnick & Fidell, 2007).

In the remainder of this article, given the presence of missing data in M, use will be made of the popular full information maximum likelihood (FIML) method (e.g., Arbuckle, 1996) for the purpose of measure correlation estimation. This method is based on the assumption of data missing at random (MAR; e.g., Little & Rubin, 2002). The assumption is fulfilled when the probability of missingness is not related to the actually missing values, but may instead be related to the observed data (e.g., Enders, 2010). Since MAR critically involves data that are not available, it is not statistically testable, that is, there is no conclusive test (of a sufficient condition) that if ‘passed’ would imply that the data are MAR (cf. Raykov, 2011). Given the lack of testability of MAR, and the fact that frequently in empirical research missingness may well depend on the actually missing values (e.g., in cases of attrition in longitudinal studies), it is important to enhance the plausibility of this assumption. This is possible by the inclusion of ‘auxiliary variables’, i.e., variables that are ‘predictive’ of the missing values (e.g., Little & Rubin, 2002). In practice, based on substantive considerations these variables may be chosen from among measures that are related to dependent variables with missing values (cf. Enders, 2010.) The rest of this paper also assumes the availability of one or more auxiliary variables, but under MAR its method could also be applied without such variables; similarly, we assume normality of the measures for the first stage of the following procedure (e.g., Arbuckle, 1996; see also illustration and conclusion section).

Estimation and Testing of Measure Correlations in the Presence of Missing Data

Estimation of observed variable correlations

In the first step of the discussed method, for the *k* manifest variables we adopt the formal factor analysis model in Raykov & Marcoulides (2010; cf. Kühnel, 1988). In the model, each of *k* latent variables is defined as identical to a corresponding from the observed measures. That is, to accomplish observed correlation estimation with missing data, we make use of the following model:

$$y = a + \Lambda f + e, \quad (1)$$

where $a = (a_1, \ldots, a_k)'$ is a vector of intercepts, Λ is the $k \times k$ identity matrix, and f is a $k \times 1$ vector of dummy latent variables, while for the measurement errors $e = \underline{0}$ is posited (and thus $Cov(e) = O_k$, where O_k denotes the $k \times k$ covariance matrix consisting of zeros only, and $\underline{0}$ is the $k \times 1$ vector of zeros). Model (1) is saturated for any number *k* of observed variables, i.e., is associated with perfect fit when fitted to data (viz. with a chi-square value of 0, while having 0 degrees of freedom). Fitting model (1) to an incomplete data set employing FIML, with auxiliary variables as indicated earlier, permits estimation of the $k(k\text{-}1)/2$ correlations between the observed variables. This estimation is possible, for instance, using the popular latent variable modeling software Mplus (Muthén & Muthén, 2010; see next section for an empirical illustration and Appendix 1 for the source code needed). Thereby, estimates of these correlations are provided in the corresponding entries of the correlation matrix of the latent variables, as are pertinent standard errors and two-tailed *p*-values.

Testing measure correlations for significance

In the second step of the method in this paper, the set of $k(k\text{-}1)/2$ hypotheses of each population correlation being 0 (or equal to another pre-specified number; see below) is tested by employing the BH multiple testing procedure. This procedure is based on the concept of false discovery rate (FDR; e.g., Wasserman, 2004). The FDR notion is particularly attractive in the present setting, and can be seen as a highly useful complement to the traditionally utilized concept of FWER in multiple testing.

The FDR is defined as the expected ratio of hypothesis rejections that are incorrect relative to all rejected hypotheses from a given set of tested null hypotheses (e.g., Benjamini & Hochberg, 1995). In the context of examining measure correlations in an exploration oriented study - in part concerned with ascertaining which among a set of observed variable interrelationships exhibit discernible linear relationships - the FDR concept is especially appealing also for the following reason. When concerned with testing multiple correlations, one is not only interested in controlling the FWER, but also in limiting the rate of incorrect rejections of null hypotheses stipulating no linear relationships. As mentioned, the BH procedure controls the FWER, and in addition limits from above the rate of incorrect rejections by the used significance level (e.g., .05; Benjamini & Hochberg, 1995). Moreover, the BH procedure is in general associated with higher power than conventional multiple testing procedures, and this advantage increases with number of incorrect null hypotheses in the studied population (Benjamini & Yakutieli, 2001).

For the sake of completeness of this discussion we provide next a brief discussion of the BH procedure, which is applicable when there are m hypotheses to be tested on a given data set ($m > 1$; see Wasserman, 2004). Denote by $p_1, p_2, \ldots, p_m$ the p-values associated with these m hypotheses, and designate by $p_{(1)} \leq p_{(2)} \leq \ldots \leq p_{(m)}$ their ascending rank-ordering. Given a pre-specified significance level α (e.g., $\alpha = .05$), one defines the m ratios

$$l_j = j\alpha/[m(1+1/2+1/3+\ldots+1/m)] \quad (j=1,\ldots,m). \qquad (2)$$

Denote next by r the largest number j for which the inequality

$$p_{(j)} \leq l_j \qquad (3)$$

holds ($j = 1, \ldots, m$); if inequality (3) turns out not to be satisfied for any j, set $r = 0$ and $p_{(r)} = p_{(0)} = 0$. This p-value $p_{(r)}$ is called the BH rejection threshold and usually denoted T, that is, $T = p_{(r)}$ ($r \geq 0$). At the decision stage, the BH procedure consists in rejecting all null hypotheses with p-values that do not exceed $p_{(r)}$, i.e., all null hypotheses with p-values smaller than or equal to the threshold T; in case $r = 0$, $T = 0$ holds as well, and none of the tested p null hypotheses is rejected (Wasserman, 2004).

In an available incomplete data set with p observed measures, after fitting model (1) using FIML with auxiliary variables as mentioned earlier, one obtains with the used software the measure correlation estimates and (two-tailed) p-values associated with each of them. An application of the BH procedure on these $m = k(k-1)/2$ correlations allows one to test the m null hypotheses of them all being 0 in the studied population. If some correlation hypothesis stipulates equality to a given non-zero value, ρ_0 say, the (corrected) difference in chi-square values can be evaluated for the two corresponding nested models and its associated p-value used in the BH procedure ($-1 < \rho_0 < 1$; see also conclusion section). Then the ‘full’ model is the one defined in (1), and the nested in it model is (1) with the correlation in question being set equal to ρ_0.

We stress that no inflation of Type I error is associated with the outlined application of the BH procedure while all these $m = k(k-1)/2$ hypotheses are tested, unlike the case when testing them at an unprotected alpha level. It is worthwhile also stressing that the BH procedure has in general higher power than an alternative Bonferroni-adjustment for correlation testing that is widely used in empirical social and behavioral research. In fact, this disadvantage of the Bonferroni adjustment method – i.e., loss in power – increases with the number of non-zero population correlations (when tested are the null hypotheses of them being 0, or in general with the number of incorrect null hypotheses). This number of positive or negative true correlations may be expected to be considerable in many empirical behavioral and social studies. Particularly for these studies (as well as others), an application of the BH procedure instead of the Bonferroni method for overall significance level protection, will be associated with marked gains in power relative to the latter method.

We illustrate next the discussed procedure for examining observed measure correlations in incomplete data sets using data from an empirical investigation.

Illustration on Data

In this section, we employ data from an aging research study. Its aims were in part to examine the relationships between measures of general health and vascular health, age, three scales of physical activity (for balance, gait, and chair), and functional independence of n = 130 urban black adults (Schneider, 2011). These variables were drawn from the World Health Organization's International Classification of Functioning, Health and Disability (WHO ICF; World Health Organization, 2002), which considers relationships between them to be bidirectional (i.e., correlational) in nature. General health and vascular health measures related to number of problematic health conditions. Physical activity was measured by the Short Physical Performance Battery (SPPB; Guralnik, Simonsick, Ferrucci, Glynn, Berkman, & Blazer, 1994), an assessment of lower-extremity functioning in which participants perform three physical tasks (i.e., chair stands, 8-foot gait speed and three standing-balance poses; see also Guralnik, & Ferucci, 2000). Functional independence was measured using the Lawton and Brody (1971) IADL scale, a self-report measure of an older adult's level of independence in performing several daily living tasks (i.e., shopping, medication management, financial management, etc). For the illustrative purposes of this section, all measures are considered approximately continuous (see next section).

In the remainder of this section, we will be interested in first estimating the correlations between all k = 7 measures, i.e., in evaluating their 21 correlation coefficients, and then testing whether each of them equals 0 in the studied population of elderly. To this end, we commence with fitting the formal factor analysis model (1) to the data from these 7 variables (see Appendix 1 for source code needed). To enhance the plausibility of the underlying MAR assumption, we use thereby gender, education, Beck Anxiety Inventory (BAI; Beck and Steer, 1993) score, and the Mini Mental State Exam (MMSE) score as auxiliary variables with the following motivation. The MMSE is a 30-item screening measure of gross cognitive functioning (Folstein, Folstein, & McHugh, 1975). The BAI is a self-report measure, which asks participants to rate the presence of and perceived distress caused by symptoms of anxiety over the past week. As demonstrated in the aging literature, gender and education have broad effects on a host of factors in later life. Specifically, individuals with higher levels of education tend to have more preserved functional status into the later years, with this relationship frequently being moderated by health. The MMSE is one of the most widely used measures of gross cognitive functioning, and several studies have demonstrated its relationship to a multitude of factors including demographic, mood and health variables. In clinical settings, the MMSE is also frequently used in the decision making process in determining an older adult's ability to live independently. In addition, presence of clinically significant anxiety among older adults is related to declines in performance of daily tasks (Okura et al., 2010), and anxiety and disability share a number of risk factors including high disease burden and cognitive impairment. With all this in mind, it is plausible that information contained in these four auxiliary variables - gender, education, BAI and MMSE scores - is related to the missing values on the above seven measures of main interest.

As indicated earlier, the fitted model (1) is saturated, associated with perfect goodness-of-fit indexes (viz. chi-square value being 0, like its degrees of freedom), and used here merely as a means of correlation estimation (cf. Raykov & Marcoulides, 2010). Table 1 presents the resulting estimates of the 21 measure correlations and their associated *p*-values.

Taking a look at the entries in Table 1, and in particular the *p*-values in parentheses beneath them, we wish to emphasize the lack of stars attached to any of these *p*-values. This is a main feature in which the method in this paper differs from traditionally used approaches to multiple correlation testing (and their widely circulated software implementations). In particular, as stressed earlier, the present procedure is not based on comparisons of software reported *p*-values to a pre-selected significance level, such as say the conventional $\alpha = .05$. Further, the procedure does not utilize an adjusted level that is constant across all correlations, like the Bonferroni-adjusted significance level (which here would be .05/21 = .0024, but is not used in this section). In contrast to much of past empirical research in the behavioral, social, educational and biomedical disciplines, the method outlined in the preceding section involves instead a comparison of each *p*-value (see Table 1) to a specific ratio corresponding to that value. This ratio is defined in Equation (2), valid only for the *p*-value considered, and compared to the latter before judgment can be made whether that *p*-value is to be declared ‘significant’ or not.

Having completed the initial step of an application of the outlined procedure, we move on to its second step consisting of testing all 21 correlations for significance (equality to 0). To this end, we need to obtain first the 21 ratios l_j in Equation (2) ($j = 1, \ldots, 21$). They can be furnished with any of a number of widely available software, or even a hand-held calculator, and are rendered easily by the freely available software R (Venables, Smith, and the R Development Core Team, 2007). The resulting *l*-ratios from Equation (2) are presented in Table 2 (along with the rank-ordered *p*-values, accounting for ties).

According to the BH procedure, the highest *p*-value needs to be found next that does not exceed its corresponding *l*-value. By within-row comparison in Table 2, we determine that all *p*-values before the 15th are smaller than their corresponding *l*-values, and all *p*-values starting with the 15th are larger than their pertinent *l*-values. Hence, $r = 14$, and thus the BH threshold is $T = p_{(14)} = .007$. Therefore, we reject all null hypotheses associated with a *p*-value not larger than $T = .007$, i.e., all null hypotheses of zero population correlation for pairs of measures whose *p*-values do not exceed .007. (We note in passing that this threshold is higher than the Bonferroni-adjusted one of .0024 mentioned above, which is not unexpected since as indicated earlier the BH procedure is in general associated with higher power.)

The significant correlations found in this way are indicated in Table 3. The latter is identical to Table 1 up to the added symbol (‡) indicating a linear interrelationship index declared significant (not 0 in the studied population) with the procedure of this paper.

From Table 3, we see that the three physical activity scale scores (for balance, gait, and chair) are significantly interrelated among themselves, and each of them is so with instrumental activities of daily living as well as with age (negatively). These findings are expected on substantive grounds and demonstrate a discernible pattern of decrease of physical activity with age, as well as of decreasing/increasing capability of managing daily activities with decreasing/increasing physical activity. Further, this capability shows a discernible decreasing relationship with number of general health problems, while number of vascular health problems has an inverse discernible relationship with the physical ability scores and capability of managing instrumental daily activities.

Conclusion

This article addressed a frequent concern in behavioral, social, educational and biomedical studies, examination of observed measure correlations in incomplete data sets. This examination is often of interest in initial exploration oriented phases of empirical research, where one may be willing to evaluate the evidence in favor of discernible interrelationship patterns, in the presence of missing data that pervade social and behavioral investigations. Traditionally widely used procedures for these aims are suboptimal and involve list-wise deletion and unprotected significance level multiple testing, or such after correction employing the widely used Bonferroni method. Under the assumption of missing at random, the article discussed a more powerful approach to measure correlation exploration. This procedure does not have the limitation of analyzing a possibly non-representative subsample from an originally available one in order to deal with the missing data, as list-wise deletion will usually tend to yield. Similarly, while the present procedure controls the family-wise error rate (at $\alpha = .05$ say), it is in general more powerful than the Bonferroni-protected multiple testing approach or other conventional multiple testing procedures. This advantage increases with increasing number of non-zero correlations (incorrect null hypotheses) in a studied population, as could be expected to be the case in many social and behavioral studies.

The method in this paper is applicable also with violations of MAR, via use of auxiliary variables (e.g., illustration section; see also Enders, 2010, on their selection), and similarly when some null hypotheses stipulate correlation equality to pre-specified non-zero values.

Further, the method is readily extended to the case of nested (clustered) data that are not infrequent in the social and behavioral sciences. One possible way of handling that case is to employ robust methods that allow one to conduct correct inferences in case of clustering effect. [1]

Limitations of the outlined correlation examination procedure in this paper stem from its requirement for large samples. The reason is the fact that the underlying method of estimation is maximum likelihood that has optimal statistical properties with large samples. This requirement will be particularly important with a large number of observed variables. While no specific guidelines are at present available for determining appropriate sample size, due to the complexity of the issues involved, it may be conjectured that with more than say 10 observed variables it would be recommendable to have at least a few hundred subjects in analyzed incomplete data sets with limited fractions of missing information (Little & Rubin, 2002). One may also submit that with fairly large fractions of missing information the results of the method of this paper should be interpreted with a great deal of caution. We strongly encourage future research for developing guidelines for sample size in relation to number of observed variables and fractions of missing information. In addition to this large sample requirement, the presently described procedure is available for (approximately) continuous observed variables, which is a case of frequent substantial interest in empirical social and behavioral research. While the normality assumption is also made in the first stage of the procedure, it may be conjectured that the latter may be somewhat robust to mild violations of normality if use is made of the robust maximum likelihood estimation method (Muthén & Muthén, 2010; Savalei, 2010). We also encourage future research examining its robustness degree as well as related conditions.

Acknowledgments

Thanks are due to K.-H. Yuan, L. Muthén, and B. Muthén for valuable discussions on analysis of missing data, and to D. Paulson for helpful information on the data set used. The research was supported in part by the National Institute on Health/National Institute on Aging and Urban Health Pre-doctoral Research Training Grant (T32 AG00275-06).

Appendix: Mplus source code for estimation of measure correlations in the presence of missing data

```
TITLE:     ESTIMATION OF OBSERVED VARIABLE CORRELATIONS IN AN
           INCOMPLETE DATA SET.
DATA:      FILE = <NAME OF RAW DATA FILE>; ! USE UNIFORM MISSING
           VALUE SYMBOL(S)
VARIABLE:  NAMES = SPPB_BP SPPB_GP SPPB_CP AGE IADL GEN_HLTH
           VAS_HLTH GENDER EDUC BAI_TOT MMSE_IAD;
           MISSING = ALL(-999); ! UNIFORM MISSING VALUE SYMBOL -999
           AUXILIARY = (M) GENDER EDUC BAI_TOT MMSE_IAD;
ANALYSIS:  ESTIMATOR = MLR;
```

[1]At the software level, one can employ the same software as in this paper, Mplus, but with robust two-level estimation (e.g., Muthén & Muthén, 2010). An alternative would be to use the Stata command 'sem', requesting robust estimation (StataCorp, 2011).

```
MODEL:     SPPBP_L BY SPPB_BP@1; SPPB_BP@0;
           SPPGP_L BY SPPB_GP@1; SPPB_GP@0;
           SPPCP_L BY SPPB_CP@1; SPPB_CP@0;
           AGE_L BY AGE@1; AGE@0;
           IADL_L BY IADL@1; IADL@0;
           GH_L BY GEN_HLTH@1; GEN_HLTH@0;
           VH_L BY VAS_HLTH@1; VAS_HLTH@0;
OUTPUT:    STANDARDIZED TECH4;
```

Note. After the title for the analysis and naming the raw data file (with uniform symbol/s for missing values used throughout), names are assigned in the VARIABLE section, the missing value symbol indicated, auxiliary variables stated, and robust maximum likelihood (FIML) requested as an estimation method. The MODEL section introduces a latent variable formally identical to each observed variable (and thus parameterizes the measure correlations in those of the latent variables). The OUTPUT requests the standardized solution with standard errors and *p*-values for all measure correlations, and their correlation matrix (end of output; for an introduction to the syntax of Mplus, see e.g. Raykov & Marcoulides, 2006).

References

Allison, PD. Missing data. Thousand Oaks, CA: Sage; 2001.

Arbuckle, JL. Full information estimation in the presence of incomplete data. In: Marcoulides, GA.; Schumacker, RE., editors. Advanced structural equation modeling. Mahwah, NJ: Lawrence Erlbaum Associates; 1996. p. 243-277.

Beck, AT.; Steer, RA. Beck Anxiety Inventory. The Psychological Corporation; 1999.

Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. 1995; 57:289–300.Series B

Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. The Annals of Statistics. 2001; 29:1165–1188.

Enders, CK. Applied missing data analysis. New York: Guilford; 2010.

Folstein MF, Folstein SE, McHugh PR. Mini-Mental State. Journal of Psychiatric Research. 1975; 12:189–198. [PubMed: 1202204]

Graham JW. Adding missing-data relevant variables to FIML-based structural equation models. Structural Equation Modeling. 2003; 10:80–100.

Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman, Blazer DG. A short physical performance battery assessing lower extremity function: Association with self-reported disability, prediction of mortality and nursing home admission. Journal of Gerontology. 1994; 49:M85–M94. [PubMed: 8126356]

Guralnik JM, Ferrucci L. Lower extremity function and subsequent disability: Consistency across studies, predictive models and value of gait speed alone compared with the Short Physical Performance Battery. Journals of Gerontology. 2000; 55:M221–M231. [PubMed: 10811152] Series A

Johnson, RA.; Wichern, DW. Applied multivariate statistical analysis. Upper Saddle River, NJ: Prentice Hall; 2002.

Kühnel S. Testing MANOVA Designs with LISREL. Sociological Methods and Research. 1988; 16:504–523.

Little, RJA.; Rubin, DB. Statistical analysis with missing data. New York: Wiley; 2002.

Lawton MP. The functional assessment of elderly people. Journal of the American Geriatrics Society. 1971; 19:465–431. [PubMed: 5094650]

Muthén, LK.; Muthén, B. Mplus user's guide. Los Angeles: Muthén & Muthén; 2010.

Okura T, Plassman BL, Steffens DC, Llewellyn DJ, Potter GG, Langa KM. Prevalence of neuropsychiatric Symptoms and their association with functional limitations in older adults in the United States: The Aging, Demographics and Memory Study. Journal of the American Geriatrics Society. 2010; 58:330–337. [PubMed: 20374406]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Raykov T. On testability of missing data mechanisms in incomplete data sets. Structural Equation Modeling. 2011; 18:419–430.

Raykov, T.; Marcoulides, GA. A first course in structural equation modeling. Mahwah, NJ: Erlbaum; 2006.

Raykov T, Marcoulides GA. Group comparisons in the presence of missing data using latent variable modeling techniques. Structural Equation Modeling. 2010; 17:135–149.

Savalei V. Small sample statistics for incomplete non-normal data: Extensions of complete data formulae and a Monte Carlo comparison. Structural Equation Modeling. 2010; 17:241–260.

Schneider, BC. Ph D Dissertation Thesis. Institute of Gerontology, Wayne State University; Detroit, MI: 2011. Refining pathways to disability in urban black older adults: The roles of cognition, health, health behaviors and depression.

Sheikh, JI.; Yesavage, JA. Clinical Gerontology: A Guide to Assessment and Intervention. N: The Haworth Press; 1986.

StataCorp. Stata: Release 12 Statistical Software. College Station, TX: StataCorp LP; 2011.

Tabachnick, BS.; Fidell, LS. Using multivariate statistics. Boston, MA: Allyn & Bacon; 2007.

Venables, WN.; Smith, DM. The R Development Core Team. An introduction to R. Bristol, UK: Network Theory Limited; 2007.

World Health Organization. Towards a common language for functioning, disability and health. ICF; Geneva, Switzerland: 2002.

Table 1

Estimated correlations among *k* = 7 observed variables, with associated *p*-values in parentheses

Measure	1.	2.	3.	4.	5.	6.	7.
1. SPPBP	1						
2. SPPGP	0.548 (.000)	1					
3. SPPCP	0.432 (.000)	0.598 (.000)	1				
4. AGE	−0.337 (.000)	−0.392 (.000)	−0.269 (.003)	1			
5. IADL	0.536 (.000)	0.485 (.000)	0.485 (.000)	−0.096 (.264)	1		
6. GH	−0.196 (.038)	−0.117 (.185)	−0.204 (.027)	−0.069 (.452)	−0.354 (.000)	1	
7. VH	−0.275 (.001)	−0.233 (.007)	−0.280 (.001)	−0.017 (.867)	−0.277 (.000)	0.177 (.041)	1

Note. SPPBP = short physical activity scale for balance, SPPGP = short physical activity scale for gait, SPPCP = short physical activity subscale for chair, short physical activity scale for gait, IADL = instrumental activities of daily living, GH = general health, VH = vascular health.

Table 2

Ascending order of *p*-values in Table 1 and comparative values l_j (j = 1, …, 21; see Equations (2))

rank	*p*-value	comparative *l*-value
1	.000	0.000653146
2	.000	0.001306293
3	.000	0.001959439
4	.000	0.002612585
5	.000	0.003265731
6	.000	0.003918878
7	.000	0.004572024
8	.000	0.005225170
9	.000	0.005878316
10	.000	0.006531463
11	.001	0.007184609
12	.001	0.007837755
13	.003	0.008490901
14	.007	0.009144048
15	.027	0.009797194
16	.038	0.010450340
17	.041	0.011103490
18	.185	0.011756630
19	.264	0.012409780
20	.452	0.013062930
21	.867	0.013716070

Note. Compare each *p*-value in the left column with the *l*-value in the right within the same row, to find out the largest *p*-value that does not exceed its pertinent *l*-value. That *p*-value is the BH-threshold T (see main text), which here is T = .007.

Table 3

Rejected null hypotheses of zero population correlations among *k* = 7 observed variables (significance designated by ‡) used in the illustration section (see Table 2)

Measure	1.	2.	3.	4.	5.	6.	7.
1. SPPBP	1						
2. SPPGP	0.548 (.000‡)	1					
3. SPPCP	0.432 (.000‡)	0.598 (.000‡)	1				
4. AGE	−0.337 (.000‡)	−0.392 (.000‡)	−0.269 (.003‡)	1			
5. IADL	0.536 (.000‡)	0.485 (.000‡)	0.485 (.000‡)	−0.096 (.264)	1		
6. GH	−0.196 (.038)	−0.117 (.185)	−0.204 (.027)	−0.069 (.452)	−0.354 (.000‡)	1	
7. VH	−0.275 (.001‡)	−0.233 (.007‡)	−0.280 (.001‡)	−0.017 (.867)	−0.277 (.000‡)	0.177 (.041)	1

Note. Abbreviations in Table 2 used. (Tested are the null hypotheses of each correlation being 0 in the studied adult population.)