Published in final edited form as:

J Nutr Health Aging. 2009 March ; 13(3): 249–255.

ANALYTIC METHODS FOR FACTORS, DIMENSIONS AND ENDPOINTS IN CLINICAL TRIALS FOR ALZHEIMER'S DISEASE

R.E. TRACTENBERG

Departments of Neurology, Biostatistics, Bioinformatics & Biomathematics, and Psychiatry, Georgetown University School of Medicine, Washington, D.C., and the Center for Integrated Latent Variable Research, University of Maryland, College Park, MD a Collaborative for Research on Outcomes and -Metrics

Abstract

Alzheimer's disease (AD) is a complex disease process, so finding a single biomarker to track in clinical trials has proven difficult. This paper describes and contrasts statistical methods that might be used with biomarkers in clinical trials for AD, highlighting their differences, limitations and interpretations. The first method is traditional regression, within which one dependent variable, the Best Empirically Supported Indicator (BESI), must be identified. In this approach one biomarker (e.g., the ratio of tau to A β 42 from CSF) is the indicator for an individual's disease status, and change in that status. The second approach is an exploratory factor analysis (EFA) to consolidate a multitude of candidate dependent variables into a sample-dependent, mathematically-optimized smaller set of 'factors'. The third method is latent variable (LV) modeling of multiple indicators of an entity (e.g., "disease burden"). The LV approach can yield a complex 'dependent variable', the Best Measurement Model Indicator (BMMI). A measurement model represents an entity that several dependent variables reflect or measure, and so can include many 'dependent variables', and estimate their relative contributions to the underlying entity. The selection of a single BESI is an artifact of regression that limits the investigator's ability to utilize all relevant variables representing the entity of interest. EFA results in sample-specific combination of biomarkers that might not generalize to a new sample – and fit of the EFA results cannot be tested. Latent variable methods can be useful to construct powerful, efficient statistical models that optimally combine diverse biomarkers into a single, multidimensional dependent variable that can generalize across samples when they are theory-driven and not sample-dependent. This paper shows that EFA can work to uncover underlying structure, but that it does not always yield solutions that 'fit' the data. It is not recommended as a method to build BMMIs, which will be useful in establishing diagnostic criteria, creating and evaluating benchmarks, and monitoring progression in clinical trials.

Introduction

Alzheimer's disease (AD) is complicated by the variety, complexity, and subtlety of contributions from many neurophysiological effects of the disease process. The Alzheimer's Disease Neuroimaging Initiative (ADNI) is a multi-center, multi-year, multi-million dollar effort sponsored by a consortium in the United States and Canada of the National Institutes of Health and pharmaceutical companies dedicated to the longitudinal collection of candidate biomarkers (from CSF/plasma, including metabolites (e.g., tau, A β); imaging

Address for correspondence and reprint requests: Rochelle E. Tractenberg, Neurology, Suite 207 Building D, 4000 Reservoir Rd., NW, Washington, DC 20057, TEL: 202.687.2247, FAX: 202.784.3504, ret7@georgetown.edu.

Financial disclosure: None of the authors had any financial interest or support for this paper.

outcomes (e.g., volumes, metabolism, perfusion), and genes (e.g., ApoE status)) from 800 individuals (1). Biomarkers in AD are an area of intense interest, because prior to clinical symptoms, the disease is affecting the brain in key ways (2) and the search for treatments of AD is shifting towards effects earlier in the disease process (recent reviews in (1–3)).

In addition to serving as endpoints in clinical trials, biomarkers could be useful in identifying persons with AD but without clinical symptoms at the time of the intervention. One of the goals of the ADNI study is to identify the set of neuroimaging, biomarker, and clinical measures that can optimally identify persons with mild cognitive impairment (MCI) who will transition to AD, distinguish AD from persons with MCI and those whose cognitive aging is “normal”, and ideally to track relevant pathophysiologic changes over time. This paper examines methods for identifying and analyzing such markers.

The data collection by ADNI investigators includes outcomes from multiple imaging modalities, blood and CSF as well as neuropsychological tests, all of which represent “disease burden” or “neurodegeneration” (1) in some way. Given the complexity of the disease, and the variety in mechanisms represented by the outcomes ADNI is collecting, it is unlikely that a single test or biomarker will reflect burden or pathology in the sensitive, specific, and longitudinally robust way that would be required of an “ideal biomarker” for AD (4).

The search for one biomarker or test may be driven, in part, by the typical/traditional biostatistical method for clinical trials for AD, regression, which has advantages such as statistical control for covariates that could interfere with the estimation of relationships between the dependent and independent variables (5) and estimation algorithms, by easily accessible statistical software (as well as Excel, Microsoft Inc.), that are robust to departures from assumptions about the variables such as non-normality or heterogeneous variances. A key disadvantage of regression is the requirement that a single dependent variable be specified. The selected dependent variable, the Best Empirically Supported Indicator (BESI (6)), is the focus of the analysis plans and power calculations; other dependent variables can be selected for analysis but these must be designated as secondary (although see (7) for analytic techniques for families of outcomes and (8) for multi-model inference). In the BESI approach, one biomarker (e.g., the ratio of tau to A β 42 from CSF) would be used as the indicator for an individual’s disease status, and change in that status. To date (September 2008), identifying a BESI for AD or MCI (or MCI that will transition to AD) has proved difficult (2,9–10).

Rather than rely on a single BESI, investigators might be interested in combining multiple observed variables in a systematic way. Exploratory factor analysis (EFA) is a multivariate approach to consolidating a large number of variables into a smaller, possibly more manageable, set while explaining as much of the variability in the original set of variables as is possible (5,11–12). If the original set of variables (e.g., a collection of biomarkers) can be consolidated into a single ‘factor’ that explains a sufficient amount (e.g., >75% (13)) of the variability in the original set, then that new single factor could be utilized as the dependent variable (BESI) in a clinical trial. This approach has the advantage of combining the multiple biomarkers into a single composite that maintains a desired amount of explanatory power –in terms of variability- relative to the original biomarker set. EFA, typically accomplished by principal components extraction, is essentially the combination of a larger set of candidate dependent variables into a single outcome, in a mathematically-optimized way. Combining multiple outcomes of interest is an important aspect of EFA that makes it a more appealing method for analyzing and incorporating biomarkers into clinical trials in AD (as endpoints or for identifying participants) than identifying or choosing a single BESI from among the candidate biomarkers. However, in addition to being sample-dependent, EFA

methods are sensitive to the strengths of associations across variables (14). This particular feature of sample-dependence can lead to solutions that do not correspond to the actual dimensions of the variables being combined. This paper will demonstrate how EFA can sometimes ‘find’ the correct underlying structure, and how it can sometimes provide a misleadingly ‘good’ solution that does not, in fact, fit the data. In the discussion, alternative methods for combining biomarkers (and other outcomes) are described.

In his seminal article LL Thurstone (15) demonstrated the method of factor analysis by obtaining length, width, and height (x, y, z) measurements for 20 boxes. In the now famous “Box Problem”, the three correlated measurements were permuted in twenty nonlinear ways. Exploratory factor analysis (EFA) recovered the original three dimensions -the manipulated variables loaded on the dimension(s) from which they were obtained (i.e, x, y, and z). This early simulation demonstrated that factor analysis can recover ‘true’ factors.

EFA methods are typically associated with ‘soft’ outcomes that cannot be directly observed; in the present study, this Thurstone simulation was replicated with a similar analysis of correlated neuroimaging (‘hard’) outcomes from the ADNI data set. The method was applied to the permuted versions of four ADNI outcomes, and then applied to the original four outcomes, to explore the applicability of EFA in the context of AD biomarkers.

Methods

Overview

Four variables were selected from the set of ADNI variables downloaded in April 2008: hippocampal volume, temporal lobe PET values, CMRglucose uptake in frontal cortex normed to pons, and entorhinal cortical volume. “Right” and “left” side volumes were combined into an average $((R+L)/2)$. These four variables were permuted and subjected to EFA, attempting to replicate the Thurstone Box Problem (15) with neuroimaging outcomes. EFA was carried out on the original (unpermuted) four ADNI variables, and confirmatory factor analysis (CFA) was then carried out to examine the fit of the exploratory analysis.

Data

Data used in these analyses were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). For up-to-date information see www.adni-info.org. Four variables (described below) is the minimum required for identification in a confirmatory factor analysis (CFA), and since the EFA results were to be followed up by CFA, four ADNI variables from the baseline visit were selected. The ADNI neuroimaging variables were obtained from 137 participants across the ADNI study groups of persons with Alzheimer’s disease, persons with mild cognitive impairment, and persons with no diagnosis. The three groups were combined to maximize correlations (greater variability increases correlations), and because these analyses were for demonstration purposes only (and not to generate a factor score that could be used in future clinical trials for AD).

Statistical methods

For the permutations, the four variables were renamed: “w” = standardized hippocampal volumes $((UCSD) (R+L)/2)$, standardized with mean = 3304.244 and SD=539.77; “x” = average of mean temporal PET values (UCB; $(R+L)/2)$; “y” = average of CMR glucose uptake in frontal cortices normed to pons (UM), and “z” = average of entorhinal cortical volume $(UCSD; (R+L)/2)$. These ‘hard’ outcomes could all potentially be used together to show how AD is progressing, or to estimate disease severity at a given visit. Thus, theoretically they should be combinable into a single ‘factor’ representing disease burden,

neurodegeneration, or AD; each of these is a reasonable latent variable responsible for the correlations among these four (observed) neuroimaging variables; that is, all four variables reflect, in different ways/areas of the brain, that atrophy has occurred or the extent of disease burden. This reflection suggests that the four observed variables have a common cause, which would be the latent variable.

The variables were relabeled w, x, y, z and were subjected to nonlinear manipulations (15) resulting in 38 new variables that were nonlinear reflections of the original four:

$$w^2, x^2, y^2, z^2, wx, wy, wz, xy, xz, yz, \sqrt{w^2+x^2}, \sqrt{w^2+y^2}, \sqrt{w^2+z^2}, \sqrt{x^2+y^2}, \sqrt{x^2+z^2}, \sqrt{y^2+z^2}, 2w+2x, 2w+2y, 2w+2z, 2x+2y, 2x+2z, 2y+2z, e^w, e^x, e^y, e^z$$

Exploratory factor analysis with principal components extraction (PCFA) was applied (SPSS v. 16.0, SPSS Inc., Chicago) to the 38×38 matrix of correlations among these permutations. Factor loadings were estimated using varimax rotation, but the oblique rotation was also carried out to estimate inter-factor correlations. Following this replication of the Thurstone Box Problem, PCFA was then also used on the original four ADNI variables, in order to determine what EFA would ‘uncover’ about the interrelationships of the original four ADNI variables. In both cases, eigenvalues were used to determine the number of factors to extract (eigenvalues greater than one represent factors which account for more variance than any single variable did before the analysis).

The ‘appropriateness’ of the PCFA results for the four ADNI variables was then explored with CFA using EQS v. 6 (16) to determine whether the PCFA results fit the four ADNI variables. Five robust fit statistics were obtained for the solution: robust model fit (Satorra-Bentler χ^2 ; non-significant p-value suggests ‘good’ fit of model to data); robust Akaike’s Information Criterion (AIC; the lower the AIC, the better); robust comparative fit index (CFI; the closer to 1.0 the better, acceptable models have $CFI \geq .95$); standardized root mean square residuals (SRMR; the smaller (and < 0.09) the better); and robust root mean square error of approximation (RMSEA; the closer to zero (and positive) the better, acceptable upper bound on the 90% CI < 0.06) (17).

Results

To replicate the Thurstone Box Problem, PCFA was applied to the 38 nonlinear permutations. The results of this analysis are shown in Table 1 (bold indicates the variable (column) on which that permutation (row) was expected to load).

The pattern of bolded loadings in Table 1 support the conclusion that PCFA with varimax rotation recovered the original four ‘hard outcomes’ (W, X, Y, Z). The pattern shows that the highest loadings were estimated for those permutations (row) involving the ‘correct’ original variable (column). The four-factor solution explained 98.2% of the variability in these 38 variables, a near-complete recovery of the variability in the original 38 transformations, in spite of the fact that the permutations were nonlinear and ‘impure’ representations of the four ADNI variables. Factor loadings were obtained from this rotated solution (Table 1) for ease of interpretation of the factors.

In order to estimate the correlations among the factors obtained from the EFA, the four-factor solution was then obliquely rotated to estimate the factor correlations. These are shown (Table 2A) together with the observed correlations between the original four variables (Table 2B).

Tables 2A and 2B reflect similar strengths of associations (in absolute value terms) between the original variables and the factors, although correlations between the factor representing glucose uptake in the frontal lobes (factor “Y”) and the other factors were all negative, while all correlations among the original ADNI measurements were all positive. This is an artifact of the principal components extraction and does not suggest that this replication of the Thurstone Box Problem was unsuccessful.

Because PCFA was able to ‘correctly’ extract the original four dimensions from the 38 nonlinear permutations, the method was next applied to the original four values. As was noted above, a single factor solution would generate a BESI that integrates these four disparate neuroimaging variables into a single indicator that might be useful in clinical trials. PCFA did find a single factor, suggesting that these four ADNI variables might be considered to reflect different, measurable aspects of one latent variable. This one-factor solution explained 59.6% of the variance in the system of four neuroimaging outcomes. The one-factor solution was then subjected to a CFA to determine the fit of the one-factor solution to the data. The fit statistics are shown in Table 3.

Table three shows the five fit statistics (four of which are robust to deviations from normality and other assumptions in the data; SRMR has no robust counterpart Hu and Bentler, 1999) describing the degree to which a single factor explains the variability observed in the four neuroimaging variables. All fit statistics support the conclusion that a single factor does not fit these four variables: Robust model fit: Satorra-Bentler χ^2 : 31.23 (6df, $p < 0.001$), robust Akaike’s Information Criterion (AIC): 27.23 for model vs. 149.1 for independence model; robust comparative fit index (CFI): 0.812; standardized root mean square residuals (SRMR): 0.116; and robust root mean square error of approximation (RMSEA): 0.328 (90%CI: .232, .432). All fit statistics were well outside acceptable ranges (17), although the robust model AIC was an improvement over the independence model. In spite of explaining nearly 60% of the variability in the four ADNI variables, a one-factor model is a poor fit to the data.

Discussion

These results show that EFA can work to recover “latent” dimensions or constructs in collections of variables –replicating the “Thurstone Box Problem” with the permuted versions of the neuroimaging outcomes from the ADNI data set – but they do not always work. Factor analytic methods often rely on principal components (PC) extraction, when the number of variables are not pre-specified (11,18); PC extraction seeks the model that maximizes variance explained (in terms of eigenvalues or variance accounted for) (see 14). When the number of factors is pre-specified and the factor loadings are estimated (as in maximum likelihood factor analysis, MLFA, see 11 and 18), the analysis estimates correlations between observed and latent variables that maximize the likelihood of the number of factors the investigator pre-specifies (11). MLFA does not have the same data-driven disadvantage that EFA does, but cannot accommodate chains of factors or higher order structure in the factors. The default setting for “factor analysis” in nearly all software that offers this function is principal components extraction, with MLFA being more difficult to obtain and interpret.

The ADNI variables chosen for the analysis were selected because they should be combinable into a measurement model representing disease burden, atrophy, or ‘AD’. One factor was extracted from the original variables that were manipulated into the nonlinear functions/combination; this single factor explained nearly 60% of the variance in the four variables. In spite of explaining so much of the variability in the four variables from different imaging modalities and areas of the brain, the fit of the hypothesized one-factor

model suggested very poor fit. This example was chosen to highlight the weaknesses of exploratory factor analysis in the combination of biomarkers for clinical trials in AD: using EFA is not recommended as a method for obtaining a BMMI for clinical research.

This study replicated the “Thurstone box problem” where nonlinear and highly correlated variables were analyzed with EFA and recovered the underlying ‘factors’. This suggests that EFA can uncover structure, even when the variables are hard endpoints that are correlated, such as the four ADNI variables whose permutations were analyzed. However, the second EFA analysis of the original four variables demonstrated that EFA does not always uncover the correct structure.

Rather than using EFA (MLFA or PCFA), a theoretically-driven measurement model (MM) would efficiently combine the information from across the variables representing AD, burden or atrophy from brain and activity (PET) or atrophy (volumes) – together with theory – to create a single outcome that would not be subject to the same disadvantages as the mathematically-optimized, atheoretical ‘model’ that might be obtained otherwise (i.e., from EFA or other multivariate techniques). A combination of theoretical and statistical insight should be applied to build a measurement model for CFA rather than using EFA.

Exploratory and confirmatory methods provide different information; the theoretical and statistical fits of any latent variable model must be established and EFA cannot provide these, but other methods, such as CFA, can. The critical point is, however, that building and testing measurement models is complex. Burnham and Anderson (2002) (8) note that a set of candidate models is important for choosing (or combining) models for inference (p. 2), and by extension, for selecting the ‘best’ measurement model as indicator for clinical trials. This means that a BMMI cannot simply be obtained by finding a single model that fits the data; the best MMI will be a model that not only fits the data but also fits better than reasonable alternative models (19,20).

This paper described some of the advantages and disadvantages of seeking/using a BESI in regression given the complexity of AD, and of EFA as a mechanism for circumventing the undesirability of one BESI in clinical trials for diseases as complex as AD. A third analytic method is available: explicit latent variable modeling such as CFA. This is a class of multivariate analytic methods that could lead to a single-variable representation of a larger set of observed variables, or ‘indicators’. Unlike EFA, the investigator builds, and then tests the fit of, a model specifying hypotheses about how the larger set of variables represents a hypothesized underlying, unobserved, entity. The model is referred to as a ‘measurement model’ (see 21). In the context of AD and finding biomarkers that are optimized for differentiating patients along the continuum of neuropathology, a set of biomarkers would be identified and hypothesized to represent one or more latent variables in specific ways. For example, within the ADNI data, ‘neurodegeneration’ might be a hypothesized clinical entity that causes volumetric imaging outcomes as well as levels of tau, to covary with low levels of glucose metabolism. In this example, variables reflecting amyloid (e.g., PIB uptake or A 42) would be expected to covary as a direct function of (i.e., caused by) neurodegeneration. This is not to imply that levels of all of these biomarkers are unrelated; only that the specific clinical entity “neurodegeneration” is hypothesized to cause decreases in volumes, tau, and lower glucose metabolism. Thus, unlike EFA, measurement models combine variables in hypothesis-driven ways– and can therefore be more generalizable across samples. When the best model has been built, tested and validated (i.e., replicated in an independent sample), it would be the “best measurement model indicator” (BMMI (6)), which itself could then be used as a BESI for regression in clinical trials.

Similar to EFA, the associations of the observed variables with the underlying entity can be estimated using maximum likelihood methods (21) and the most straightforward modeling

will follow from relationships that are linear (although nonlinear relationships can be modeled with latent variable modeling techniques; see (22) for a variety of complex latent variable methods and techniques). Unlike EFA and MLFA, causal chains and higher-order latent variable models can be hypothesized and tested in LV methods other than EFA/MLFA.

Latent variable (LV) methods simultaneously model multiple indicators of an entity (e.g., “disease burden”) by regressing observed variables on the hypothesized unobserved, underlying, or ‘latent’ one(s). It is recommended that searches for biomarker measurement models focus on causal models, where the unobserved clinical entity is the hypothesized cause of the covariances among the observed variables (biomarkers) (see (21) for discussion of causal vs. emergent latent variables). Importantly, in a causal model, the extent to which the causal factor (latent clinical entity) does not cause the variability in any observed variable is explicitly modeled and estimated – as ‘measurement error’. This is an important aspect of a latent variable causal model, since in standard linear/multiple regression, the residuals in the model represent the error with which the independent variable(s) represent or predict the dependent variable plus the error with which the dependent variable represents whatever it is supposed to represent. Within a LV model, the latter source of error is modeled explicitly, so that the former type of error can be estimated. This is not the case in EFA (or MLFA) models, nor is it a feature of any composite-forming method of multivariate analysis.

Once a causal model is hypothesized and the relations between the observed variable and the latent cause are estimated using specialized software (EQS, SAS, SPSS/AMOS, MPlus, R), the fit of the model to the observed data is estimated. This is roughly equivalent to obtaining the R² for a regression model, and in some software the R² is computed for the regression of each observed variable onto the latent variable(s). However, in addition to these indicator-specific summaries, many indices of overall fit of the model to the data are computed – such as those represented in Table 3, including areas of particular misfit (e.g., if the hypothesized relationship between the latent cause and one indicator is unsupported by the data).

The LV approach can yield a complex ‘dependent variable’, and given adequate fit of the measurement model to the data – as well as better fit than reasonable alternatives - this new dependent variable can be considered the ‘Best Measurement Model Indicator’ (BMMI). As described above, a causal measurement model represents a latent entity that several observed variables reflect or measure. The BMMI can accommodate all indicators of the underlying entity (e.g., “neurodegeneration”), and so can incorporate multiple ‘dependent variables’ into a single dependent ‘model’ variable, as well as estimating the relative contributions of the latent factor to each indicator. This is in contrast to regression based techniques (including those underlying model averaging and other combinations of regressions) where independent variables must (by assumption) be independent (orthogonal) to one another. Thus, it is inappropriate to include correlated variables as independent variables within linear regression, whereas a measurement model approach takes advantage of the correlations among variables.

The selection of a single-variable BESI is an artifact of regression that limits the investigator’s ability to utilize all relevant variables representing the entity of interest. EFA and other data-driven, atheoretical multivariate methods result in sample-specific single (composite)-variable combinations of biomarkers that might not generalize to a new sample and can sometimes uncover the correct structure, but not always. By contrast, the BMMI approach is a theory- and hypothesis-driven simultaneous analysis of multiple ‘dependent variables’ which are indicators of the underlying clinical entity. It requires extra work, but its accommodation of multiple and correlated variables, and its explicit modeling of error,

make this extra modeling effort worthwhile. This is particularly true in cases where, as in AD research, previous research has shown that no single variable can serve as the best biomarker.

As Box said, “all models are wrong, but some models are useful” (23). The assumptions, implications, and penalties for building and testing a BMMI are similar to those common to regression, multivariable systems, and measurement modeling found in any multivariate statistical textbook (more technical (5); more accessible: (24)). The main disadvantage of using BMMIs is that they must be built, tested and validated (see, e.g., (20,25)), which is far more time consuming than selecting a BESI, using EFA, or creating a composite or index. However, combination-of-models methods also require this attention to alternative models (8).

A measurement model is derived from the fact that it conceptualizes a construct, such as “neuropathology”, that can be measured in several different ways, all of which are subject to some type and extent of measurement error, and all of which are of interest/important to a complete appreciation and neuropathology in AD, MCI, and normal cognitive aging. The measurement model will not be “true” or “correct”, but it represents the optimal combination of theory and statistics. Thus, the Best Measurement Model Indicator (BMMI (6)) will ideally articulate a unidimensional (latent) construct to be measured, which in the current example could be “neuropathology”. “Neuropathology” can ONLY be estimated, it can never be directly (or completely) quantified or observed; moreover, its estimation/quantification will be optimized by increasing the number and quality of indicators that are hypothesized to be caused by it (26). By incorporating uncertainty and permitting multiple indicators, a measurement model improves estimation of the ‘truth’ or ‘true level’ of that construct in which we are most interested.

The BESI does not permit the simultaneous analysis of these indicators, but only the combination (through argument) of results from multiple regressions (on BESIs). For this reason, quite apart from the failures of any genetic, imaging, or biologic measure to attain the sensitivity and specificity that a biomarker for AD requires (1,4), building, testing and validating a BMMI is recommended over choosing a BESI from among the collection of ADNI measures.

Latent variable methods have a natural place in biomedical research. The results presented here show that using data-driven methods such as exploratory factor analysis will not necessarily uncover the ‘true’ relations among a set of biomarkers. Instead, a BMMI represents a combination of theory and statistical support for that theory, taking advantage of all relevant indicators-even if they are correlated, and importantly, the fit of model to data can be quantified, and replicated in new samples. A BMMI will explicitly model measurement error of indicators and, if a causal BMMI is built and validated, then a single target for any intervention would be identified.

Acknowledgments

This research was supported by NIH K01 AG027172 to RET; the ADNI data collection was supported by U01 AG024904-01.

References

1. Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack CR, Jagust W, Trojanowski JQ, Toga AW, Beckett L. Ways toward an early diagnosis in Alzheimer’s disease: The Alzheimer’s Disease Neuroimaging Initiative (ADNI). *Alzheimers Dement*. 2005; 1(1):55–66. [PubMed: 17476317]

2. Shaw LM, Korecka M, Clark CM, Lee VMY, Trojanowski JQ. Biomarkers of neurodegeneration for diagnosis and monitoring therapeutics. *Nature Reviews Drug Discovery*. 2007; 6:295–303.
3. Skovronsky DM, Lee VM, Trojanowski JQ. Neurodegenerative diseases: new concepts of pathogenesis and their therapeutic implications. *Annual Review of Pathology*. 2006; 1:151–70.
4. The Ronald and Nancy Reagan Research Institute of the Alzheimer's Association and The National Institute on Aging Working Group. Consensus report of the working group on molecular and biochemical markers of Alzheimer's disease. *Neurobiology of Aging*. 1998; 19(2):109–116. [PubMed: 9558143]
5. Johnson, RA.; Wichern, DW. *Applied multivariate statistical analysis*, 6E. New Jersey: Prentice Hall; 2007.
6. Tractenberg RE. Beyond factor analysis: Analytic methods for factors, dimensions, and endpoints in clinical trials for Alzheimer's disease. First Conference on Clinical Trials for Alzheimer's disease, Montpellier, France. *Journal of Nutrition, Health and Aging*. 2008 September.12(8) [section 3].
7. Dmitrienko A, Tamhane AC. Gatekeeping procedures with clinical trial applications. *Pharm Stat*. 2007; 6(3):171–80. [PubMed: 17583553]
8. Burnham, KP.; Anderson, DR. *Model selection and multimodel inference: A practical information-theoretic approach*, 2E. New York, NY: Springer-Verlag; 2002.
9. Thal LJ, Kantarci K, Reiman EM, Klunk WE, Weiner MW, Zetterberg H, Galasko D, Praticò D, Griffin S, Schenk D, Siemers E. The role of biomarkers in clinical trials for Alzheimer disease. *Alzheimer Dis Assoc Disord*. 2006; 20(1):6–15. [PubMed: 16493230]
10. Ridha BH, Anderson VM, Barnes J, Boyes RG, Price SL, Rossor MN, Whitwell JL, Jenkins L, Black RS, Grundman M, Fox NC. Volumetric MRI and cognitive measures in Alzheimer disease: comparison of markers of progression. *J Neurol*. 2008; 255:567–574. [PubMed: 18274807]
11. Bartholomew, DJ.; Knott, M. *Latent Variable Models and Factor Analysis*, 2E. London, UK: Arnold; 1999.
12. Afifi, A.; Clark, VA.; May, S. *Computer-Aided Multivariate Analysis*, 4E. Boca Raton: Chapman & Hall/CRC; 2004.
13. Stevens, JP. *Applied multivariate statistics for the social sciences*, 4E. Mahwah, NJ: Lawrence Erlbaum Associates; 2001.
14. Wickens, TD. *The Geometry of Multivariate Statistics*. Mahwah, NJ: Lawrence Earlbaum; 1994.
15. Thurstone LL. Current issues in factor analysis. *Psychological Bulletin*. 1940; 37(4):189–236. See pp 223–227 for box problem setup.
16. Bentler, PM.; Wu, E. *EQS structural equations program manual*. Encino, CA: Multivariate Software Inc; 1995.
17. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6(1):1–55.
18. Widaman, KF. Common factors versus components: Principals and principles, errors and misconceptions. In: Cudeck, R.; MacCallum, RC., editors. *Factor Analysis at 100: historical developments and future directions*. Mahwah, NJ: Lawrence Erlbaum; 2004. p. 177-204.
19. Jöreskog, KG. Testing structural equation models. In: Bollen, KA.; Long, JS., editors. *Testing Structural Equation Models*. Thousand Oaks, CA: Sage; 1993. p. 294-316.
20. Tractenberg RE, Aisen PS, Hancock GR, Rebeck GW. cross-sectional influence of APOE ε4 dose on clinical tests in Alzheimer's disease. *Neurobiology of Aging*. 200710.1016/j.neurobiolaging.2007.11.006
21. Bollen, KA. *Structural equations with latent variables*. New York, NY: Wiley Interscience; 1989.
22. Hancock, GR.; Samuelsen, KM., editors. *Advances in Latent Variable Mixture Models*. Charlotte, NC: Information Age Publishing; 2007.
23. Box, GEP.; Draper, NR. *Empirical Model-Building and Response Surfaces*. New York, NY: Wiley; 1987.
24. Kline, RB. *Principles and Practice of Structural Equation Modeling*, 2E. New York, NY: The Guilford Press; 2005.

25. Tractenberg RE, Aisen PS, Weiner MF, Cummings JL, Hancock GR. Independent contributions of neural and 'higher order' deficits to symptoms in Alzheimer's disease: a latent variable modeling approach. *Alzheimer's and Dementia*. 2006; 2(4):303–313.
26. Nunnally, JC.; Bernstein, IH. *Psychometric Theory*, 3E. McGraw-Hill; 1994.

Table 1

Exploratory factor analysis (EFA) with principal components extraction (varimax rotation) of 38 variables representing permutations of four ADNI neuroimaging outcomes

	Component			
	W	Z	X	Y
w2	.964	.125	.057	.037
x2	.249	.194	.922	.216
y2	.243	.178	.459	.835
z2	.154	.984	.076	.043
wx	.932	.181	.276	.114
wy	.940	.181	.159	.228
wz	.950	.275	.078	.073
xy	.262	.210	.773	.536
xz	.253	.613	.727	.172
yz	.258	.662	.365	.600
SQRT(w2+x2)	.889	.215	.371	.113
SQRT(w2+y2)	.886	.205	.214	.337
SQRT(w2+z2)	.564	.814	.090	.059
SQRT(x2+y2)	.271	.224	.744	.566
SQRT(x2+z2)	.198	.939	.268	.086
SQRT(y2+z2)	.201	.934	.176	.235
2w+2x	.909	.222	.320	.118
2w+2y	.914	.215	.184	.277
2w+2z	.822	.550	.094	.081
2x+2y	.268	.228	.764	.539
2x+2z	.229	.840	.476	.121
2y+2z	.232	.852	.257	.392
log w	.783	.250	.081	.082
log x	.230	.237	.920	.156
log y	.255	.223	.449	.817
log z	.166	.978	.085	.055
wxy	.881	.171	.336	.255
wxz	.913	.258	.267	.110
xyz	.269	.465	.687	.480
SQRT(w2+x2+y2)	.808	.226	.422	.325
SQRT(w2+x2+z2)	.562	.786	.231	.086
SQRT(x2+y2+z2)	.228	.881	.331	.247
SQRT(w2+x2+y2+z2)	.555	.750	.282	.211
e**w	.957	.112	.054	.025
e**x	.250	.188	.919	.224
e**y	.236	.166	.461	.834
e**z	.137	.977	.064	.030

Table 2A

Correlations estimated between the four factors recovered in the PCFA on 38 permuted variables

	W	X	Y	Z
W	1			
X	0.378	1		
Y	-0.395	-0.601	1	
Z	0.425	0.326	-0.367	1

Table 2B

Correlations observed between the four original ADNI variables from which the 38 permuted variables were derived

	Standardized hippocampal volumes ((R+L)/2) (W)	Average of mean temporal PET values ((R+L)/2) (X)	Average of CMR glucose uptake in frontal cortices normed to pons (Y)	Average of entorhinal cortical volume ((R+L)/2) (Z)
Standardized hippocampal volumes	1.0			
Average of mean temporal PET values	0.384	1.0		
Average of CMR glucose uptake in frontal cortices normed to pons	0.388	0.562	1.0	
Average of entorhinal cortical volume	0.672	0.385	0.341	1.0

Table 3

Fit of the one-factor model (CFA) to the four ADNI neuroimaging variables

Fit Criteria				
Satorra-Bentler 2 (df, p)	AIC	CFI	SRMR	RMSEA (90% CI)
31.23 (6, p<0.001)	27.23 (independence=149.1)	.812	.116	.328 (.232, .432)

Fit criteria: Satorra-Bentler χ^2 : general robust model fit statistic, with the associated p-value for the degrees of freedom shown. Non-significant p-value suggests 'good' fit of model to data. AIC: robust Akaike's Information Criterion; the lower, the better. CFI: Robust Comparative fit index; the closer to 1.0 the better, acceptable models have CFI $\geq$.95. SRMR: standardized root mean square residuals, the smaller (and <0.09) the better. RMSEA: Robust root mean square error of approximation; the closer to zero (and positive) the better, acceptable upper bound on the 90% CI <0.06.